Published in final edited form as:

Neuropsychol Rehabil. 2015 January ; 25(1): 15–52. doi:10.1080/09602011.2014.932290.

Contextual Constraint Treatment for coarse coding deficit in adults with right hemisphere brain damage: Generalization to narrative discourse comprehension

Margaret Lehman Blake,

Department of Communication Sciences & Disorders, University of Houston, 100 Clinical Research Services, Houston, TX 77204-6018, 713-743-2894

Connie A. Tompkins,

Department of Communication Science & Disorders, University of Pittsburgh, 4033 Forbes Tower, Pittsburgh, PA 15260, 412-383-6536

Victoria L. Scharp,

Department of Communication Science & Disorders, University of Pittsburgh, 4033 Forbes Tower, Pittsburgh, PA 15260, 412-383-6536

Kimberly M. Meigh, and

Department of Communication Science & Disorders, University of Pittsburgh, 4033 Forbes Tower, Pittsburgh, PA 15260, 412-383-6536

Julie Wambaugh

Communication Sciences & Disorders, University of Utah and VA Salt Lake City Healthcare Systems, 151 A, 500 Foothill Blvd., Salt Lake City, UT 84148, 801-582-1565 ext. 1363

Margaret Lehman Blake: mtblake@uh.edu; Connie A. Tompkins: tompkins@pitt.edu; Victoria L. Scharp: scharpvl@hotmail.com; Kimberly M. Meigh: kimmeigh@gmail.com; Julie Wambaugh: Julie.Wambaugh@health.utah.edu

Abstract

Coarse coding is the activation of broad semantic fields that can include multiple word meanings and a variety of features, including those peripheral to a word's core meaning. It is a partially domain-general process related to general discourse comprehension and contributes to both literal and non-literal language processing. Adults with damage to the right cerebral hemisphere (RHD) and a coarse coding deficit are particularly slow to activate features of words that are relatively distant or peripheral. This manuscript reports a pre-efficacy study of Contextual Constraint Treatment (CCT), a novel, implicit treatment designed to increase the efficiency of coarse coding with the goal of improving narrative comprehension and other language performance that relies on coarse coding. Participants were four adults with RHD. The study used a single-subject controlled experimental design across subjects and behaviors. The treatment involves pre-stimulation, using a hierarchy of strong- and moderately-biased contexts, to prime the intended distantly-related features of critical stimulus words. Three of the four participants exhibited gains in auditory narrative discourse comprehension, the primary outcome measure. All participants exhibited generalization to untreated items. No strong generalization to processing nonliteral language was

evident. The results indicate that CCT yields both improved efficiency of the coarse coding process and generalization to narrative comprehension.

Keywords

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

It has been well established that the right cerebral hemisphere (RH) plays a role in language comprehension (see reviews in Beeman & Chiarello, 1998; Chiarello, 2003) and that damage to the right side of the brain (RHD) can result in language comprehension deficits (Blake, 2011; Benton & Bryan, 1996; Côté, Payer, Giroux & Joanne, 2007; Joanne, Goulet & Hannequin, 1990; Myers, 1979; Tompkins, Klepousniotou & Gibbs Scott, 2012). Important questions remain about the specific nature of the processing deficits, the effects of the hypothesized deficits on broader language abilities, and how to treat the deficits.

Evidence is fairly consistent for a RH role in semantic processing. In various priming studies of healthy young adults, results suggest that the RH activates and/or maintains activation of related word meanings and features that are relatively distant (Atchley, Burgess & Keeney, 1999; Burgess & Simpson, 1988; Klepousniotou & Baum, 2005; but see Fassbinder & Tompkins, 2006, for cautions in interpreting priming results). These include subordinate, secondary, or less-common meanings of ambiguous words (Burgess & Simpson, 1988; Klepousniotou & Baum, 2005); weakly-related meanings; and distant features (Atchley et al., 1999). Beeman (1998; Jung-Beeman, 2005) incorporated these findings into his fine/coarse coding model of semantic processing. This model suggests that the left hemisphere (LH) uses fine semantic coding in which, during initial processing of a word, the LH strongly activates a small semantic field and then quickly selects the most common meaning and/or closely related features. In contrast, the RH is proposed to use coarse coding to weakly activate a broader semantic field that can include multiple word meanings and a variety of features, including those that are quite peripheral to the word's core meaning. Activation of this broad field also is maintained temporarily in the RH. As an example, given the word "piano," the left hemisphere might quickly activate features like keys, music and black. The right hemisphere would more broadly activate semantic features such as melody, song, and lessons.

The activation and maintenance of broad semantic fields, including distantly-related, peripheral meanings and features, has been hypothesized to be a partially domain-general process (e.g., Anaki, Faust & Kravetz, 1998; Beeman, 1993; Beeman, 1998; Mashal & Faust, 2008; Richards & Chiarello, 1995; Titone, 1998). This means that variations in coarse coding efficiency could affect multiple comprehension processes and performances. In this theoretical construction, coarse coding underlies various aspects of language comprehension, such as deriving figurative or implicit meanings, interpreting jokes, generating inferences that rely on overlapping semantic fields, and revising interpretations.

The majority of the semantic processing research has been conducted with healthy young adults, and results are extrapolated from processes observed in a healthy RH to provide post-

hoc explanations of RHD deficits. Additionally, most studies focus on generation of multiple meanings of homonyms (e.g., Burgess & Simpson, 1988; Klepousniotou & Baum, 2005). Tompkins and colleagues (Tompkins, Fassbinder, Scharp & Meigh, 2008; Tompkins, Scharp, Meigh & Fassbinder, 2008) conducted two direct studies of coarse coding in adults with RHD focusing on activation of distantly-related, or peripheral, features of unambiguous words. To maximize the likelihood that coarse coding processes were being tapped, extensive piloting was conducted to verify that the features were only peripherally related. Distantly-related features were defined using methods from Atchley and colleagues (1999). First, participants generated features of common, unambiguous words (e.g., apple). Features mentioned by 10–24% of participants were considered distantly-related. These features were then rated in terms of how compatible they were with the most common mental image of the target concept. Those rated as incompatible (e.g., “rotten”) were considered to be particularly distant. Results from the first study (Tompkins, Fassbinder et al.) indicated that adults with RHD do activate and maintain a variety of peripheral meanings and features, including those that are distant but compatible with a common mental image (e.g., “crunchy”), but they have difficulty with features that are particularly remote from (and incompatible with) the core concept. Tompkins and colleagues’ (Tompkins, Scharp et al.) second study indicated that coarse coding deficits were related to general comprehension of narratives, particularly comprehension of implied meanings. Some stimuli and validation procedures from these studies were used in the current treatment, to verify that coarse coding processes were being tapped.

In the vein of classical stimulation-facilitation approaches to treatment, Contextual Constraint Treatment (CCT) was designed to facilitate basic language comprehension processes, including coarse coding, in adults with RHD. CCT is a novel, implicit, contextual stimulation treatment, which is intended to speed the targeted basic comprehension process (i.e., coarse coding). The implicit nature of CCT, which requires no direct responses to the treatment stimuli or post-stimulus questions, mimics natural comprehension and minimizes the metacognitive and metalinguistic demands that often negatively affect performance by adults with RHD. Evidence from several studies suggests that adults with RHD tend to have difficulty when typical deficit areas are assessed metacognitively or metalinguistically, yet perform within the range of healthy adults when the same processes are assessed implicitly (e.g., Tompkins, 1990, 1991). For example, difficulties are noted in defining idioms (Kempler, Van Lancker, Merchman & Bates, 1999; Myers & Linebaugh, 1981; Tompkins, Boada & McGarry, 1992), yet in an implicit measure involving word monitoring, the same participants performed as well as a control group, indicating rapid processing of idiomatic meanings (Tompkins et al., 1992). Another reason for using an implicit treatment is that participants with coarse coding deficits are generally accurate on coarse coding assessment tasks, but slow to respond. Thus, an implicit task requiring speeded responses should increase the speed and efficiency of a process that already functions sufficiently for accurate responding.

Contextual constraint is used in CCT to pre-stimulate or prime interpretations and features of words that adults with CC deficits would not otherwise activate in a timely manner. Blake (2009; Blake & Lesniewicz, 2005) and others have reported that adults with RHD, despite having other language processing deficits, are sensitive to strong contextual cues. This

sensitivity allows them to perform as well as adults without brain injury on tasks requiring determination of an appropriate referent for ambiguous pronouns (Leonard, Waters & Caplan, 1997a, b), generation of predictive inferences (Lehman-Blake & Tompkins, 2001), and comprehension of potentially ambiguous phrases or sentences (Brownell, Potter, Bihrlé & Gardner, 1986; Tompkins, Baumgaertner, Lehman & Fassbinder, 2000; Tompkins, Lehman-Blake, Baumgaertner & Fassbinder, 2001). The contextual constraint in CCT is designed to establish the focus of each stimulus item, which is the distantly related feature of a common noun. For example, sentences about a melody with words stimulate the concept of SONG that is a distantly-related feature of piano.

Another reason for incorporating pre-stimulation in CCT is to assist in minimizing erroneous or slow responses, by providing strong contextual support that is reduced only after fast and accurate responses. This taps into one of the components of the errorless (or error-limited) learning approach – beginning with maximal cues that are gradually faded after correct responses (e.g., Fillingham, Hodgson, Sage & Lambon Ralph, 2003; Fillingham, Sage, & Lambon Ralph, 2006; Kessels & deHaan, 2003). While there is conflicting evidence regarding whether errorless or errorful learning is more effective for language rehabilitation, both have been shown to have positive effects (Conroy, Sage & Lambon Ralph, 2009; Middleton & Schwartz, 2012).

The goal of CCT for participants in the current study was to increase the efficiency of the coarse coding process itself, in an attempt to reap broad-based benefits of the treatment. Thus the aim of CCT was to facilitate narrative comprehension and other comprehension outcomes that are hypothesized or demonstrated to be related to coarse coding. Generalization to narrative comprehension was the primary outcome of interest. In addition, response generalization was assessed to untreated items, to document improvement in the underlying coarse coding process. While the magnitude of gains on treatment probes typically is a major focus in language treatment studies, treatment-related effect sizes are of much less importance for this study than the generalization measures. Additionally, because the focus of CCT is to speed accurate but inefficient comprehension processes, the typical measures of treatment-contingent effect sizes may not be as meaningful as they are for aphasia or other language treatments in which the purpose is to improve accuracy.

To examine the theoretical claim that coarse coding is a partially domain-general process, we also assessed generalization to the processing and interpretation of non-literal language. It has been proposed that adults with RHD have a specific deficit in processing metaphors and other non-literal language (e.g., Klepousniotou & Baum, 2005; Lundgren, Brownell, Cayer-Meade, Milione & Kearns, 2011). However, in line with several theories on the continuity of literal and nonliteral language processing (Coulson & Van Petten, 2002; Grady, Oakley & Coulson, 1999; Pynte, Besson, Robichon & Poli, 1996), it is also possible that difficulties with the activation and maintenance of metaphoric meanings reflect a special case of coarse coding in which the non-literal meanings reflect subordinate or distant meanings (Tompkins, Fassbinder et al., 2008; Tompkins, Scharp et al., 2008). In this case, improving the coarse coding process should facilitate access to metaphoric meanings, and improve interpretation of metaphors and perhaps other non-literal language forms. We evaluated this hypothesis by (a) tracking response generalization to lexical metaphor stimuli

and (b) examining stimulus generalization to the metalinguistic interpretation of phrasal metaphors (e.g., The teacher is a sleeping pill) and indirect language.

This manuscript extends the investigation of CCT that is described in previous reports, including Phase I results (Tompkins, Blake, Meigh & Wambaugh, 2011) and treatment outcomes and generalization to auditory narrative comprehension for one participant (Tompkins, Scharp, Meigh, Blake & Wambaugh, 2012). This pre-efficacy study (Phase II based on Robey, 2004 and Robey & Schultz, 1998) was designed to validate procedures and evaluate the potential benefits of the treatment for broader comprehension processes prior to investing in more complex and costly case series or randomized controlled treatment studies (Gonzalez Rothi, 2006; Robey & Schultz, 1998). Fully expecting that CCT will result in treatment-contingent gains, as indicated by faster responses to the probe stimuli, the major hypotheses for this study were as follows: (1) Gains from CCT will generalize to narrative comprehension; (2) Response generalization will occur to untreated items and lexical metaphoric meanings, in the latter case consistent with the hypothesis that activating these meanings is a special case of coarse coding; (3) Gains from CCT will generalize to the explicit interpretation of non-literal language (non-lexical metaphors and indirect requests); (4) No change in control task performance will be observed, indicating that CCT stimulates coarse coding processes and that the gains of interest are not due simply to practice effects or general cognitive stimulation.

Methods

Participants

Participants were three males and one female with a cerebrovascular accident (CVA or stroke) in the right cerebral hemisphere verified by CT and/or MRI scan report, recruited from the two study sites: Pittsburgh, PA and Houston, TX. Three participants were included to investigate across-participant replication (Thompson, 2006) of the results from the fourth, Mr. R, whose positive response to treatment was reported in a previous manuscript (Tompkins et al., 2012). Mr. R's data are included here as well, for comparison purposes. All participants were right handed, native speakers of American English with no history of psychiatric or neurological disease (other than stroke) or substance abuse. Demographic and clinical data are provided in Table 1.

An extensive hearing screening was conducted on each participant to ensure adequate hearing for the auditorily-presented treatment tasks. Study participants had to have pure tone thresholds within one standard deviation of age and sex-related norms (Beaver Dam Norms; Cruikshanks, Wiley, Tweed, Klein, Klein, Mares-Perlman et al., 1998) at 500, 1000, 2000, and 4000Hz. Additionally, all participants had to pass a speech perception screening using the Northwestern University Test #6 (Tillman & Carhart, 1966), presented at a level 40dB greater than the pure tone average (PTA) for the better ear. "Passing" was based on the percent correct falling within the 95% confidence interval norms published by Dubno and colleagues (Dubno, Lee, Klein, Matthews & Lam, 1995). To ensure audibility of study stimuli, hearing profiles based on PTA results were created for all participants who had a hearing loss but no hearing aids. All auditory stimuli for this study were frequency shaped in Adobe Audition to match each hearing profile. Specifically, for each profile, the intensity of

the stimulus signal was adjusted within eight frequency bands, to mimic the results a hearing aid would provide. Intensity was increased at 4 and 8 KHz for all hearing profiles, and also at 2 KHz for profiles 2 (mild-moderate loss), 3 (moderate-severe loss) and 4 (severe loss). For profiles 3 and 4, intensity was simultaneously decreased for the remaining frequency bands (63, 125, 250, 500, and 1000 Hz).

Ancillary tasks, including measures of working memory, neglect, and immediate and delayed memory, were administered to characterize the participants (see Table 2). On all of the ancillary measures, all four participants fell within the range of performance for the RHD group from the original studies of coarse coding (Tompkins, Fassbinder et al., 2008).

Diagnostic Task and Procedures

Coarse coding deficits were identified using the auditory lexical decision task and methods from Tompkins and colleagues' original study (Tompkins, Fassbinder et al., 2008). The probe stimuli were 15 short, semantically-neutral sentences that ended with a common concrete, unambiguous noun (e.g., "There was a piano") followed by a 175 ms inter-stimulus interval, then a phoneme string (see Appendix A for all diagnostic and treatment stimuli). The participants' task was to indicate whether the phoneme string was a real word or a nonword by pressing either "yes" or "no" on a manual response box. For experimental items, the phoneme string was a common, concrete, unambiguous noun that was distantly related to the sentence-final noun (e.g., SONG).

"Yes" was the correct response for all experimental coarse coding items. The 15 experimental items were divided into two sets; each set was interspersed with 20 filler items. The filler items were similar neutral sentences ("There was mustard") followed either by non-words (e.g., FURNIBO), more closely related words (SPICY), or unrelated words (TEAMMATE). Overall, 60% of the diagnostic task items required Yes responses (of these 50% were experimental items; 20% closely related; 30% unrelated). To remind participants to keep responding quickly, a "ding" (standard Microsoft Windows bell sound) was presented after the phoneme string on 60% of the filler items.

Extensive validation of the probe stimuli was conducted (see Tompkins et al., 2012, for details). Distantly-related features were determined using the procedures described in the introduction. Additionally, for experimental items, sociodemographically appropriate raters verified that sentence-final nouns and real-word targets were semantically distant. Real-word phoneme strings were matched on a variety of lexical properties, including abstractness, prosodic features, log frequency, and mean reaction time (e-lexicon; Balota et al., 2002).

A secondary word monitoring task was added to encourage participants to attend to the sentences, given that the lexical decision could be made without doing so. Before the list of stimulus probes was presented, participants were told to listen for a specific word that was printed on a 3×5 card and set on top of the response box, in clear view for the participant. This word was always the sentence-final word of a filler item that appeared within the last five items of a set. This word monitoring task has been used successfully in previous studies of language processing in RHD without interfering with fast, accurate performance on the

primary task (Tompkins Fassbinder et al., 2008; Tompkins, Scharp et al., 2008). Responses to the word monitoring task were uniformly highly accurate, and will not be discussed further.

All stimuli were presented auditorily using EPrime software (Schneider, Eschman & Zuccolotto, 2002) on a Dell Latitude E6400 laptop. Signals were routed through an Anchor Audio AN 130 BK speaker. Auditory presentation levels were set individually based on each participant's hearing profile and the ambient environmental noise¹. Participants made yes/no responses on an EPrime serial response box, using a single finger to respond to all items. Between items, participants rested that finger on a space equidistant from the yes and no response buttons. Participants were instructed to respond as quickly and accurately as possible.

Participants who were too slow or inaccurate on over half (e.g., 8 or more) of the experimental diagnostic items were identified as having a coarse coding deficit. None of the participants made more than five errors, indicating that their lexical decision performance was clearly above chance. "Too slow" was defined with reference to the median reaction time (RT) for each experimental item obtained from the healthy adult control group from the original coarse coding study (Tompkins, Fassbinder et al., 2008), that was then adjusted for the degree of generalized slowing that has characterized RHD participants in our prior research (e.g., Tompkins et al. 2000; Tompkins, Scharp, Gibbs-Scott & Meigh, in preparation). Specifically, the median per-item RTs of the control group were multiplied by 1.5. For the current study, RTs that were at least 1.5 SD slower than this adjusted median were considered "too slow," and potentially indicative of a coarse coding deficit. Slow responses in and of themselves may not reflect a language processing deficit; however, each of the RHD participants who had particularly slow responses to the Coarse Coding task, even after adjusting for generalized slowing in RHD, also had impaired discourse comprehension.

In addition to the Coarse Coding task, participants completed the Immediate and Delayed story retell task from the Arizona Battery for Communication Disorders (ABCD, Bayles & Tomoeda, 1993) to rule out incipient dementia, and a baseline manual reaction time task. The diagnostic session lasted approximately 90 minutes.

Treatment Tasks and Procedures

Experimental design—This study used a single-subject controlled experimental design across subjects and behaviors. The design included both a multiple baseline across subjects component and a multiple probe across behaviors (treatment lists) component. The experiment consisted of four phases: baseline, treatment, maintenance (sessions that occurred immediately after treatment of one list and prior to initiation of treatment on another list) and follow-up (see Appendix B for outline of sample sessions across phases). The primary outcome was stimulus generalization to narrative comprehension. Secondary

¹Presentation level was 15dB greater than the ambient environmental noise (per sound level meter), except when environmental noise was low. In that case, presentation level was either 65dB (for participants with normal or corrected to normal hearing), or 80 dB (for participants with hearing loss and no hearing aid).

outcomes were response generalization to untrained items and lexical metaphor processing, and stimulus generalization to interpretation of nonliteral language (phrasal metaphors and indirect speech acts). Acquisition and maintenance of treated items also were tracked over time.

Per expectations of single subject experimental designs (e.g., Byiers, Reichle & Symons, 2012; Kratochwill, 2010; Thompson, 2006), the study included repeated baseline probing until stability was observed, to use as reference points for evaluating the nature of changes on the treatment items themselves and, most importantly in this study, the generalization effects of the treatment. Repeated measurements/probes within and between phases of the design were administered to allow valid interpretations of observed changes. Generalization probes at each change of phase allowed documentation of when reliable change has occurred. The probe and treatment tasks were identical to the lexical decision tasks used in the diagnostic session, as derived from our prior studies of coarse coding. We note that lexical decision performance in treatment itself is not a critical element or outcome of CCT – it is self-evident that a response that is slow will get faster as it is practiced over time. The lexical decision task is merely the engine through which we aim to facilitate the coarse coding process. The active ingredient of the treatment is the repeated presentation and removal of the context sentences that bias particular semantic features.

For all probe and treatment lists, accuracy and RT were monitored. Although accuracy varied across participants in the diagnostic session (see Results section), accuracy was high (minimum of 88%) throughout study phases for all participants, thus it will not be discussed further. All subsequent references to RTs are for accurate responses only.

Probe task—The probe stimuli (neutral sentences followed by a phoneme string) and task (lexical decision) were based on the experimental/diagnostic items described above. All stimuli were extensively piloted and validated using methods similar to those for the diagnostic items. Three stimulus lists were created, each with 15 experimental and 10 filler items. Sixty percent of items required “Yes” responses. Across lists, items were matched in terms of semantic fit (based on ratings from piloting) and response time to the real-word phoneme string (e-lexicon, Balota et al., 2002). Additionally, three versions of each list were created by changing the order of stimuli following a set of controls, including beginning/ending each list with filler items, changing protagonist gender (he/she/it) in successive items, using no more than three experimental items in succession, and separating themes that could be construed to be related. The presentations of the lists, and versions of each list, were counterbalanced within and between participants.

Lists 1 and 2 were treatment lists. List 3 was a response generalization list, described below in the Outcome Measures section. The dependent variable was the percentage of accurate responses to experimental probe stimuli that met a pre-set response time criterion (%Crit). This criterion, calculated independently for each list and ranging from 368–385ms, was one standard deviation slower than (i.e., worse than) the mean achieved by the non-brain-damaged control participants in a prior study of coarse coding (Tompkins, Fassbinder et al., 2008). This criterion was considered to reflect a substantial change in response time, and was used to make objective decisions about when to begin different phases of treatment

(ending baseline and beginning treatment, ending treatment and beginning maintenance, etc.).

Baseline phase—Baseline sessions began with presentation of each of the three probe lists. The outcome measures and ancillary tasks to measure clinical variables were distributed across the first five baseline sessions. These sessions lasted approximately 90 minutes each.

A minimum of five baseline sessions was conducted with each participant and additional sessions were conducted as needed until stable performance was obtained for at least one of the two treatment lists. Stability was defined as improvement in %Crit by no more than 1 item (6.7%) or a decrease in %Crit over three consecutive baseline sessions. The four participants required from 5 to 8 sessions to meet the stability criterion on at least one of the two treatment lists.

Treatment and maintenance phases

Treatment stimuli and task: As described in the introduction, CCT is designed to facilitate engagement of bottom-up, context-independent coarse coding processes to activate the distantly related features of nouns (e.g., SONG as a feature of PIANO). The context sentences in the treatment hierarchy were designed to strengthen the neural pathways that support coarse coding, by repeatedly pairing the nouns of interest with their semantically-distant target features, and eliciting fast and accurate connections between them.

The treatment involved two levels of contextual constraint preceding the probe stimuli. Context sentences were created to strongly or moderately bias interpretation towards the distantly-related semantic features of the sentence-final word of the probe stimulus (see Table 3 and Appendix A). The first sentence was strongly biased, and the second sentence moderately biased toward the target feature. Strength of bias was established via extensive piloting and validation, in which healthy older adults rated how strongly a sentence was related to its corresponding target word. On a scale of 1–10, sentences with average ratings between 8–10 were defined as strongly related, and those with average ratings of 5–7.9 were defined as moderately related. Mean ratings were 8.80 and 6.67 for strong- and moderate-bias sentences, respectively.

A treatment hierarchy was created to exploit the use of contextual bias to prime the target semantic features. In the first level (Strong bias), both context sentences were presented prior to the probe sentence and target: “She heard the melody. She forgot the words. There was a piano – SONG.” If a participant’s (accurate) response was sufficiently fast, the strongly-biased sentence was removed, and the probe sentence was preceded only by the moderate-bias sentence: “She forgot the words. There was a piano – SONG.” If the response again met RT criterion, then only the probe sentence was presented prior to the target phoneme string. If the response once more met criterion, the treatment moved back up the hierarchy, and onto the next item.

If the participant responded too slowly at the strongest level of contextual bias, the full item was repeated. If the response still did not meet criterion the treatment moved to the next

item. If a slow response was given at the other levels of the hierarchy, the treatment moved back up the hierarchy one level (e.g., a too slow response on the moderate biasing level triggered a re-presentation of the strongly biasing level). Sufficiently fast responses triggered a move back down the hierarchy.

The RT criterion for progression through the treatment hierarchy was calculated individually for each participant and each treatment list. Specifically, the target RT for a treatment list was the average of a participant's RTs, during the baseline phase of the study, to the experimental items within the probe list associated with that treatment list.

Participants were instructed that they would hear "up to 2 sentences" preceding the final sentence and target word. They were not instructed to pay attention to these additional sentences; instead, they were told to answer as quickly as possible whether the target was a real word. The word monitoring task also was included in the treatment solely to induce participants to carefully listen to the stimuli. As with the diagnostic task, responses to word monitoring were consistently accurate and will not be discussed further.

Probing schedule: An outline of the probing schedule is provided in Appendix B. One of the three versions of the experimental list that was entered into treatment was probed at the beginning of each treatment session. The second experimental list was scheduled for probing once at the estimated midpoint of treatment (estimated based on each participants' daily performance on the treated list) and then was probed in the maintenance phase until stability was obtained. List 3 was probed approximately every third session during treatment.

Treatment methods: Treatment sessions were scheduled 2–3 times per week. After all scheduled probes, treatment commenced. Each treatment list included 15 experimental items and 10 filler items arranged according to the same criteria used for probe lists. The order of stimuli in a treatment list differed from the order in the probe list. Each treatment session lasted approximately 30 minutes, although the total amount of time depended on how many iterations of the treatment hierarchy were triggered for each item.

Treatment sessions continued until %Crit on the associated probe list reached and remained at or above 80% for three consecutive sessions. At that point, the first treated list was put into maintenance, and baseline probing began for the second treatment list. Once stability was obtained, treatment began on that second list and continued in the same manner until %Crit on its associated probe list again reached and remained at 80% over three consecutive sessions.

Follow-up phase—Follow-up sessions were conducted at approximately 1, 2, and 3–4 weeks after completion of treatment on the second of the two lists. During each follow-up session all three lists were probed. The outcome measures were administered in at least one of the three follow-up sessions. Follow-up sessions each lasted approximately 90 minutes.

General procedures—Six examiners were trained to conduct the diagnostic and treatment sessions. All had to achieve at least 80% accuracy on a written test of procedures and demonstrate competency in conducting a session to at least one other team member. All

sessions were video recorded for later reliability analyses. Each participant was tested by no more than two examiners. Testing sessions were conducted either in a quiet room at the participants' homes or in the first author's laboratory at the University of Houston. Equipment set-ups were the same at both sites. Recording of the stimuli as well as EPrime programming and data analysis took place in the second author's laboratory in Pittsburgh.

Outcome measures

Stimulus generalization: The primary measure of generalization to broader outcomes was the Discourse Comprehension Test (DCT; Brookshire & Nicholas, 1993). The DCT consists of short (14 sentence) humorous stories about familiar activities, followed by eight true/false questions. The questions probe details and main ideas that are either explicitly mentioned or implied. There are 10 stories divided into two sets of five. The entire test was administered once during the baseline sessions. Probing in subsequent phases was conducted on only one set, to reduce the potential effects of repetition. The order of set presentation was counterbalanced within and across participants. The dependent variable was the total number of correct responses to the comprehension questions.

To assess generalization to interpretation of non-literal language, posited to be subserved by the partially domain-general coarse coding process, additional measures included the Metaphor Interpretation and the Speech Act Interpretation subtests of the Protocole Montreal d'Evaluation de la Communication (MEC; experimental English version; Joanne, Ska & Côté, 2004). The Metaphor Interpretation task consists of 10 novel metaphors (e.g., The teacher is a sleeping pill) and 10 idioms (e.g., Mary let the cat out of the bag). If participants do not generate a complete, accurate meaning, they are presented with multiple choice options. The Speech Act Interpretation subtest consists of 20 2-sentence vignettes in which a character makes a statement or asks a question. In half of the items the statement/request is direct, and in the other half it is indirect. Participants are informed that some sentences contain an insinuation while others do not. If the participant gives an incomplete or incorrect interpretation of what the character meant, multiple choice options are provided. All items in these subtests were scored in accordance with test scoring guidelines (correct = 2 points, incomplete/imprecise or requiring the multiple choice option=1, incorrect=0). The dependent variable for the Metaphor Interpretation subtest was the total score for the metaphors only (maximum = 20), and for the Speech Act Interpretation subtest only the indirect statements/request items were scored (maximum = 20).

Response generalization: Response generalization to untrained items was measured on a Coarse Coding Generalization Task, which consisted of five probe stimuli that were in the diagnostic set but were not included in treatment. Response generalization would be evident if, after treatment, speed of response on these items exceeded the criterion used in the diagnostic task.

Response generalization to lexical metaphor processing was assessed using the untreated List 3 (see Appendix A). Each sentence in List 3 ended with a homophone with a metaphoric subordinate meaning (e.g., "It was plush"). The target phoneme string reflected the subordinate, metaphorical sense of the word (e.g., FANCY). In order to provide a test of

whether access to metaphoric meanings might be a special case of coarse coding deficit, the relative semantic distance of the metaphoric meaning to the sentence-final word (using Maki, McKinley & Thompson, 2004 norms), and a variety of other lexical properties were matched to the literal features of the items in the treatment lists. Just as for the two treatment lists, the participants' task was a lexical decision, and the correct response to all experimental items was "YES." Filler items were followed by a non-word, for which the correct response was "NO."

Control measures: One of three possible control measures was selected for each participant. Given the heterogeneity of the RHD population and the lack of consistent pattern of deficit co-occurrence, a variety of potential control tasks were administered: the Emotional Prosody Production subtest from the MEC (Joanette et al., 2004); the Judgment of Line Orientation test (JLO; Benton, Hamsher, Varney & Spreen, 1983) and the Visual Form Discrimination task (VFD; Benton, Sivan, Hamsher, Varney & Spreen, 1994). These tasks were administered twice during the baseline sessions to assess stability on the measures because none has a reported standard error of measurement. Ideally, the measure on which a participant demonstrated impairment yet relatively stable performance was selected as the control measure for that individual. However, some participants were impaired on only one of the three measures, and thus it was selected even if performance was quite variable. Post-treatment scores (assessed in the phase change/maintenance period immediately post-treatment and in the follow-up phase) were interpreted in comparison to baseline administrations such that interpretable change was only considered to have occurred if the scores exceeded the variability exhibited in baseline sessions.

Results

Diagnostics

Out of the 15 experimental diagnostic items, Mr. R was too slow/inaccurate on 12 items (5 errors, 7 correct but slow); P101 on all 15 (1 error, 14 correct but slow); P113 on 13 (5 errors, 8 correct but slow); and P206 on 12 of 13 items (2 errors, 10 correct but slow; no response to 2 items).

Probe List Gains

Faster response times on probe lists were expected from the repetition of the probe lists and repeated exposure to the treatment hierarchy. As noted earlier, the primary outcome was narrative comprehension, and the more important and interesting findings relate to generalization to broader outcomes. However, for completeness, and per convention for single subject experimental design, results and figures for probe list gains are provided in Appendix C.

Broader Outcome Measures

Stimulus generalization—Data for all generalization and control measures are in Table 4. The primary generalization measure was percent correct performance on the DCT. Improvement on the task was interpreted only if the changes exceeded the standard error, which was one point (i.e., 1.25%), to account for potential effects of repetition. Changes

over time apparent for the four participants are not likely to be due just to repetition, as the task was divided into two sets of five stories and each set typically was administered only twice throughout the study (a third administration occasionally occurred at final follow-up). Additionally, there was both positive and negative change across participants over time, suggesting that participants were not simply remembering the stories.

Generalization to narrative comprehension varied across participants (see Table 4). Mr. R evidenced improvement from baseline to post-treatment on the overall DCT score and the gains were maintained into the follow-up period. P101 evidenced some improvement from baseline to the first follow-up, with additional gains at the third follow-up. P113 improved from baseline to the first follow-up, but then declined back to baseline levels at the third follow-up. P206 exhibited a decline in DCT scores immediately following treatment, but then evidenced improvement beyond baseline levels by the third follow-up. For all participants, performance on the subset of questions that required an inference followed the same general pattern as the overall scores.

The other two language measures hypothesized to be related to coarse coding were interpretation of phrasal metaphors and indirect speech acts. Neither of these measures have standard error data, so they were administered twice during baseline sessions to obtain an estimate of variability. Participants were particularly variable on repeated administrations of the metaphor subtest, making it difficult to interpret any potential changes related to treatment. P113 was the only participant for whom positive change on the Speech Acts task exceeded the range of variability obtained in the baseline sessions. No clear gains were seen on either measure from baseline to post-treatment or follow-up phases for the other participants.

Response generalization—One measure of response generalization was the Coarse Coding Generalization Task. All four participants evidenced generalization on this task, improving substantially from scores of 0/5 or 1/5 during the diagnostic session. All but P101 achieved 5/5 in at least one post-treatment session (either immediately post-treatment or during follow-up sessions).

Response generalization to lexical metaphor processing was assessed by %Crit on List 3. P101 was at 0% on this list during baseline, along with the two treatment lists. She clearly evidenced improvement related to treatment. For the other three participants, performance on List 3 was quite variable during the baseline sessions, but plateaued during treatment, with a slight increase in %Crit for P113 but little or no change for Mr. R and P206. In broad terms, the effect sizes for List 3 were similar to those for the treated lists for all participants, suggesting that there was only slight generalization to lexical metaphor processing, if any.

Control measures—There were no systematic changes in the control measures for any participant. The control task for Mr. R was the MEC Emotional Prosody production, in which the change seen at follow-up was within the range of variability obtained during baseline. The VFD was the control task for P101 and P113. Performance by these two participants was variable during baseline, and positive changes seen post-treatment were within the range of their baseline variability. The JLO was the control measure for P206. He

showed no change immediately following treatment, when stimulus-and response-contingent generalization was observed. He did exhibit a 5-point increase between the first and third follow-up sessions.

Discussion

There are large gaps in the literature and in our understanding of treatment for language processing deficits associated with RHD. This pre-efficacy study provides evidence of generalization to narrative comprehension from four participants treated with an implicit, contextual constraint stimulation treatment for a basic language processing deficit – coarse coding.

The purpose of CCT is to provide implicit contextual stimulation to facilitate inefficient basic comprehension processes, including coarse coding. This treatment differs from traditional language treatments, both for deficits associated with RHD and aphasia, in several ways. First, it is implicit, tapping more natural comprehension processing by removing metalinguistic and memory demands required in other treatments, such as the metaphor treatment for adults with RHD (Lundgren et al., 2011). Second, the intent is to improve the speed and efficiency of coarse coding with outcomes based on speed rather than accuracy. Third, the overall focus is on a partially domain-general process with the potential to generalize broadly. Thus, the goal is to effect change in narrative comprehension and other broader outcomes.

As expected due to repeated stimulation with the same items over time, participants evidenced faster RTs and higher %Crit on probe lists. The gains are not simply due to repeated exposure to the items, as there was response generalization to untrained stimulus probes. Three of the four participants evidenced 5/5 fast, accurate response to these items following treatment (either immediately post-treatment or during the follow-up). These results indicate that the implicit stimulation-facilitation treatment increased the efficiency of the coarse coding process. Nonetheless, these effect sizes and response generalization data do not get to the heart of the results: that meaningful changes in narrative comprehension were observed for three of the four participants.

The increased efficiency of the domain-general coarse coding process positively affected narrative comprehension. Generalization to the DCT was evident for three of the four participants. In each case, participants evidenced fewer errors on the DCT questions following treatment. There was some variability in performance in the maintenance and follow-up phases, which is not unexpected in single-subject experimental designs. While there were some instances of decreased performance by two participants, it is possible that those changes were due to reduced attention or fatigue during those particular sessions, as the trend over time for all participants was improvement in discourse comprehension. As noted above the improvements on this task are not likely due to repetition effects because all reported changes were greater than the standard error for the test. Mr. R and P101 exhibited improvements immediately after treatment that were maintained through the follow-up period. P113 evidenced gains during follow-up. P206 also evidenced improvement, but not until the final follow-up session. The reason for his post-treatment gains (in narrative

comprehension as well as in the control measure) is unclear, but does not seem to be related to the treatment itself, as the gains did not appear until approximately one month following the end of treatment. P206 was the only participant who had visuospatial neglect, and perhaps the broader attentional deficits associated with neglect somehow affected his response to treatment and/or contributed to his fluctuating performance on several of the outcome measures.

The inability to attribute probe list improvement to treatment (per the CDC method) for 3 of 4 participants raises the question of whether the observed generalization effects would occur without the provision of CCT per se, and with only simple repetition of the probes over time. The strongest argument against the effect of simple repetition is that improvement on untreated items on the CC generalization task was exhibited by all four participants. The first two administrations of this task (in baseline and maintenance phases for most participants) were separated by a minimum of 2 months, reducing the chance that a repetition effect was in play. Additionally, the changes between repeated administrations of the probe lists generally did not exceed 2 out of 15 items, while changes seen in the generalization block ranged from 3–5 items out of 5 possible.

Despite this strong generalization, the effect of repetition is an interesting empirical issue. It is noteworthy that the most impaired participant, who did not improve with repeated exposure during the baseline sessions, was the one for whom probe list gains could be attributed to treatment. Perhaps patients with baseline profiles like hers need the CCT treatment package, and patients who make baseline improvements could make generalized gains more efficiently, without the actual treatment component.

Given the partially domain-general nature of coarse coding, it was predicted that benefits of more efficient coarse coding also would generalize to non-literal language processing. The results for List 3 are inconclusive. The effect sizes are of the same general magnitude as those on the treated lists. Inspection of the figures, however, indicates that while P101 evidenced steady improvement throughout the course of treatment, the other three participants seemed to maintain the level obtained during the baseline phase. It is not possible to determine whether the maintenance is a result of CCT, or if the probing was just frequent enough for the participants to sustain performance over time.

List 3 was specifically designed to assess the hypothesis that non-literal meanings of lexical metaphors may be a special case of coarse coding. The semantic distance between the metaphoric meaning and the core meaning of the sentence-final noun was matched to that of the literal features and core meanings in the treated lists so the absence of strong generalization cannot be attributed to metaphoric meanings that were too peripheral. One possible conclusion is that metaphoric meanings are not processed in the same manner as distant literal features. According to Giora's (1997) Graded Salience Hypothesis, the RH is important for processing novel, unfamiliar metaphoric meanings (e.g., Giora, Zaidel, Soroker, Batori & Kashner, 2000; Mashal, Faust, Hendler & Jung-Beeman, 2007). It is possible that the lexical metaphors were too common (too well-known), and did not rely on RH processing. However, the metaphorical meanings were determined to be of similar semantic distance compared to the literal features in the experimental lists. Additionally, if

the issue was related to familiar, common metaphorical meanings that are processed in the intact left hemisphere, then the participants should not have exhibited any deficits on List 3.

Another possible reason for lack of strong generalization to List 3 results is that metaphoric meanings are more “complex” in some ways than literal meanings. It has been established that treatment of more complex items generalizes to more simple items, but not vice versa (Ballard & Thompson, 1999; Geirut, 2001; Kiran & Thompson, 2003; Thompson, Shapiro & Sobecks, 2003). Perhaps treatment of metaphoric meanings would generalize to distant literal meanings. To test this, future studies may involve treatment of List 3 and assessment of generalization to List 1 or 2.

The predicted generalization to non-linguistic processing, as measured by the MEC subtests of metaphor and indirect speech act interpretation, was not apparent. There are several possible explanations for this. First, the selected subtests require participants to explain what a phrase or statement means. This type of task is clearly metalinguistic, potentially creating extra difficulty for adults with RHD. It is also less natural to explain meaning or intent as opposed to responding to meaning, and likely requires additional processing that was not targeted in the treatment. However, the indirect speech act stimuli were quite conventional and three of the four participants were generally quite accurate at the task. There also appeared to be a learning effect. In the task, if the participant’s response is incomplete or incorrect, multiple choice options are provided. The participants with the lowest initial scores had relatively large increases in the second baseline administration, suggesting that they remembered the presented options and used those to create their answers the second time. Questions about the stability of the measure also come from the substantial variability (up to 15% change) on repeated administrations within the baseline sessions.

In future work, other tasks will be administered to test the domain-generalty of the coarse coding process via treatment effects. One possibility is a test of the generation of tool-based inferences. These inferences occur when comprehenders infer that a particular tool was used although it is not mentioned (e.g., when presented with “he pounded in the nail,” most comprehenders will infer that he used a hammer). Activation of other possible tools (e.g., brick, rock, wrench) that could be used to pound in a nail may be linked to coarse coding, as they are distantly-related to the action. Another possible generalization measure is comprehension of inferential ambiguities (e.g., She slowly walked down the aisle – grocery store vs. wedding interpretation; Tompkins et al., 2001, Tompkins, Fassbinder, Blake, Baumgaertner & Jayaram, 2004), in which multiple possible interpretations are generated. Previous work has suggested that adults with RHD can generate multiple contextually-appropriate alternatives (Tompkins et al., 2001, 2004). However, it is becoming apparent that coarse coding deficits are less common than suppression deficits², so the group results from the studies of inferential ambiguity processing may obscure problems from the few participants with coarse coding deficits. Individuals with both suppression and coarse coding deficits may generate multiple inferences, but still not generate those that are particularly remote or peripheral.

²Our diagnostic procedure involves assessment of both coarse coding and suppression. To date, of the 47 participants tested, 2 had only a coarse coding deficit and 17 had only a suppression deficit. An additional 14 exhibited both deficits.

Importantly, the fact that there was no meaningful change in the control tasks (production of emotional prosody and visuospatial processing) suggests that the effects of the Contextual Constraint Treatment were not simply due to generalized cognitive stimulation. The one exception is P206, who had a large increase in the number of correct items on the JLO in the final follow-up session, along with improved DCT scores. These changes are not likely to reflect a generalized benefit to cognitive processing because they occurred only during the final session which took place approximately one month after the final treatment session. Additionally, there were no corresponding increases in performance for List 2 or 3 in that final testing session, as might be expected if there were generalized improvements.

In summary, the results from the four participants indicate that CCT resulted in generalized gains to narrative comprehension for three of the four participants. As expected, treatment contingent gains were obtained on the probe stimuli and there was strong generalization to untreated stimuli. There was no strong generalization to lexical metaphor processing or to the other non-literal language assessed.

Some limitations of the current study that need to be addressed in future work include the inclusion of a control group with comprehension deficits not related to coarse coding, to determine whether the treatment specifically facilitates coarse coding processes. Additionally, in future studies inclusion criteria should include measures of comprehension, to ensure that all participants indeed have the deficits that the treatment targeted. Finally, a larger set of generalization probes would provide more solid evidence for the efficacy of the treatment.

Several extensions of this treatment are currently underway, including the enrollment of more participants so we can better understand potential individual differences in responses to the treatment. We are also examining dose-response patterns and generalization to aspects of social communication. The number of stimulus repetitions and/or treatment sessions required to achieve treatment gains is important to gauge the efficiency as well as the effectiveness of a treatment. To examine dose-response patterns at a more precise level than number of treatment sessions, the number of “failed” treatment trials – those for which a participant responds inaccurately or too slowly even at the strongest level of contextual constraint – also will be counted. It is predicted that participants who have more failed trials will take longer to reach criterion on the probe task, and will require more trials and sessions to show generalization to discourse comprehension. Examination of broader social communication outcomes, as measured by participant and significant other reports on the Social Relations scale of the Burden of Stroke Scale (BOSS; Doyle et al., 2004, 2007) and Conversational Focus and Conversational Effectiveness factors from the LaTrobe Communication Questionnaire (Douglas, Bracy & Snow, 2007; Douglas, O’Flaherty & Snow, 2000) also is underway. If, as hypothesized, coarse coding is domain-general and affects comprehension processes on multiple levels, then communication breakdowns and misinterpretations that commonly occur after RHD could be positively affected by treatment that facilitates coarse coding processes.

Future work could examine the neural correlates of treatment to explore possible relationships between lesion size/location and response to treatment. The use of functional

brain imaging pre- and post-treatment would help elucidate neural changes that occur as a result of the treatment, and would provide greater understanding of the processes of interest and the effects of the treatment.

Acknowledgments

This work was supported in part by grants DC01820 and DC010182-01 from the National Institutes of Health (National Institute on Deafness and Other Communication Disorders) awarded to the second author.

APPENDIX A. Coarse Coding Diagnostic and Treatment Probe Stimuli for all Probe Lists and Constraint Sentence Contexts for Lists 1 and 2

DIAGNOSTIC STIMULI

1. I saw the cotton – Field.
2. He has an apple – Rotten.
3. He inspected the sofa – Spring.
4. He had a car – Hood.
5. She used the garlic – Powder.
6. He liked the milkshake – Calories.
7. She didn't like the rice – Bland.
8. He ate the potatoes – Fluffy.
9. He drank some coffee – Beans.
10. He visited the castle – Royal.
11. He moved the oak – Furniture.
12. He has a cabin – Cramped.
13. Here is a shirt – Wrinkled.
14. There is the mustard – Plant.
15. She spit out the milk – Spoiled.

COARSE CODING LIST 1

<u>Probe Stimulus</u>	<u>Constraint Sentences (Strong, Moderate)</u>
1. There was an airplane – Captain.	He sat in the cockpit. He announced the final destination.
2. It was a milkshake – Calories.	The drink had extra whipped cream. The drink was bad for her.
3. It was rice – Bland.	The food had no flavor. She forgot to add salt.
4. It was a shirt – Wrinkled.	The garment needed ironing. She smoothed the fabric.
5. There was a castle – Royal.	The king sat on his throne. They put the crown on his head.
6. There was a carrot – Salad.	She cut up the lettuce. She got out some vegetables.
7. It was fish – Smell.	The aroma was too strong. The food was in the garbage.

8. There was a hotel – Dirty.	The resort was filthy. There were bugs in the room.
9. There was a piano – Song.	She played the melody. She forgot the words.
10. There was an owl – White.	The animal blended in with the snow. Its feathers were very clean.
11. It was leather – Couch.	She lay down on the sectional. The upholstery was soft.
12. It was a horse – Wild.	The animal was untamed. The animal ran freely.
13. There was a farm – Dairy.	The cheese was fresh. The products came from an animal.
14. It was tuna – Oil.	The canned fish was packed in canola. She removed the fish from the liquid.
15. It was a lamb – Shave.	He used electric shears on the animal. The fur fell to the ground.

COARSE CODING LIST 2

Probe Stimulus	Constraint Sentences (Strong, Moderate)
1. It was potatoes – Fluffy.	She beat the food until it was airy. The vegetable was whipped with a whisk.
2. It was a sofa – Springs.	The metal squeaked under his weight. He bounced on the furniture.
3. There was mustard – Plant.	She placed the seeds in the soil. She watered the soil.
4. It was cotton – Field.	He plowed an acre of land. The crops needed fertilizer.
5. It was chicken – Baked.	It was in the oven for some time. The skin was golden brown.
6. There was coffee – Beans.	They were picked in order to be ground. The grindings smelled like hazelnut.
7. There was a car – Hood.	The mechanic opened the engine compartment. The engine needed to be looked at.
8. It was a poem – School.	The boy listened to his teacher. She handed him a paper.
9. There was a ship – Pirate.	He had a wooden leg and eye patch. The men seized the passengers.
10. It was soup – Salty.	The meal had a lot of sodium. She added too much seasoning.
11. There was a stone – Beach.	The children built sandcastles. The children wore flip flops.
12. It was a rose – Green.	The flower leaves were the color of grass. The plant was freshly cut.
13. There was a camera – Landscape.	She enjoyed the scenery. She wanted to remember the scenery.
14. There was a cavern – Moist.	The walls were damp. Moss was everywhere.
15. It was pine – Clean.	The bathroom was spotless. The odor on the mop still lingered.

COARSE CODING LIST 3

(Note: Treatment was not conducted on List 3, hence there are no constraint sentences for this list.)

Probe Stimulus

1. It was a bluff – Cards.
2. There was a unit – Hospital.
3. It was a nucleus – Science.
4. It was plush – Fancy.
5. It was a tablet – Paper.
6. It was quiet – Shy.
7. It was respect – Manners.
8. There was a bond – Love.

9. There was scum – Person.
10. There was an outline – Class.
11. It was a flutter – Stomach.
12. It was a branch – Bank.
13. There was a ribbon – Typewriter.
14. There was a server – Tennis.
15. There was a jewel – Helpful.

APPENDIX B. Organization of Study Phases and Tasks Administered

BASELINE PHASE (minimum of 5 sessions): Sample session organization (order of presentation of lists and versions was counterbalanced across sessions and participants)

Session 1 List 1, version 1 – List 2, version 1 – List 3, version 1

Generalization tasks (e.g., DCT Set A, MEC indirect speech acts)

Control tasks (e.g., JLO, VFD, MEC prosody)

Ancillary tasks (e.g., ABCD immediate & delayed memory)

Session 2 List 2, version 3 – List 3, version 3 – List 1, version 3

Generalization tasks (e.g., MEC metaphor)

Ancillary tasks (e.g., BIT, PPVT)

Session 3 List 3, version 2 – List 1, version 2 – List 2, version 2

Generalization tasks (e.g., DCT Set B)

Control tasks (e.g., JLO, VFD, MEC prosody)

Ancillary tasks (e.g., Cookie Theft, working memory)

TREATMENT PHASE: Sample session organization for treatment of List 1

Treatment 1: Probe List 1, version 1 – Treatment List 1, version 1

Treatment 2: Probe List 1, version 3 – Treatment List 1, version 3

Treatment 3: Probe List 1, version 2 – Probe list 3, version 3 – Treatment List 1, version 2

MAINTENANCE PHASE: In the first maintenance session, all three probe lists were administered. For the remaining sessions, only the as-yet-untreated list was probed in each session until stability was obtained.

Maintenance 1: Probe List 1, version 1 – Probe List 2, version 1 – Probe List 3, version 2

Generalization task (e.g., CC generalization, DCT A)

Maintenance 2: Probe List 2, version 3

Control task (JLO, VFD, or MEC prosody, per baseline performance)

Maintenance 3: Probe List 2, version 2

Generalization tasks (e.g., MEC metaphor and indirect speech acts)

FOLLOW-UP PHASE: The counterbalanced order and versions of lists was maintained throughout.

Session 1 List 2, version 1 – List 3, version 1 – List 1, version 1

Generalization tasks (e.g., DCT Set A, MEC indirect speech acts)

Control task (JLO, VFD, or MEC prosody, per baseline performance)

Session 2 List 1, version 2 – List 2, version 2 – List 3, version 2

Generalization tasks (e.g., MEC metaphor)

Session 3 List 3, version 3 – List 1, version 3 – List 2, version 3

Generalization tasks (e.g., DCT Set B)

Control task (JLO, VFD, or MEC prosody, per baseline performance)

APPENDIX C. Probe List Gains

Probe %Crit data for all probe lists for all phases of treatment are provided in Figures 1–4. Two methods were used to interpret the treatment effects (data provided in the Table below). The conservative dual criterion (CDC) method (Fisher, Kelly & Lomas, 2003) controls for Type I error even when there is a high degree of autocorrelation. Using the CDC method, two criterion lines are created from the baseline data: a mean line and a trend line. Both lines are then adjusted by 0.25 standard deviations (SD calculated from the baseline data) in the expected direction – in this case, upward, to reflect a positive treatment effect – and are extended into the treatment phase on the graph. Fisher et al. (2003) provide criteria for the number of data points that must fall above both lines in order to conclude that positive change on the probe items is systematic and associated with the treatment.

For the second method, effect sizes for change from baseline to treatment and from baseline to the follow-up were calculated using the d-Index (Bloom, Fischer & Orme, 2003). This statistic uses the pooled standard deviation from baseline, treatment, and follow-up phases, which provides a relatively conservative estimate of effect size. The true/original baseline, which includes all probes prior to the start of any treatment with any list, was used. Beeson and Robey's (2006) recommendations for interpreting effect sizes (2.6=small, 3.9=medium, and 5.8=large) are used as a starting point, however they are interpreted cautiously given that they are based on treatments for aphasia in which accuracy is the most important variable, and, more importantly, because the focus of the current study is on generalization outcomes.

Experimental probe lists %Crit

All but P101 demonstrated some improvement (higher %Crit) during the baseline phase, but our stability criterion was obtained for all participants on at least one of the two treatment lists within 5–8 sessions. All participants made additional gains in %Crit once treatment began.

As shown in the table below, for the first treatment list, small to medium effect sizes were obtained for three of the participants. Effect sizes for P101 are inflated due to the lack of variability in baseline performance. Based on the CDC method, only P101 had enough data points above both lines to conclude that the systematic change in probe item responses was associated with the application of treatment. For the other participants, changes could not be unequivocally associated with treatment due to the rising trends in the baseline data.

Treatment on the second list began after our stability criterion was met on that list during the maintenance phase. Mr. R was continuing to improve %Crit on List 2 throughout the maintenance sessions, and as he was at or exceeding the 80% level, and (more importantly) had shown improvement on the DCT, treatment was not initiated for List 2. Small to medium effect sizes were obtained for P113 and 206, and an inflated effect again was apparent for P101 due to lack of baseline variability. As for the first treated list, only P101 had enough data points above both CDC lines to suggest that the systematic change in probe performance was associated with the treatment.

Three follow-up sessions were conducted. In the first session, all participants had %Crit similar to the last treatment session, although P206's decreased slightly. The pattern of the remaining follow-up session probes varied across the four participants. The effect sizes for baseline to follow-up were small to moderate for most of the participants, and inflated for P101.

*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 three probe lists across experimental design phases. Adapted from Tompkins, C.A., Scharp, V.L., Meigh, K., Blake, M.L., & Wambaugh, J.L. (2012). Generalization of a novel, implicit treatment for coarse coding deficit in right hemisphere brain damage: A single subject experiment. *Aphasiology*, 26, 689–708. (Taylor & Francis Ltd, www.tandfonline.com). Reprinted by permission of Taylor and Francis.

Figure 2.

P101's performance on three probe lists across experimental design phases. Adjusted level and trend lines calculated per the conservative dual criterion (CDC) method (Fisher et al., 2003).

Figure 3. P113's performance on three probe lists across experimental design phases. Adjusted level and trend lines calculated per the conservative dual criterion (CDC) method (Fisher et al., 2003).

Figure 4.

P206's performance on three probe lists across experimental design phases. Adjusted level and trend lines calculated per the conservative dual criterion (CDC) method (Fisher et al., 2003).

Appendix C Table.d-Index and CDC statistics for baseline to treatment and baseline to follow-up comparisons, for three treatment lists

	Mr. R	101	113	206
d-Index ^a				
1 st list Baseline-Tx	2.52	34.2	2.2	4.2
2 nd list Baseline-Tx	N/A *	12.4	2.7	3.3
1 st list Baseline-FollowUp	1.66	34.5	2.1	3.6
2 nd list Baseline-FollowUp	N/A *	8.7	1.8	3.0
List 3 Baseline-FollowUp	1.54	9.9	1.9	1.9
CDC method data ^b				
1 st List actual (criterion)		12/12 (9/12)	0/12 (9/12)	2/6 (6/6)
2 nd List actual (criterion)		6/7 (6/7)	2/3 (N/A **)	0/9 (8/9)

Note. 1st list, 2nd list refer to first and second treated lists, respectively.

^aBloom, Fischer & Orme, 2003

^bFisher, Kelly & Lomas, 2003

* Required treatment on only one list.

** Minimum of 5 data points needed to calculate CDC lines

References

- Anaki D, Faust M, Kravetz S. Cerebral hemisphere asymmetries in processing lexical metaphors. *Neuropsychologia*. 1998; 36:691–700. [PubMed: 9723940]
- 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]
- Ballard KJ, Thompson CK. Treatment and generalization of complex sentence production in agrammatism. *Journal of Speech, Language, and Hearing Research*. 1999; 42:690–707.
- Balota, DA.; Cortese, MJ.; Hutchison, KA.; Neely, JH.; Nelson, D.; Simpson, GB., et al. The English Lexicon Project: A web-based repository of descriptive and behavioral measures for 40,481 English words and nonwords. Washington University; 2002. <http://lexicon.wustl.edu/>
- Bayles, KA.; Tomoeda, CK. Arizona Battery for Communication Disorders of Dementia. Tucson, AZ: Canyonlands Publishing; 1993.
- Beeman M. Semantic processing in the right hemisphere may contribute to drawing inferences from discourse. *Brain and Language*. 1993; 44:80–120. [PubMed: 8467379]
- Beeman, M. Coarse semantic coding and discourse comprehension. In: Beeman, M.; Chiarello, C., editors. *Right hemisphere language comprehension: Perspectives from cognitive neuroscience*. Mahwah, NJ: Lawrence Erlbaum Associates Inc; 1998. p. 255-284.
- Beeman, M.; Chiarello, C. *Right hemisphere language comprehension: Perspectives from cognitive neuroscience*. Mahwah, NJ: Lawrence Erlbaum Associates Inc; 1998.
- Beeson PM, Robey RR. Evaluating single-subject treatment research: Lessons learned from the aphasia literature. *Neuropsychological Review*. 2006; 16(4):161–169.
- Benton E, Bryan K. Right cerebral hemisphere damage: Incidence of language problems. *International Journal of Rehabilitation Research*. 1996; 19:47–54. [PubMed: 8730543]
- Benton, AL.; Hamsher, KS.; Varney, NR.; Spreen, O. *Contributions to neuropsychological assessment*. New York: Oxford University Press; 1983. Judgment of Line Orientation; p. 44-54.

- Benton, AL.; Sivan, AB.; Hamsher, KS.; Varney, NR.; Spreen, O. Contributions to Neuropsychological Assessment. 2. New York: Oxford University Press; 1994. Visual Form Discrimination; p. 65-72.
- Blake, ML. Cognitive Communication Disorders Associated with Right Hemisphere Brain Damage. In: Kimbarow, ML., editor. Cognitive Communication Disorders. Plural Publishers; 2011. p. 119-168.
- 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; 19:423–434.
- Bloom, M.; Fischer, J.; Orme, JG. Evaluating practice—Guidelines for the accountable professional. 4. Boston, MA: Allyn & Bacon, Pearson Education; 2003.
- Brookshire, RH.; Nicholas, LE. Discourse Comprehension Test. Tucson: Communication Skill Builders; 1993.
- Brownell HH, Potter HH, Bihrlé AM, Gardner H. Inference deficits in right brain-damaged patients. *Brain and Language*. 1986; 27:310–321. [PubMed: 3955344]
- Burgess C, Simpson GB. Cerebral hemispheric mechanisms in the retrieval of ambiguous word meanings. *Brain and Language*. 1988; 33:86–103. [PubMed: 3342321]
- Byiers BJ, Reichle J, Symons FJ. Single-subject experimental design for evidence-based practice. *American Journal of Speech Language Pathology*. 2012; 21:397–41. [PubMed: 23071200]
- Chiarello, C. Parallel system for processing language: Hemispheric complementarity in the normal brain. In: Banich, MT.; Mack, M., editors. Mind, brain & language: Multidisciplinary perspectives. Erlbaum; Mahwah, NJ: 2003. p. 229-247.
- Côté H, Payer M, Giroux F, Joanne Y. Towards a description of clinical communication impairment profiles following right-hemisphere damage. *Aphasiology*. 2007; 21(6–8):739–749.
- Conroy, Sage; Lambon, Ralph. Errorless and errorful therapy for verb and noun naming in aphasia. *Aphasiology*. 2009; 23:1311–1337.
- Coulson S, Van Petten C. Conceptual integration and metaphor: An event-related potential study. *Memory & Cognition*. 2002; 30(6):958–968. [PubMed: 12450098]
- Cruikshanks KJ, Wiley TL, Tweed TS, Klein BE, Klein R, Mares-Perlman JA, Nondahl DM. Prevalence of hearing loss in older adults in Beaver Dam, Wisconsin. The Epidemiology of Hearing Loss Study. *American Journal of Epidemiology*. 1998; 148(9):879–86. [PubMed: 9801018]
- Douglas JM, O’Flaherty CA, Snow PC. Measuring perception of communicative ability: The development and evaluation of the La Trobe communication questionnaire. *Aphasiology*. 2000; 14(3):251–268.
- Douglas JM, Bracy CA, Snow PC. Exploring the factor structure of the La Trobe Communication Questionnaire: Insights into the nature communication deficits following traumatic brain injury. *Aphasiology*. 2007; 21(12):1181–1194.
- Doyle PJ, McNeil MR, Mikolic JM, Prieto L, Hula WD, Lustig AP, Elman RJ. The Burden of Stroke Scale (BOSS) provides valid and reliable score estimates of functioning and well-being in stroke survivors with and without communication disorders. *Journal of Clinical Epidemiology*. 2004; 57:997–1007. [PubMed: 15528050]
- Doyle PJ, McNeil MR, Bost JE, Ross KB, Wambaugh JL, Hula WD, Mikolic JM. The Burden of Stroke Scale (BOSS) provided valid, reliable, and responsive score estimates of functioning and well-being during the first year of recovery from stroke. *Quality of Life Research*. 2007; 16:1389–1398. [PubMed: 17665314]
- Dubno JR, Lee F, Klein AJ, Matthews LJ, Lam CF. Confidence limits for maximum word-recognition scores. *Journal of Speech and Hearing Research*. 1995; 38:490–502. [PubMed: 7596114]
- Dunn, L.; Dunn, L. Peabody Picture Vocabulary Test-Revised. Circle Pines, MN: American Guidance Service; 1981.
- Dunn, LM.; Dunn, LM. Peabody Picture Vocabulary Test: Third Edition. Circle Pines, MN: American Guidance Service; 1997.

- Fassbinder W, Tompkins C. Hemispheric differences in word-meaning processing: Alternative interpretations of current evidence. *Aphasiology*. 2006; 20(2/3/4):110–122. [PubMed: 20054427]
- Fillingham JK, Hodgson C, Sage K, Lambon Ralph MA. The application of errorless learning to aphasic disorders: A review of theory and practice. *Neuropsychological Rehabilitation*. 2003; 13(3):337–63. [PubMed: 21854317]
- Fillingham, Sage; Lambon, Ralph. The treatment of anomia using errorless learning. *Neuropsychological Rehabilitation*. 2006; 16(2):129–54. [PubMed: 16565031]
- Fisher WW, Kelley ME, Lomas JE. Visual aids and structured criteria for improving visual inspection and interpretation of single-case designs. *Journal of Applied Behavior Analysis*. 2003; 36:387–406. [PubMed: 14596583]
- Gierut JA. Complexity in phonological treatment: Clinical factors. *Language, Speech, and Hearing Services in Schools*. 2001; 32:229–241.
- Giora R. Understanding figurative and literal language: The graded salience hypothesis. *Cognitive Linguistics*. 1997; 7:183–206.
- Giora R, Zaidel E, Soroker N, Batori G, Kashner A. Differential effect of right- and left-hemisphere damage on understanding sarcasm and metaphor. *Metaphor and Symbol*. 2000; 15:63–83.
- Gonzalez-Rothi LJ. Cognitive rehabilitation: The role of theoretical rationales and respect for the maturational process needed for our evidence. *Journal of Head Trauma Rehabilitation*. 2006; 21:194–197. [PubMed: 16569992]
- 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.; Goulet, P.; Hannequin, D. Right hemisphere and verbal communication. New York: Springer-Verlag; 1990.
- Joanette, Y.; Ska, B.; Côté, H. Protocole Montréal d' évaluation de la communication(MEC). Isbergues: Ortho Édition; 2004.
- Jung-Beeman M. Bilateral brain processes for comprehending natural language. *Trends in Cognitive Sciences*. 2005; 9:712–718.
- Kempler D, Van Lancker D, Merchman V, Bates E. Idiom comprehension in children and adults with unilateral brain damage. *Developmental Neuropsychology*. 1999; 15:327–349.
- Kessels RP, deHaan EH. Implicit learning in memory rehabilitation: a meta-analysis on errorless learning and vanishing cues methods. *Journal of Clinical & Experimental Neuropsychology*. 2003; 25(6):805–14. [PubMed: 13680458]
- Kiran S, Thompson CK. The Role of Semantic Complexity in Treatment of Naming Deficits: Training Semantic Categories in Fluent Aphasia by Controlling Exemplar Typicality. *Journal of Speech, Language, and Hearing Research*. 2003; 46:773–787.
- Klepousniotou E, Baum SR. Processing homonymy and polysemy: Effects of sentential context and time-course following unilateral brain damage. *Brain and Language*. 2005; 95:365–382. [PubMed: 16298667]
- Kratochwill, TR.; Hitchcock, J.; Horner, RH.; Levin, JR.; Odom, SL.; Rindskopf, DM.; Shadish, WR. Single-case designs technical documentation. 2010. Retrieved from What Works Clearinghouse website: http://ies.ed.gov/ncee/wwc/pdf/wwc_scd.pdf
- 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.
- Leonard CL, Waters GS, Caplan D. The use of contextual information by right-brain damaged individuals in the resolution of ambiguous pronouns. *Brain and Language*. 1997a; 57:309–342. [PubMed: 9126419]
- Leonard CL, Waters GS, Caplan D. The use of contextual information related to general world knowledge by right brain-damaged individuals in pronoun resolution. *Brain and Language*. 1997b; 57:343–359. [PubMed: 9126420]
- Lundgren K, Brownell H, Cayer-Meade C, Milione J, Kearns K. Treating metaphor interpretation deficits subsequent to right hemisphere brain damage: Preliminary results. *Aphasiology*. 2011; 25:456–474. [PubMed: 22837588]

- Maki WS, McKinley LN, Thompson AG. Semantic distance norms computed from an electronic dictionary (WordNet). *Behavior Research Methods, Instruments, & Computers*. 2004; 36(3):421–431.
- Mashal N, Faust M. Right hemisphere sensitivity to novel metaphoric relations: Application of the signal detection theory. *Brain & Language*. 2008; 104:103–112. [PubMed: 17445878]
- Mashal N, Faust M, Hendler T, Jung-Beeman M. An fMRI investigation of the neural correlates underlying the processing of novel metaphoric expressions. *Brain and Language*. 2007; 100:115–126. [PubMed: 16290261]
- Middleton EL, Schwartz MF. Errorless learning in cognitive rehabilitation: A critical review. *Neuropsychological Rehabilitation*. 2012; 22:138–168. [PubMed: 22247957]
- Myers, PS. Profiles of communication deficits in patients with right cerebral hemisphere damage. In: Brookshire, RH., editor. *Clinical Aphasiology: Conference proceedings*. Minneapolis, MN: BRK; 1979. p. 38-46.
- Myers, PS.; Linebaugh, CW. Comprehension of idiomatic expressions by right-hemisphere-damaged adults. In: Brookshire, RH., editor. *Clinical Aphasiology: Conference proceedings*. Minneapolis, MN: BRK; 1981. p. 254-261.
- 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]
- Richards L, Chiarello C. Depth of associated activation in the cerebral hemispheres: Mediated versus direct priming. *Neuropsychologia*. 1995; 33:171–179. [PubMed: 7746362]
- Robey RR. A five-phase model for clinical-outcome research. *Journal of Communication Disorders*. 2004; 37:401–411. [PubMed: 15231420]
- Robey RR, Schultz MC. A model for conducting clinical-outcome research: An adaptation of the standard protocol for use in aphasiology. *Aphasiology*. 1998; 12:787–810.
- Schneider, W.; Eschman, A.; Zuccolotto, A. E-prime user's guide. Pittsburgh, PA: Psychology Software Tools Inc; 2002.
- Thompson CK. Single subject controlled experiments in aphasia: The science and state of the science. *Journal of Communication Disorders*. 2006; 39:266–291. Published online April 10, 2006. PMID: PMC1847620 NIHMSID: NIHMS18308. 10.1016/j.jcomdis.2006.02.003 [PubMed: 16635494]
- Thompson C, Shapiro S, Sobecks J. The role of syntactic complexity in treatment of sentence deficits in agrammatic aphasia: The complexity account of treatment efficacy (CATE). *Journal of Speech, Language, and Hearing Research*. 2003; 46(3):591–607.
- Tillman, T.; Carhart, R. An expanded test for speech discrimination utilizing CNC monosyllabic words. Brooks Air Force Base, TX: USAF School of Aerospace Medicine, Aerospace Medical Division (AFSC); 1966. (Northwestern University Auditory Test No. 6) (Technical Report No SAM-TR-66-55)
- Titone D. Hemispheric differences in context sensitivity during lexical ambiguity resolution. *Brain and Language*. 1998; 65:361–394. [PubMed: 9843609]
- 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*. 1991; 34:820–830. [PubMed: 1956190]
- Tompkins CA, Baumgaertner A, Lehman MT, Fassbinder W. Mechanisms of discourse comprehension impairment after right hemisphere brain damage: Suppression and enhancement 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, Bloise CG, Timko ML. Working memory and inference revision in brain damaged adults. *Journal of Speech and Hearing Research*. 1994; 37:896–912.

- 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; 47:1380–1395.
- 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; 44:400–415.
- Tompkins CA, Scharp VL, Gibbs Scott A, Meigh KM. Suppression in polarized ambiguity processing after right hemisphere brain damage. (in preparation).
- Tompkins CA, Scharp VL, Meigh KM, Fassbinder W. Coarse coding and discourse comprehension in adults with right hemisphere brain damage. *Aphasiology*. 2008; 22:204–223. [PubMed: 20037670]
- Tompkins CA, Scharp VL, Meigh K, Blake ML, Wambaugh JL. Generalization of a novel, implicit treatment for coarse coding deficit in right hemisphere brain damage: A single subject experiment. *Aphasiology*. 2012; 26:689–708. [PubMed: 22837589]
- Wilson, B.; Cockburn, J.; Halligan, P. *Behavioural Inattention Test*. Titchfield, England: Thames Valley Test Company; 1987.

Table 1

Demographic and stroke data for 4 participants.

ID	Age	Education	Sex	MPO	Lesion Characteristics
Mr.R	75	10	M	72	Posterior limb of the internal capsule and centrum semiovale of parietal lobe
101	79	12	F	87 (1 st stroke) 19 (2 nd stroke)	Stroke 1: temporal, occipital/PCA distribution/ thalamus
113	85	16	M	66	Stroke 2: MCA distribution/temporal lobe
206	88	18	M	68	MCA distribution: majority basal ganglia; minimal involvement of frontal, parietal, temporal lobes Temporal, frontal, and parietal lobes; old left basal ganglia lacune

Note: ID = Participant identification number; Age, Education in years; MPO = Months post-onset; Lesion Characteristics from MRI/CT scan report, all right cerebral hemisphere unless otherwise noted.

Table 2

Clinical characteristics of the four participants and comparison groups.

	Mr.R	P101	P113	P206	RHD* N=32	NBD* N=38
Auditory Working Memory for Language ^a						
Word Recall errors (max 42)	15	21	18	16	13.2(7.0)	5 (4.6)
Behavioural Inattention Test ^b (max 146, neglect cutoff 129)	146	137	129	111	137 (13.5)	144 (2.8)
ABCD Story Recall ^c (max 17)						
Immediate Retell	15	11	13	10	13.2 (2.5)	14.4 (2.1)
Delayed Retell	14	13	12	12	12.7(3.1)	13.6 (2.5)
PPVT-III ^d raw score (max 204)	162	145	189	188	157(11.3)	163 (11.2)
PPVT-R ^e raw score (max 175)						

* Means and standard deviations of RHD and NBD participants from prior studies (Tompkins, Scharp et al., 2008)

^aTompkins, Bloise & Timko (1994).

^bWilson, Cockburn & Halligan (1987)

^cABCD = Arizona Battery for Communication Disorders in Dementia; Bayles & Tomoeda (1993)

^dPPVT-III = Peabody Picture Vocabulary Test-3rd Edition; Dunn & Dunn (1997)

^ePPVT-R = Peabody Picture Vocabulary Test -Revised; Dunn & Dunn (1981)

Table 3

Treatment stimulus sample.

She played the melody	strong bias sentence
She forgot the words	moderate bias sentence
There was a piano – SONG	probe stimulus

Table 4

Results for broader outcome measures for four participants across experimental phases.

	Dx	Baseline	Maintenance	FollowUp1	FollowUp2	FollowUp3
Mr.R						
DCT Total		55/80 (69%)	31/40 (78%)*		62/80 (78%)*	
DCT Implied		27/40 (67.5%)	15/20 (75%)*		30/40 (75%)*	
MEC Speech Acts (of 20)		18, 20	20		20	
MEC Metaphor (of 20)		10, 17			16	
CC Generalization	0/5				5/5	
MEC emotional (control) (of 10)		2, 4		4	5	
PI01						
DCT Total		61/80 (76%)	35/40 (88%)*	32/40 (80%)*		34/40 (85%)*
DCT Implied		27/40 (68%)	18/20 (90%)*	14/20 (70%)*		16/20 (80%)*
MEC Speech Acts (of 20)		18	19	19		
MEC Metaphor (of 20)		15, 17	17	15		
CC Generalization	0/5		3/5	3/5		
VFD (control) (of 32)		20, 29	28	31		24
PI13						
DCT Total		71/80 (89%)	35/40 (88%)	40/40 (100%)*	37/40 (93%)*	36/40 (90%)*
DCT Implied		34/40 (85%)	16/20 (80%)	20/20 (100%)*	17/20 (85%)*	17/20 (85%)*
MEC Speech Acts (of 20)		19, 19	18	20		
MEC Metaphor (of 20)		18, 18	16	17		16
CC Generalization	1/5		5/5	3/5		
VFD (control) (of 32)		16, 20	22	21	24	18
P206						
DCT Total		61/80 (76%)	27/40 (68%)	30/40 (75%)*		32/40 (80%)*
DCT Implied		28/40 (70%)	14/20 (70%)	13/20 (65%)*		17/20 (85%)*
MEC Speech Acts (of 20)		15, 18	16	15		

	Dx	Baseline	Maintenance	FollowUp1	FollowUp2	FollowUp3
MEC Metaphor (of 20)		12, 17	17	18		
CC Generalization	0/5		3/5	5/5		
JLO (control) (of 30)		18, 20	21	22		27^

Notes: Dx = diagnostic session; DCT = Discourse Comprehension Test (Brookshire & Nicholas, 1993); MEC = *Protocole Montréal d' évaluation de la communication* (Joanette et al., 2004); emotional = emotional prosody repetition; CC = coarse coding; VFD = Visual Form Discrimination task (Benton et al., 1994); JLO = Judgment of Line Orientation (Benton et al., 1983)

* Change exceeding standard error

^ Change exceeding baseline variability