Published in final edited form as:

Aphasiology. 2012 May 1; 26(5): 689–708. doi:10.1080/02687038.2012.676869.

Generalization of a Novel, Implicit Treatment for Coarse Coding Deficit in Right Hemisphere Brain Damage: A Single Subject Experiment

Connie A. Tompkins^a, Victoria L. Scharp^a, Kimberly Meigh^a, Margaret Lehman Blake^b, and Julie Wambaugh^b

^aCommunication Science & Disorders, University of Pittsburgh, Pittsburgh, PA, USA

^bCommunication Sciences & Disorders, University of Houston, Houston, TX, USA

^cCommunication Sciences & Disorders, University of Utah, and VA Salt Lake City Healthcare Systems, Salt Lake City, UT, USA

Abstract

Background—This manuscript reports generalization effects of Contextual Constraint Treatment for an adult with right hemisphere brain damage (RHD). Contextual Constraint Treatment is designed to stimulate inefficient language comprehension processes implicitly, by providing linguistic context to prime, or constrain, the intended interpretations of treatment stimuli. The study participant had a coarse coding deficit, defined as delayed mental activation of particularly distant semantic features of words (e.g., *rotten* as a feature of “apple”). Treatment effects were expected to generalize to auditory comprehension of narrative discourse, and perhaps to figurative language interpretation, because coarse coding has been hypothesized and/or demonstrated to support these abilities.

Aims—This treatment study aimed to induce generalization of Contextual Constraint Treatment in an adult with RHD with inefficient coarse coding.

Methods & Procedures—The participant in this study was a 75 year old man with RHD and a coarse coding deficit. A single subject experimental design across behaviors (stimulus lists) was used to document performance in baseline, treatment, and follow-up phases. Treatment consisted of providing brief, spoken context sentences to prestimulate, or constrain, intended interpretations of stimulus items. The participant made no explicit associations or metalinguistic judgments about the constraint sentences or stimulus words; rather, these contexts served only as implicit primes. Probe tasks were adapted from prior work on coarse coding in RHD. The dependent measure was the percentage of responses that met predetermined response time criteria. There were two levels of contextual constraint, Strong and Moderate. Treatment for each item began with the provision of the Strong constraint context, to minimize the production or reinforcement of erroneous or exceedingly slow responses. Generalization was assessed to a well-standardized measure of narrative discourse comprehension and to several metalinguistic tasks of figurative language interpretation.

Outcomes & Results—Treatment-contingent gains, associated with respectable effect sizes, were evident after a brief period of treatment on one stimulus list. Generalization occurred to untrained items, suggesting that the treatment was facilitating the underlying coarse coding process. Most importantly, generalization was evident to narrative comprehension performance,

for both overall accuracy and accuracy answering questions about implied information, and all of these gains maintained through three follow-up sessions.

Conclusions—Though the results are still preliminary, this single-subject experimental design documents the potential for meaningful gains from a novel treatment that implicitly targets an underlying language comprehension process in an adult with RHD.

Keywords

language comprehension; language comprehension treatment; language therapy; right hemisphere; brain damage; coarse coding

Right hemisphere brain damage (RHD) in adults can cause deficits in language comprehension processes that ultimately impair the understanding of various aspects of discourse (Tompkins, 2008; Tompkins, Klepousniotou & Gibbs Scott, 2012). One such deficit affects the processing of meanings and features of words that are especially semantically remote (Klepousniotou & Baum, 2005a; Tompkins, Fassbinder, Scharp, & Meigh, 2008). For example, participants with RHD and without brain damage primed equally well for subordinate features of words like “apple” that were not particularly distant from a typical instantiation of that word (e.g., *crunchy*), but for more remote subordinate features (e.g., *rotten*), the RHD group was impaired (Tompkins et al., 2008). This deficit reflects difficulty with a normal comprehension process that has been dubbed “coarse semantic coding” (e.g., Beeman, 1998; aka coarse coding). In traditional models of language comprehension, coarse coding occurs in an early phase, before the full context is taken into account (e.g., Gernsbacher, 1990; Kintsch, 1988, 1998). As such, coarse coding generates mental activation of wide-ranging aspects of word meaning, some of which may conflict with the broader context. In a later phase of comprehension, activation is dampened for concepts that are superfluous or irrelevant to the fuller context (e.g., Gernsbacher, 1990; Kintsch, 1988, 1998).

Recommendations for treating coarse coding deficits typically involve metalinguistic tasks, such as asking patients to generate alternative meanings or features of words or phrases. Such recommendations raise several concerns. First, because this area of research and clinical practice is so new, these recommendations are as yet untested. In addition, adults with RHD have well-documented difficulties with metalinguistic tasks (e.g., Tompkins, 1990, 1991a; Tompkins, Boada, & McGarry, 1992; Tompkins et al., 2004; Vanhalle, Lemieux, Joubert, Goulet, Ska, & Joannette, 2000). As a result, such tasks may confound the attempt to improve targeted underlying language processes. Finally, treatment that targets the generation of alternative meanings may be erroneously applied with individuals for whom it would be counterproductive. There is no test yet for use in clinical practice to distinguish a coarse coding deficit from a deficit in another crucial comprehension process, suppression (Gernsbacher, 1990; Gernsbacher & Faust, 1991). In normal comprehension, the suppression mechanism acts during the later phase of comprehension noted above, and is essential for reducing mental activation of concepts that are eventually determined to be contextually superfluous or incompatible. To exemplify, consider a narrative context like the following: “They were walking in an orchard. There was a lot of old, mushy fruit on the ground. He stepped on an apple.” Upon encountering the word “apple,” individuals with a suppression deficit would mentally activate various features, including both *crunchy* and *rotten*, and would be abnormally slow to inhibit the concept of crunchiness. Many adults with RHD have no difficulty activating multiple meanings or features of words (Klepousniotou & Baum, 2005a, b; Tompkins, Baumgaertner, Lehman, & Fassbinder, 2000), or alternate inferences (Tompkins, Lehman-Blake, Baumgaertner, & Fassbinder, 2001; Tompkins, Fassbinder, Blake, Baumgaertner, & Jayaram, 2004). Rather, these

individuals have difficulty suppressing meanings, features, or inferences that are incompatible with or irrelevant to the broader context (Klepousniotou & Baum, 2005a; Tompkins et al., 2000; 2001; 2004). Thus, treatment tasks that require the generation of alternatives would do nothing for the suppression deficit, and could potentially exacerbate it. Further complicating this situation, adults with RHD can have co-occurring deficits in coarse coding and suppression that would manifest as (a) difficulty activating aspects of lexical meaning that are particularly remote from those words' dominant interpretations, together with (b) difficulty dampening activation of less-distant concepts that have been mentally activated.

These complexities notwithstanding, there are several reasons that coarse coding is an important treatment target for adults with RHD. First, the efficiency of coarse coding predicts aspects of narrative comprehension. For example, individuals with RHD who were most impaired in the priming study of coarse semantic coding (Tompkins et al., 2008) were poorer comprehenders of implied information in discourse than other adults with RHD, even after accounting for factors like vocabulary knowledge and working memory capacity for language (Tompkins, Scharp, Meigh, & Fassbinder, 2008). Coarse coding also has been hypothesized to support the comprehension of figurative language, including phrasal metaphors (e.g., Beeman, 1998; for a summary, see Tompkins et al., 2012). As such it is possible that improving the efficiency of coarse coding will yield generalized gains in a range of communicative outcomes.

This manuscript reports early results of a novel language processing treatment that targets language comprehension processes in adults with RHD. A prior paper (Tompkins, Blake, Wambaugh, & Meigh, 2011) presented Phase I data that demonstrated promising improvements on the treatment tasks themselves. The current paper reports evidence of generalization of a refined version of the treatment, to a standard measure of narrative discourse comprehension, in a new study participant with a coarse coding deficit.

The treatment, Contextual Constraint Treatment, has two versions, one each for coarse coding and suppression processes. It is a novel treatment in several ways. First, rather than focusing on particular language structures or forms like metaphors, the treatment targets underlying language comprehension processes that should have some generalized effects on comprehension. The treatment is also unique in avoiding explicit, metalinguistic task demands which otherwise might confound treatment of the target process itself. Rather, the treatment targets the comprehension process of interest implicitly, via contextual prestimulation. The treatment was designed this way because adults with RHD who have difficulty with metalinguistic tasks can perform well on assessments of the same processing operations that are implicit (or nearly so; see summary in Tompkins et al., 2012). In addition, the crucial impairments for adults with RHD are overwhelmingly in the *speed* of coarse coding and suppression processes, not in their accuracy, so explicit training should not be necessary to initiate these processes.

Method

Participant

Mr. R is a 75-year-old, community-dwelling African-American male with 10 years of formal education, who worked as an operating engineer until he retired of his own volition. He had a thromboembolic right hemisphere cerebrovascular accident (CVA) 6 years prior to his participation in this study. His CT scan showed infarcts in the posterior limb of the internal capsule and in the centrum semiovale area of the parietal lobe. According to his medical chart, he reported at the time of his CVA that "I've been having trouble with my speech, losing my balance, and weakness in my left leg and arm." Acutely, he was

diagnosed with dysphagia, hemispatial neglect, and left hemiparesis. Nowadays he is proud to state "There is nothing wrong with my brain." At the time of testing for this study, he had no overt signs of dysphagia or neglect but the hemiparesis remained, and dysarthria was clearly evident.

By self-report, Mr. R. is a right-handed, monolingual, native speaker of American English without prior neurological or psychiatric conditions or learning disabilities. He screens as free of depression risk on the Center for Epidemiologic Studies – Depression Scale (CES-D; Radloff, 1977). He has corrected normal vision, more than adequate to read aloud the labels "Yes" and "No" on the response box used in diagnosis and treatment. He also has normal hearing in his better ear, as evidenced by a pure tone average (at 500, 1000, and 2000 Hz) of 20 dB and 96% on the Northwestern University Test #6 measure of speech discrimination (Tillman & Carhart, 1966).

Table 1 provides current data for Mr. R. on various clinical measures of language and cognition, obtained in the baseline sessions at the initiation of this study. For purposes of comparison, the table also provides means and standard deviations for the RHD participants in our study of coarse coding and discourse comprehension (Tompkins, Scharp, et al., 2008). Mr. R. was about 1 standard deviation older than that participant group, and had about 1 standard deviation fewer years of formal education.

Diagnosis

The diagnostic assessment of coarse coding is derived from our original experimental priming task (Tompkins, Fassbinder, et al., 2008). The assessment includes 15 critical stimuli in which a 1–2 syllable unambiguous noun (e.g., "apple") is followed by a target word that represents a particularly distant semantic feature of that noun (e.g., *rotten*). The noun itself is embedded in a short, semantically neutral sentence (e.g., "He has an apple"). Soon (175 ms) after the offset of the sentence-final noun, a spoken phoneme string is presented for timed lexical decision. The participant presses a button on a response box (Yes or No) to indicate as quickly as possible whether the phoneme string is a real word. Filler stimuli have nonword targets (e.g., */bImu/*), targets that are less semantically distant from their corresponding noun (e.g., "There is the mustard" - *spicy*), or targets that are unrelated to the sentence or its final noun (e.g., "He has a cabin" - *teammate*).

The diagnostic items are divided into two blocks of trials for administration. All stimuli were presented at 65dB SPL via an Anchor Audio AN 130BK speaker, as verified by a reading taken at the better ear using a Digital-Display Sound Level Meter (2103667). Stimuli are delivered via E-Prime 2.0 software (Schneider, Eschman, & Zuccolotto, 2002) and a Dell Latitude E6400 laptop computer. Manual responses are made on the E-Prime serial response box, and E-Prime records both response accuracy and speed.

A secondary word monitoring task is incorporated in the diagnostic assessment, to keep participants attending to the semantic content of the sentences. Each block of diagnostic trials has a single word, for which subjects are instructed to listen, and that word either does or does not occur in the block of trials. When it is present, it is a sentence-final noun from the last 4–5 trials in the block. Just prior to beginning the diagnostic block, the participant is told what word to listen for, and a printed reminder is placed directly on the response box, where it remains for the duration of the block of trials. Immediately after the participant's response to the final diagnostic item, the examiner asks "Did you hear the word X?"

In addition, a response deadline is signaled periodically on filler stimuli, to remind participants to continue to respond quickly. This deadline is indicated by a standard Microsoft Windows bell sound.

The criterion for diagnosis of a coarse coding deficit is impaired lexical decision performance on more than half of the critical trials (i.e., at least 8 items). Both accuracy and reaction time (RT) data are considered in the diagnostic process. RT data are judged with reference to criteria established for each individual stimulus, based on the median RTs of the group of non-brain-damaged participants (NBD; $N = 38$) from our original studies of coarse coding (Tompkins, Fassbinder, et al., 2008). To account for well-documented differences in simple reaction times between RHD and NBD groups, each median RT was multiplied by a correction factor of 1.5, which represents the proportional group difference that we have identified in previous studies. Mr. R was inaccurate on 5 of the 15 critical trials, and his RTs were more than 2 standard deviations below the item-specific criterion on 7 additional trials. Thus he met our established criterion for coarse coding deficit.

Experimental Design and Outcome Measures

This study employed a single-subject, multiple baseline design, across behaviors (Probe Lists). Lexical decision performance was probed repeatedly for two lists of sentence stimuli similar to those in the diagnostic task (Lists 1 and 2), and for one list of generalization stimuli (List 3), in 3 design phases: pre-treatment baseline, treatment, and follow-up. At follow-up, we also re-administered a block of 5 trials from the diagnostic task that are not included in the treatment lists (Coarse Coding Generalization task), to evaluate the extent to which any post-treatment improvements in performance were item-specific or more generalized.

In addition, performance was evaluated periodically on a number of broader outcome measures, including several assessments of language comprehension. The primary outcomes of interest are performance on the auditory version of the Discourse Comprehension Test (DCT; Brookshire & Nicholas, 1993). The DCT is a well-standardized test of the understanding of main ideas and details, as directly stated or implied in 10 narrative stimuli. The narratives, each 14 sentences long, describe “humorous situations that would be familiar to most adults in America” (Brookshire & Nicholas, 1993, p. 6). As audiorecorded for a prior study, the narratives average 63.2 seconds in duration.

The DCT controls for numerous important influences on comprehension and has strong psychometric properties, including test-retest and standard error data. Because coarse coding skill has been associated with performance on the DCT (Tompkins, Scharp, et al., 2008), we were particularly interested in DCT-based measures to assess generalization of treatment effects. The DCT-based measures for this purpose were *Total accuracy* on all comprehension questions (both stated and implied main ideas and details), and *Accuracy on questions about implied information*.

Coarse coding has been hypothesized to underpin figurative language processing as well (e.g., Beeman, 1998; Jung-Beeman, 2005). To assess potential generalization of treatment to figurative language interpretation, we administered the *Metaphor interpretation* and *Speech act interpretation* subtests of the experimental English-language version of the Protocole Montreal d’Evaluation de la Communication (MEC; Joannette, Ska, & Côté, 2004). The MEC is a comprehensive test of communication difficulties that typify those experienced by adults with RHD, and its authors graciously supplied the English-language subtests for our use in this study. In both subtests, the participant is asked to explain the meaning of a given sentence or short text. The Metaphor subtest assesses familiar, non-lexical metaphors (e.g., “The teacher is a sleeping pill”), while the Speech act subtest assesses direct and indirect statements or requests (e.g., indirect: “Nancy is very busy at work. She calls her husband and says: ‘Tonight, I won’t have time to pick the kids up at school.’ What do you think Nancy means?”). In both subtests, the examiner reads the stimulus to the participant, and if the participant’s answer is incorrect, the examiner offers several response options (3 for the

Metaphor subtest; 2 for the Speech act subtest). Because these tasks are metalinguistic in nature, we did have some question as to whether they would capture improvements in the underlying comprehension process of coarse coding. However, there are no implicit tasks of figurative language processing available at this time.

Finally, the Emotional Prosody Production subtest of the experimental English MEC was also administered periodically, as a control measure. Because coarse coding comprehension processes should be unrelated to the production of emotional prosody, we expected performance on this measure to be unaffected by the treatment. In this task, the participant hears a recording of four emotionally-neutral sentences (e.g., “Sally is eating bread”) intoned by a male speaker to convey three emotions (Happy, Sad, Angry). The participant is asked to “repeat each sentence with the exact same voice and intonation.” One replay is allowed if the participant requests it.

The first author, a native American English speaker who was not the primary examiner and therefore was not present on the days the task was administered, judged the participant’s responses. The audio or video recordings from each session were randomized for judgment. For each trial, the judge listened to the original stimulus as presented to the participant and scored the participant’s response into one of four categories: (a) neutral or wrong emotion conveyed; (b) conveys intended emotion on less than 50% of the utterance; (c) conveys intended emotion on 50% or more of the utterance; (d) conveys intended emotion on (essentially) 100% of the utterance. “Accurate” responses were defined as those in the latter two categories. A second naïve judge scored the same recordings, for inter-rater reliability purposes. Although this second judge categorized 25% more items as “accurate” than the first judge, on average, the rank order of sessions by total accuracy was the same for both judges.

Probe List Stimuli and Task

General Description—The coarse coding version of the Contextual Constraint Treatment has two lists of 15 critical probe stimuli (Lists 1 and 2), similar to those in the diagnostic task and designed to be entered into treatment, and a third list of 15 stimuli (List 3) to assess generalization of treatment results to the processing of figurative meanings of lexical metaphors. Between lists, stimuli are balanced for (a) the rated semantic fit, or semantic distance, between sentence-final nouns and their critical targets; and (b) the mean RTs of the target words, as determined by e-lexicon (<http://elexicon.wustl.edu/>; List 1, 2, and 3 $M = 613, 618, \text{ and } 595$ ms, respectively). In all probe sessions, the appropriate list(s) of stimuli were presented in the auditory implicit priming task used for diagnostics. Because all targets are real words, all critical probe stimuli require a ‘Yes’ response. Thus, each list also includes 10 filler stimuli with nonword targets that require a ‘No’ response. Filler items come from the same general categories as the experimental stimuli (clothing, transportation, places, food, animals), to help further disguise the experimental stimuli.

Lists 1 and 2—As in the diagnostic task, each probe stimulus for Lists 1 and 2 consists of a 3–4 word, spoken, semantically-neutral sentence that ends with a 1–3 syllable, concrete, common noun (e.g., “There was a piano”), followed by a spoken target word (e.g., *song*; see Appendix A for all probe stimuli). The critical stimuli were developed after many rounds of pilot testing to establish dominant and subordinate features/meanings of the nouns, and the rated semantic fit between each noun and its target word (See Appendix B for summary). The critical targets for these lists are words that represent semantically-remote subordinate features of their corresponding nouns (e.g., *song* for “piano”). Many of the critical items on these lists come directly from the diagnostic task.

List 3—List 3 stimuli are designed to assess generalization of treatment effects to the processing of lexical metaphors. Homophones whose subordinate senses are metaphoric were placed in the sentence final-position of neutral sentence frames (e.g., “There was a jewel”). The target for each homophone represents the subordinate, metaphoric sense of the homophone (e.g., *helpful*). Because we predict that Contextual Constraint Treatment for coarse coding will improve activation of semantically distant aspects of words, and because some evidence documents the processing similarities of literal and nonliteral language (Coulson & Van Petten, 2002; Grady, Oakley, & Coulson, 1999; Pynte, Besson, Robichon, & Poli, 1996), treatment gains on Lists 1 and/or 2 may generalize to List 3. If generalization does not occur, this would be consistent with the claim that adults with RHD can have a metaphor-specific processing deficit (Klepousniotou & Baum, 2005a) that is separate from, but could co-occur with, a coarse coding or suppression deficit.

Probe List Composition—Three presentation orders were developed for each Probe List. Across presentation orders, the 15 experimental and 10 filler trials from a single probe list are distributed so that each individual trial occurs in a different third of each list order (e.g., Item 1 in Order 1 is in the bottom third of Order 2 and the middle third of Order 3). Within these parameters, no two experimental trials are repeated in the same sequence in any of the orders for a single list. The other controls on list/order construction were: (a) each list starts and ends with filler trials to allow for warm-up and cool-down effects; (b) no more than 3 experimental trials occur sequentially in any list/order; (c) at least 3 items intervene between any two trials that could form a coherent story (e.g., could be construed as continuing a prior theme, or as having the same protagonist); and (d) successive stimuli are arranged so that an immediately preceding trial does not prime a theme or a word in the subsequent experimental trial, and so that there is a change in the protagonist (He, She, It).

Probe Task Dependent Variable

The outcome measure for the probe tasks is the percentage of accurate lexical decision responses to experimental probe stimuli that meet a preset RT criterion (%Crit). The criterion was established with reference to data from the 38 NBD control participants in our initial work on RHD and coarse coding (Tompkins, Fassbinder, et al., 2008), and represents a value that is essentially 1 standard deviation slower than those participants’ performance. Specifically, this value was calculated by: (1) identifying the mean RTs achieved by these 38 participants for each of the experimental items in the initial study, i.e., the same items as those in the diagnostic task; (2) taking the median of those mean RTs; and (3) adjusting this median for each Probe List, based on differences in lexical properties between the experimental target words in the diagnostic task and the experimental targets on each of the three lists. This correction factor was determined by consulting e-lexicon (<http://elexicon.wustl.edu/>) for the baseline RT of each experimental target word. This RT-based measure was chosen because deficient performance of adults with RHD on the lexical decision task has been overwhelmingly accurate, but delayed.

Probe Schedule

Probe sessions during the baseline and treatment phases were conducted 2–3 times a week. Follow-up probes were administered 2 weeks, 3 weeks, and 1 month after the termination of treatment.

Treatment

General Description—The goal of Contextual Constraint Treatment is to facilitate language comprehension processes (coarse coding or suppression) that often are impaired after RHD in adults, via repeated stimulation of the neural networks that support these

comprehension processes. Facilitation is operationalized as increasing the speed of response, because data from our past experiments suggest that most treatment participants will be highly accurate but abnormally delayed on the treatment tasks. For participants whose performance accuracy is less than 80%, the treatment initially targets accuracy until it reaches 80% or above, with a subsequent focus on response speed.

The crux of the treatment is to supply contextual bias prior to each probe stimulus, to pre-stimulate for each sentence final noun (e.g., “piano”) the semantic feature that is represented by the target word (e.g., *song*). We chose this approach because contextual bias is well-documented to benefit language comprehension in adults with RHD (Blake, 2009; Blake & Lesniewicz, 2005; Lehman-Blake & Tompkins, 2001; Tompkins, 1991b). When each acquisition stimulus is first introduced into treatment, it is preceded by two sentences that form a Strong biasing context, which is expected to prompt a quick and accurate response. With this approach, we expect to minimize the production or reinforcement of erroneous and/or exceedingly slow responses.

Treatment Hierarchy—A three-level hierarchy of constraint is utilized in treatment: Strong, Moderate, and Minimal-to-none. Two cueing sentences were developed to precede each probe stimulus, and to provide the intended level of contextual constraint. These cueing sentences were designed to fit Swinney’s (1991) Category 4 of context-producing procedures. That is, they rule out the context-inappropriate feature or meaning of the probe sentence’s final word, and provide a basis for anticipating the intended meaning. The first cueing sentence is strongly related to the target concept (and thus the intended feature or meaning of the probe’s sentence-final noun), and the second cueing sentence is moderately related to the target concept. *Strong constraint contexts* consist of both cueing sentences presented together, while *Moderate constraint contexts* consist of only the second sentence of the pair (see Appendix A for all cueing sentences; see Appendix C for description of procedures and results for validating strength of cueing sentence bias). The *Minimal-to-none* condition consists of presenting the probe stimulus in isolation.

The treatment hierarchy is programmed into E-Prime 2.0 software (Schneider et al., 2002). The program first checks the participant’s response for accuracy, and if accurate, compares the RT against a predetermined RT criterion to determine whether to initiate the treatment hierarchy. The RT criterion for triggering the hierarchy is determined individually for each Probe List. In addition, this criterion is customized for each participant, and represents the average of the participant’s baseline phase RTs for each list. To remind the participant to continue to respond quickly, 60% of the filler items include the response deadline signal that is also implemented in the diagnostic task (a standard Windows bell).

Treatment for each experimental item begins with the auditory presentation of the Strong constraint context (the strong plus moderate cueing sentences), prior to the probe stimulus. If the (accurate) response meets the RT criterion, the Moderate constraint context is provided prior to the probe stimulus. If RT criterion is met at this level, the probe stimulus is presented in isolation, without prior context. If the participant reaches RT criterion at this level, the program moves back up the constraint hierarchy and then goes on to the next item.

If the participant makes an error or fails to reach RT criterion with the Strong constraint context, the trial repeats with the Strong context. If the response is still inaccurate or too slow, the program moves to the next item. If a participant is inaccurate or fails to meet the RT criterion at either of the lower two levels of the hierarchy (Moderate constraint or isolated probe stimulus), E-Prime moves to the next higher level of constraint until performance meets criterion again. If accuracy or RT criterion is not achieved after two repetitions of this process, treatment moves to the next stimulus item.

Filler trials move down and back up the constraint hierarchy one time, regardless of accuracy or RT performance, before the program moves on to the next stimulus item. This procedure was designed to limit the overall length of the treatment session.

Treatment Lists—Two treatment lists (A and B) were derived from each of the 3 orders of Probe Lists 1 and 2, generating 6 orders for each list (List 1A, 1B, etc.). In developing these treatment lists, we implemented the same controls that were used in developing the alternate orders of each Probe List.

Initiating and Terminating Treatment—Stable performance on the probe task was necessary before treatment could be initiated on List 1 or 2 stimuli. In the pre-treatment baseline phase of this study, each Probe List was administered for a minimum of 5 sessions. The *a priori* criterion for this study was that performance on a given list is stable when, for the final 3 baseline sessions, there is (a) no more than a 1-item change in %Crit (increase or decrease) from one session to the next ($1/15 = 6.67\%$); or (b) a consistent decrease in %Crit by 1 or more items. When stable performance was achieved, treatment was initiated on one of the two lists. Treatment ended on that list when %Crit for accurate responses reached 80% for 3 consecutive sessions (with a minimum accuracy of 80%). Before the next list was treated, it needed to meet the same criteria for stability.

Treatment and Probe List Schedule—Treatment sessions were conducted 2–3 times per week. In the treatment phase of the design, the Probe List associated with the list currently in treatment was administered each session, before treatment itself began. After the baseline phase, probes of untreated lists were taken less frequently, as illustrated in the Results section.

Task Construction

A practiced female speaker audio recorded all sentence stimuli (probe sentences and cueing sentences), and a practiced male speaker the probe words, all at an average speaking rate of about four syllables/second. Stimuli were produced without undue emphasis on any lexical element. All recordings were made with an Audio-Technica ATR20 vocal/instrument microphone with a constant microphone-to-mouth distance (~4 inches). Recording was done in a double-walled, sound-treated booth. Stimuli were recorded onto a Dell Optiplex 760 computer with an Analog Devices audio soundcard (ADI 198x Integrated HD), using Adobe Audition 3.0 software at a sampling rate of 44.10 KHz with 16-bit resolution. Several of the authors collaborated with the speakers to achieve recording consistency.

Stimuli were then assembled using E-Prime software (Schneider et al., 2002). Probe task stimuli consisted of a trial number, a 500-ms pause, the probe sentence, a 175-ms pause, and the spoken target word. A single treatment trial consisted of the cueing sentence(s) with 500 ms between each one, the probe sentence, an interval of 175 ms, and then the target.

General procedures

Baseline and treatment sessions, including all relevant probe and generalization tasks, lasted up to 90 minutes. The duration of the treatment itself depended on Mr. R's progression through the constraint hierarchy, and usually took about 30–40 minutes. Probe and treatment lists orders were counterbalanced for administration. Mr. R. was tested in a quiet room in his house, where he was alone with the examiner(s) for all but the second follow-up probe session.

Three examiners were trained to perform the testing and the treatment. In the diagnostic and initial baseline sessions, the examiner provided Mr. R with extensive orientation, instruction,

and practice in performing the lexical decision task, first without the word-monitoring component and later with it. He was trained to indicate his lexical decision for each trial with a single finger of his right hand, and to return that finger between trials to a designated location equidistant from the labeled response buttons. He was instructed to respond as quickly and accurately as possible, and both spoken and gestural reminders about speed and consistency were provided throughout the training. Training and practice continued until he responded with assurance and his RTs had stabilized. In each new session, spoken and gestural reminders about response speed and consistency were provided prior to each block of trials.

Results

Probe List Performance

Probe List Accuracy—Although Mr. R. was only 67% accurate on the lexical decision task in diagnostic testing, he achieved a minimum of 14/15 correct for all lists in the first baseline session, and stayed above 12/15 (80%) correct throughout baseline for all lists and all further probes of Lists 1 and 2. Thus, the probe results below are presented and interpreted with reference to %Crit, the measure of the percentage of accurate responses that met the predetermined response time criterion.

Probe List %Crit—Figure 1 represents Mr. R.'s probe data for Lists 1–3, for all phases of the experiment. By the sixth baseline session, performance on List 1 probes had stabilized (i.e., %Crit changed by no more than 1 item, or 6.67%, from one session to the next, for 3 consecutive sessions). By contrast, List 2 probe data remained highly variable. Thus, treatment began on List 1 with the plan to treat List 2 second.

Lists 1 and 2—With treatment, List 1 probe performance improved rapidly, leveling off at 79% for 3 consecutive sessions by experimental session 12 (treatment session 6), after about 4 total hours of treatment. At this point, List 1 treatment ended, and probing with List 2 increased in frequency to every session, to try to achieve stability so treatment could be initiated on that list. As is evident in the figure, probe performance for List 2 never stabilized. After six List 2 probe sessions in which probe performance was increasing steadily, we abandoned the plan to treat List 2 and moved on to the follow-up phase of the experiment.

Three follow-up probe sessions were conducted. In the first follow-up session, List 1 probe performance remained as good as it was in the final sessions of the treatment phase. Performance fell off during follow-up session 2, when Mr. R. was distracted by family who were visiting, but rebounded nicely in follow-up session 3.

To augment visual inspection of List 1 results, we calculated effect sizes using the *d*-Index statistic (Bloom, Fischer, & Orme, 2003). This statistic is more conservative than some estimates of *d* in that it uses the pooled standard deviation from baseline and intervention phases. Bloom et al. also indicate that the *d*-Index is more precise in cases with a small number of baseline observations, as in this study. Based on all 6 baseline data points and all treatment data points, the *d*-index for the treatment phase was 2.52 (an increase of 49.4%) and for the follow-up phase was 1.66 (an increase of 45.2%). Calculated using only the final 3 baseline points, when probe performance had stabilized, the treatment phase *d*-index increased to 3.86 (an increase > 50%), while remaining essentially the same for the follow-up data (1.63).

List 3—During treatment of List 1, List 3 performance remained stable, per our pre-established criterion for stability. Thus, the treatment appears not to have generalized to the processing of figurative meanings of lexical metaphors. Although there is an uptick on List 3 in the first follow-up session, that is hard to interpret given the instability of the baseline performance.

Coarse Coding Generalization task—The %Crit measure for the Coarse Coding Generalization trials improved from 0/5 in the diagnostic session to 5/5 in the two follow-up sessions in which these trials were re-administered.

Performance on Broader Outcome Measures

Table 2 outlines Mr. R's performance on the various broader outcome measures. Performance on the DCT, both Total accuracy and Accuracy for questions about implied information, increased substantially from baseline to immediate post-treatment measures, far exceeding the standard error of 1 point for the standardization sample of adults with RHD. These improvements remained evident at follow-up.

It is more difficult to interpret the performance changes on the MEC Metaphor and Speech act interpretation subtests, the other two measures of language processing that were hypothesized to be related to coarse semantic coding. These measures lack standard error data, so each was administered twice during the baseline phase of the study, to estimate test-retest change. Performance on the Speech act subtest varied by 3 points (7.5%) over these two administrations, and the follow-up results, collected after List 1 treatment and during the extended probing of List 2, fall within this test-retest difference. The Metaphor subtest was particularly unstable during baseline, improving by 7 points (17.5%) from test to retest. Thus, the 2-point improvement noted at follow-up is well within the estimated error of measurement.

Performance on the control measure, MEC Emotional Prosody Production, was similarly variable in the two baseline administrations, increasing by 2 of 12 points (16.7%). However, performance on this task was unchanged from the second baseline administration to the maintenance assessment, and the 1-point gain at follow-up was within the error variability established during baseline.

Discussion

This study was performed to investigate whether a novel, implicit, stimulation-facilitation treatment for coarse coding processes could yield generalized gains to broader measures of language comprehension. Though obviously quite preliminary, the results are in some ways very promising.

Mr. R. demonstrated treatment-contingent gains, associated with respectable effect sizes, and these gains maintained through the follow-up phase of the study. These gains appear to reflect improvements in the underlying comprehension process of interest, coarse coding, rather than just item-specific improvements due to repeated exposures, because generalization was evident to untrained items in the Coarse Coding Generalization trials. The gains in treatment do not appear to be due to some form of generalized improvement, as performance did not change from baseline to later phases for either the control measure of Emotional Prosody Production or for List 3. Most importantly, the effects of treatment generalized to the DCT-based measures of narrative comprehension, both Total accuracy and Accuracy for questions about implied information, and these generalized improvements lasted into the follow-up phase.

Due to the performance variability in the pre-treatment baseline phase, cautious interpretation is warranted for List 3 results. The apparent lack of generalization to the processing of lexical metaphors raises intriguing questions about the limits on theories of continuity between literal and nonliteral language processing (Coulson & Van Petten, 2002; Grady et al., 1999; Pynte et al., 1996). In addition, as noted earlier, this result would be consistent with the claim that some adults with RHD may have a metaphor-specific processing deficit (Klepousniotou & Baum, 2005a), separate from but potentially co-occurring with coarse coding and/or suppression deficits. This finding, if sufficiently replicated, could reinforce the common clinical practice of evaluating and treating metaphoric stimuli separately from literal stimuli in adults with RHD.

Extensive pre-treatment variability makes it impossible to interpret the results for the measures of Metaphor and Speech act interpretation. As we gather data from more participants we will be able to ascertain whether the estimated re-test reliability of these measures is sufficient for tracking change in the manner that we wish to.

Overall, the results of this study support the potential for generalization to meaningful comprehension measures from treatment that implicitly targets an underlying comprehension process. Contextual Constraint Treatment is very different from the metalinguistic association tasks that typify most clinical interventions for neurologically-based language disorders. Provision of the constraint contexts is the sum total of the treatment; the contexts serve only as primes, and are never explicitly associated with a probe stimulus, either by the examiner or the participant. In addition, the treatment's focus on underlying processes should, in theory, yield broader-based generalization than does treatment that targets a specific language structure or form, such as metaphor interpretation. If promising results for this treatment approach continue to accrue, future work can compare it with more typical metalinguistic approaches that engage participants in guided problem-solving and self-discovery.

The RHD treatment literature contains a gaping hole where an evidence base should be. Evidence has begun to accrue for treatment of expressive aprosodia (e.g., Leon & Rodriguez, 2008; Rosenbek, Crucian, Leon, Hieber, Rodriguez, Holiway, B., . . . Gonzalez-Rothi, L., 2004; Rosenbek, Rodriguez, Hieber, Leon, Crucian, Ketterson, . . . Gonzalez-Rothi, 2006) and a metalinguistic treatment has been developed for the oral interpretation of metaphoric phrases (e.g., "a family is a cradle," Lundgren, Brownell, Cayer-Meade, Milione, & Kearns, 2011). Otherwise, treatment for communication disorders in adults with RHD has been largely untested. In addition, as described earlier in this paper, such treatment is often based on tenuous assumptions. Our investigation of Contextual Constraint Treatment is ongoing, and includes individuals with suppression deficits, for whom another version of the treatment has been developed. With these efforts we hope to bolster the scientific underpinnings of the treatment of language comprehension difficulties in adults with RHD.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Preparation of this manuscript was supported by Grant #DC010182 from the National Institutes of Health (National Institute on Deafness and Other Communication Disorders). I would also like to acknowledge Bob Marshall, who was instrumental in changing the course of my career. Quite a while ago, Bob gave me an opportunity that provided a major impetus for me to return to school for my PhD. He had hired me, a new Master's graduate, as a research assistant and CFY in his shop at the Portland, Oregon, Veterans Hospital. Several months after I started my job, he

took me to the Clinical Aphasiology Conference and allowed me to present the research on which I had been assisting. I had done a master's thesis and had been told that I should go on for a PhD, but attending that first CAC – and witnessing the spirited intellectual exchanges there – was a major factor in my eventually taking the plunge. I will always be grateful to Bob for that experience in particular, and for his support in general.

References

- Atchley RA, Burgess C, Keeney M. The effect of time course on the facilitation of semantic features in the cerebral hemispheres. *Neuropsychology*. 1999; 13(3):389–403. [PubMed: 10447300]
- Beeman, M. Coarse semantic coding and discourse comprehension. In: Beeman, M.; Chiarello, C., editors. *Right hemisphere language comprehension: Perspectives from cognitive neuroscience*. Mahwah, NJ: Erlbaum; 1998. p. 255–284.
- Blake ML. Inferencing processes after right hemisphere brain damage: Effects of contextual bias. *Journal of Speech, Language and Hearing Research*. 2009; 52(2):373–384.
- Blake ML, Lesniewicz K. Contextual bias and predictive inferencing in adults with and without right hemisphere brain damage. *Aphasiology*. 2005; 19(3–5):423–434.
- Bloom, M.; Fischer, J.; Orme, JG. *Evaluating practice – Guidelines for the accountable professional*. 4th ed.. Boston: Allyn & Bacon, Pearson Education, Inc; 2003.
- Brookshire, RH.; Nicholas, LE. *Discourse Comprehension Test*. Tucson, AZ: Communication Skill Builders; 1993.
- Coulson S, Van Petten C. Conceptual integration and metaphor: An event-related potential study. *Memory & Cognition*. 2002; 30(6):958–968.
- Gernsbacher, MA. *Language comprehension as structure building*. Hillsdale, NJ: Erlbaum; 1990.
- Gernsbacher MA, Faust ME. The mechanism of suppression: A component of general comprehension skill. *Journal of Experimental Psychology: Learning, Memory, & Cognition*. 1991; 17:245–262.
- Grady, J.; Oakley, T.; Coulson, S. Conceptual Blending and Metaphor. In: Gibbs, R.; Steen, G., editors. *Metaphor in cognitive linguistics*. Amsterdam & Philadelphia: John Benjamins; 1999. p. 101–124.
- Joanette, Y.; Ska, B.; Côté, H. *Protocole Montréal d'Évaluation de la Communication (Protocole MEC)*. Isbergues, France: Ortho-Edition, 2004; 2004.
- Jung-Beeman M. Bilateral brain processes for comprehending natural language. *Trends in Cognitive Sciences*. 2005; 9:712–718.
- Kintsch W. The use of knowledge in discourse processing: A construction-integration model. *Psychological Review*. 1988; 95:163–182. [PubMed: 3375398]
- Kintsch, W. *Comprehension: A paradigm for cognition*. Cambridge, UK: Cambridge University Press; 1998.
- Klepousniotou E. The processing of lexical ambiguity: Homonymy and polysemy in the mental lexicon. *Brain & Language*. 2002; 81(1–3):205–223. [PubMed: 12081393]
- Klepousniotou E, Baum SR. Processing homonymy and polysemy: Effects of sentential context and time-course following unilateral brain damage. *Brain and Language*. 2005a; 95(3):365–382. [PubMed: 16298667]
- Klepousniotou E, Baum S. Unilateral brain damage effects on processing homonymous and polysemous words. *Brain & Language*. 2005b; 93(3):308–326. [PubMed: 15862856]
- Lehman-Blake MT, Tompkins CA. Predictive inferencing in adults with right hemisphere brain damage. *Journal of Speech, Language, and Hearing Research*. 2001; 44:639–654.
- Leon SA, Rodriguez AD. Aprosodia and its treatment. *Perspectives on Neurophysiology and Neurogenic Speech and Language Disorders*. 2008; 18:66–72.
- Lundgren K, Brownell HH, Cayer-Meade C, Milione J, Kearns KP. Treating metaphor interpretation deficits subsequent to right hemisphere brain damage: Preliminary results. *Aphasiology*. 2011; 25(4):456–474.
- Pynte J, Besson M, Robichon F, Poli J. The time-course of metaphor comprehension: An event-related potential study. *Brain & Language*. 1996; 55:293–316. [PubMed: 8954602]
- Radloff LS. The CES-D Scale: A self-report depression scale for research in the general population. *Journal of Applied Psychological Measurement*. 1977; 1(3):385–401.

- Rosenbek JC, Crucian GP, Leon SA, Hieber B, Rodriguez AD, Holiway B, Gonzalez-Rothi L. Novel treatments for expressive aprosodia: A phase I investigation of cognitive linguistic and imitative interventions. *Journal of the International Neuropsychological Society*. 2004; 10(5):786–793. [PubMed: 15327724]
- Rosenbek JC, Rodriguez AD, Hieber B, Leon SA, Crucian GP, Ketterson TU, Ciampitti M, Singletary F, Heilman KM, Gonzalez-Rothi L. The effects of two treatments for aprosodia secondary to acquired brain injury. *Journal of Rehabilitation Research and Development*. 2006; 43:379–390. [PubMed: 17041823]
- Schneider, W.; Eschman, A.; Zuccolotto, A. E-prime user's guide. Pittsburgh, PA: Psychology Software Tools Inc.; 2002.
- Swinney, DA. The resolution of indeterminacy during language comprehension: Perspectives on modularity in lexical, structural, and pragmatic processes. In: Simpson, GB., editor. *Understanding word and sentences*. North-Holland: Elsevier Science Publishers; 1991. p. 367-385.
- The New Oxford American Dictionary. Oxford: Oxford University Press, Inc.; 2001.
- Tillman, T.; Carhart, R. An expanded test for speech discrimination utilizing CNC monosyllabic words (Northwestern University Auditory Test No. 6) (Technical Report No. SAM-TR-66-55). Brooks Air Force Base, TX: USAF School of Aerospace Medicine, Aerospace Medical Division (AFSC); 1966.
- Tompkins CA. Knowledge and strategies for processing lexical metaphor after right or left hemisphere brain damage. *Journal of Speech and Hearing Research*. 1990; 33:307–316. [PubMed: 2359271]
- Tompkins CA. Automatic and effortful processing of emotional intonation after right or left hemisphere brain damage. *Journal of Speech and Hearing Research*. 1991a; 34:820–830. [PubMed: 1956190]
- Tompkins CA. Redundancy enhances emotional inferencing by right- and left- hemisphere-damaged adults. *Journal of Speech and Hearing Research*. 1991b; 34:1142–1149. [PubMed: 1749244]
- Tompkins CA. Theoretical considerations for understanding “understanding” by adults with right hemisphere brain damage. *Perspectives on Neurophysiology and Neurogenic Speech and Language Disorders*. 2008; 18:45–54. [PubMed: 20011667]
- Tompkins CA, Baumgaertner A, Lehman MT, Fassbinder W. Mechanisms of discourse comprehension impairment after right hemisphere brain damage: Suppression in lexical ambiguity resolution. *Journal of Speech, Language, and Hearing Research*. 2000; 43:62–78.
- Tompkins CA, Blake ML, Meigh KM, Wambaugh J. A novel, implicit treatment for language comprehension processes in right hemisphere brain damage: Phase I data. *Aphasiology*. 2011; 25(6–7):789–799. [PubMed: 22368317]
- Tompkins CA, Boada R, McGarry K. The access and processing of familiar idioms by brain-damaged and normally aging adults. *Journal of Speech and Hearing Research*. 1992; 35:626–637. [PubMed: 1608254]
- Tompkins CA, Fassbinder W, Blake ML, Baumgaertner A, Jayaram N. Inference generation during text comprehension by adults with right hemisphere brain damage: Activation failure versus multiple activation. *Journal of Speech, Language, and Hearing Research*. 2004; 47(6):1380–1395.
- Tompkins CA, Fassbinder W, Scharp VL, Meigh KM. Activation and maintenance of peripheral semantic features of unambiguous words after right hemisphere brain damage in adults. *Aphasiology*. 2008; 22:119–138. [PubMed: 20011607]
- Tompkins, CA.; Klepousniotou, E.; Gibbs Scott, A. Nature and assessment of right hemisphere disorders. In: Papathanasiou, I.; Coppens, P.; Potagas, C., editors. *Aphasia and related neurogenic communication disorders*. Sudbury, MA: Jones & Bartlett; 2012. p. 297-343.
- Tompkins CA, Lehman-Blake MT, Baumgaertner A, Fassbinder W. Mechanisms of discourse comprehension impairment after right hemisphere brain damage: Suppression in inferential ambiguity resolution. *Journal of Speech, Language, and Hearing Research*. 2001; 44(2):400–415.
- Tompkins CA, Scharp VL, Meigh KM, Fassbinder W. Coarse coding and discourse comprehension in adults with right hemisphere brain damage. *Aphasiology*. 2008; 22(2):204–223. [PubMed: 20037670]
- Vanhalle C, Lemieux S, Joubert S, Goulet P, Ska B, Joanne Y. Processing of speech acts by right hemisphere brain-damaged patients: An ecological approach. *Aphasiology*. 2000; 14:1127–1142.

Webster's Encyclopedic Unabridged Dictionary of the English Language: New revised edition.
Avenel, NJ: Gramercy Books; 1994.

% Crit = percentage of correct responses that met response time criterion

* Probe #19 conducted on day prior to Thanksgiving;
probe condition less than optimal due to family members creating distractions

Figure 1.
Mr. R.'s Performance on 3 Probe Lists Across Experimental Design Phases

Table 1

Clinical characteristics for Mr. R. and participants with RHD from a prior study of coarse coding and discourse comprehension (Tompkins, Scharp, et al., 2008)

	Mr. R.	Prior RHD Participants (N = 32)
Auditory Working Memory for Language ^a		
Word recall errors	15	M = 13.2; SD = 7.0
Behavioural Inattention Test ^b	146	M = 137; SD = 13.5
Visual Form Discrimination ^c	27	M = 28.1; SD = 3.5
Judgement of Line Orientation ^d	25	M = 22.2; SD = 5.2
ABCD ^e Story Retell		
Immediate Retell	15	M = 13.2; SD = 2.5
Delayed Retell	14	M = 12.7; SD = 3.1
PPVT-R ^f raw score	162	M = 157.3; SD = 11.3

Note. RHD = Right Hemisphere Damaged ; M = Mean ; SD = Standard Deviation

^aTompkins et al. (1994; maximum errors = 42).

^bWilson, Cockburn, & Halligan (1987; maximum = 146; neglect cutoff = 129).

^cBenton, Sivan, Hamsher, Varney & Spreen, (1983; maximum = 32; cutoff for defective performance = 23).

^dBenton, Hamsher, Varney, & Spreen (1983; age & gender corrected score maximum = 35).

^eABCD = Arizona Battery for Communication Disorders in Dementia; Bayles & Tomoeda, (1993; maximum = 17)

^fPPVT-R = Peabody Picture Vocabulary Test--Revised; Dunn & Dunn (1981; maximum = 175).

Table 2

Mr. R's performance on broader outcome measures

	Pre-treatment	Immediate post-treatment	Maintenance	Follow-up
DCT Total (80 possible) ^a	55 (69%)	31/40 ^b (78%)	N/A	62 (78%)
DCT Implied information (40 possible)	27 (68%)	15/20 ^b (75%)	N/A	30 (75%)
MEC Speech Acts ^c (40 possible)	35, 38	N/A	37	40
MEC Metaphor ^c (40 possible)	27, 34	N/A	N/A	36
MEC Emotional Prosody Production ^c (12 possible)	2, 4	N/A	4	5

Note. DCT = Discourse Comprehension Test (Brookshire & Nicholas, 1993); MEC = experimental English-language version of Protocole Montreal d'Evaluation de la Communication (Joanette et al., 2004); %Crit = percentage of accurate lexical decisions that met pre-established response time criterion; N/A = not administered.

^aStandard error for RHD standardization sample = 1 point

^bOnly Set A administered (half the total items)

^cAdministered twice during baseline phase to estimate variability associated with simple re-test