Treating metaphor interpretation deficits subsequent to right hemisphere brain damage: Preliminary results

By: Kristine Lundgren, Hiram Brownell, Carol Cayer-Meade, Janet Milion, Kevin Kearns

Lundgren, K., Brownell, H., Cayer-Meade, C., Milione, J., & Kearns, K. (2011). Treating Metaphor Interpretation Deficits Subsequent to Right Hemisphere Brain Damage: Preliminary Results. *Aphasiology, 25(4)*, 1-19. doi: 10.1080/02687038.2010.500809.

Abstract:

Aims: This investigation sought to determine whether a structured intervention focused on improving use of semantic associations could improve patients' ability to provide oral interpretations of metaphors following right hemisphere damage (RHD).

Methods & Procedures: Principles of single participant experimental design provided the basis for the study. Five patients received either 10 or 20 baseline assessments of oral metaphor interpretation and, as a control, assessments of line orientation skill. They then received approximately 10 one-hour sessions of structured intervention to improve oral metaphor interpretation followed by post-training assessments and a 3-month follow up.

Outcomes & Results: Patients' performances revealed evidence of good response to training as shown by patients' ability to reach criterion on all intervention tasks and by their significant improvement on oral metaphor interpretation. There was relatively little improvement on the line orientation task.

Conclusions: The results of this study support the clinical usefulness of this new approach to treating communication deficits associated with RHD due to stroke, even years post-onset. There are, however, questions that remain unanswered. For example, additional data will be needed to gauge how a patient's severity of impairment relates to the potential for improvement, to chart the durability and scope of improvement associated with the training, and to determine the type of visuospatial ability needed for using this type of pictorial material.

Keywords: Right hemisphere disorder | Stroke | Treatment | Metaphor | Cognitive-linguistic deficits

Article:

Communication often involves interpretation of figurative language in order to understand a speaker's intended meaning. For example, a co-worker might provide a warning that the boss is "growling like a rabid dog" that day. If a listener is unable to infer the intended meaning of this utterance, he or she may not realise that the boss is in a very bad mood and that any request for a raise or for extra time off should be postponed. Consistently failing to apprehend figurative meaning represents a significant communication problem. In this paper we examine one form of figurative language: metaphor. We begin with a brief and selective review to highlight some of the effects of brain injury on patients' metaphor processing. In the major portion of the paper we present a rationale and a description of a novel programme for training skills to support processing metaphor. We identify some patients with RHD and deficits in metaphor processing and evaluate their response to a new training protocol.

Several decades of research have established that damage to the right hemisphere (RHD) can impair individuals' performance in a host of language-relevant tasks. Individual studies have documented impairments involving comprehension of humour (Happe, Brownell, & Winner, 1999; Shammi & Stuss, 1999), irony (McDonald, 1999; Winner, Brownell, Happe, Blum, & Pincus, 1998), indirect requests (Brownell & Stringfellow, 1999; Stemmer, Giroux, & Joanette, 1994), prosody (Baum & Dwivedi, 2003; Lundgren, Moya, & Bennowitz, 1984), and proverb interpretation (Van Lancker, 1990). Several reviews cite a range of linguistic deficits that can result even when the basic building blocks of language are not obviously impaired (e.g., Beeman, 1998; Joanette, Goulet, & Hannequin, 1990; Kempler, 2005; McDonald, 1993; Myers, 1999; Tompkins, 1995). Of particular interest are deficits in comprehension of figurative language, particularly idioms and metaphor (e.g., Brownell, Simpson, Bihrle, Potter, & Gardner, 1990; Myers & Linebaugh, 1981; Van Lancker & Kempler, 1987; Winner & Gardner, 1977), and how best to understand these impairments in theoretical terms.

Early reports suggested that adults with acquired RHD have difficulty with metalinguistic processing of domain-general meaning. Gardner and Denes (1973), for example, observed a selective reduction in sensitivity to connotative dimensions of meaning associated with RHD rather than with left hemisphere damage (LHD). Their task involved matching connotative values across widely disparate domains, for example matching a verbal concept like "wealth" with a drawing of an arrow pointing up rather than down. Brownell, Potter, Michelow, and Gardner (1984) examined patients' use of connotative meaning carried by adjectives such as "warm" and "deep". The major findings were that the patients with RHD based their judgements relatively less often on connotation (e.g., a shared positive association for "warm" and "wise") or metaphoric equivalence (e.g., between "deep" and "wise") than patients with LHD (with aphasia) and the non-brain-damaged participants.

Other studies of patients with RHD have reported problems in interpreting frozen metaphors and idioms such as "a heavy heart". Patients with RHD were relatively less likely to select the conventional metaphoric interpretations and more likely to opt for literal meanings, even though one might expect that over-learned phrases would be stored as a single lexical unit and, for that reason, would be easy to understand in figurative terms (Myers & Linebaugh, 1981; Van Lancker & Kempler, 1987; Winner & Gardner, 1977). The explanations for these abnormal performances are varied. One possibility is that the abnormal performances of patients with RHD on figurative language tasks are due to a general deficit in processing non-literal language. Another is that deficient performances may, to some degree, reflect problems using context to guide interpretation of utterances (e.g., Weylman, Brownell, Roman, & Gardner, 1989). Despite the range of accounts, the basic deficit has been shown for many patients using a variety of language materials.

Among the examples of language impairment exhibited by many patients with RHD, metaphor provides an excellent foundation for examining the potential of cognitive-linguistic training. A conceptual analysis of metaphor makes it relatively straightforward to dissect instances of metaphor into constituent semantic components and then to illustrate how the parts fit together. As will be shown below, the transparency of the analysis supports efforts to train component skills.

The theoretical analysis of metaphor that we use grows out of the domains interaction approach (Hillson & Martin, 1994; Tourangeau & Sternberg, 1981). On this view the communicative impact of a metaphor derives, first, from the discrepancy between semantic domains for the topic and vehicle, and second from the degree to which the topic and vehicle have similar values on some semantic dimension(s) relevant to both, i.e., the ground for the metaphor. Often the ground rests on connotative meaning that can apply across different semantic domains (e.g., Aitchison, 1987; Brown, 1958; Lyons, 1968; Osgood, Suci, & Tannenbaum, 1957).

Consider the example “Microsoft Corporation is the tiger of the business world”. This metaphor links a disparate topic (Microsoft Corporation) and vehicle (tiger) on the basis of a shared ground (ferocity, tenacity) that is, by definition, applicable across different semantic domains. First, greater discrepancy between the semantic domains of a topic and vehicle results in a better metaphor. “Microsoft is the tiger of the business world” is more compelling than “the barracuda is the tiger of the fish world” or, worse yet, “the mountain lion is the tiger of North American cats”. Once a discrepancy between topic and vehicle is noted, problem solving can be used to identify the basis of similarity. Connotations associated with tiger, for example, include elements such as strength, energy, and especially ferocity that can apply within a wide variety of superordinate concepts from very different semantic domains. Graded variation along the connotative dimension of ferocity, for example, can carry distinctions within category concepts as different as animals, people, corporations, and weather patterns. A particularly apt metaphor is one that forces a listener to bridge two very dissimilar domains and to appreciate an especially close equivalence between a topic and vehicle, once that initial discrepancy is acknowledged.

An analysis of these component elements of metaphor suggests processing components that might lend themselves to remediation. Another core idea is choosing from among associations. Consider, as an example, the concept *lawyer*, which elicits strong positive and negative feelings. Those who have had good experiences with lawyers may report associations such as smart, powerful, protector, hard working, justice, advocacy, and good. Those who have had less-positive experiences with a lawyer may generate a different set of associations: aggression, loss, threat, danger, and bad. To construct a metaphor one needs a second noun from a different semantic domain to elicit associations that a speaker seeks to emphasise. Associations to the concept *lion*might include strong, fierce, proud, dignified, brave, jungle, or good. The metaphor “The lawyer is a lion” might convey a strong positive feeling about a lawyer who protects a client from attack. A speaker who has been successfully sued recently might base a metaphor on a different concept such as *shark*, often associated with qualities of aggressiveness and viciousness. For both metaphors, the foundation for comprehension for a listener is awareness of the range of associations to the two nouns. A second important process involves selection. For example, an association with jungles and Africa applies to lions but is not relevant to lawyers or to the metaphor. Furthermore, a listener must select one or more associations that apply to both nouns and that convey a meaning relevant to the context and consistent with the speaker's intended meaning.

In a series of papers Beeman has offered an account for the collection of language-related findings associated with RHD that fits well with the analysis of metaphor outlined above. He suggests that the intact right hemisphere supports “coarse-grained” semantic processing while the intact left hemisphere supports “fine-grained” processing (for reviews see Beeman, 1998; Jung-Beeman, 2005; see also Burgess & Simpson, 1988; Chiarello, 1998; Shibata, Abe, Terao, & Miyamoto, 2007; Stringaris, Medford, Giampetro, Brammer, & David, 2007). The left hemisphere, according to Beeman's model, is best suited for literal language processing. On receipt of a word the left hemisphere mediates focused activation of the limited subset of semantic content that is tied to the most frequent meaning or to the immediate linguistic context. The activation of context-relevant meaning keeps that meaning available for subsequent language comprehension. This processing focus of the left hemisphere does not facilitate the integration of meaning from different domains, as is required for most metaphor, even though it (the LH's processing focus) is ideal for understanding many other aspects of language. In contrast, on receipt of a word the right hemisphere characteristically supports diffuse activation related to the word's semantic content, including even weakly associated components that can apply across semantic domains.

A second, critical part of Beeman's model is that the right hemisphere appears good at maintaining the activation over a relatively long period of time. On receipt of a potential metaphoric topic such as “Microsoft Corporation”, weak activation spreads to many elements. Upon hearing a potential vehicle such as “tiger”, some of the same elements are activated. The activation summed over the topic and vehicle—what Jung-Beeman refers to as integration—is sufficient to identify the metaphoric ground as a candidate for interpretation. The non-directed, apparently effortless nature of spreading activation and integration mediated by the right hemisphere fits with the inevitable ("automatic") nature of much metaphoric processing: Glucksberg, Gildea, and Bookin (1982), for example, report that non-brain-damaged participants are unable to inhibit their appreciation of metaphoric meaning. The anatomical regions within the right hemisphere most important to coarse-grained activation and integration of semantic information are not yet corroborated by results from several laboratories, though Jung-Beeman (2005) suggests an important role for right temporal structures.

Beeman's framework has been extended to include detail on the selection from among alternatives (Jung-Beeman, 2005; see also Faust & Gernsbacher, 1996; Tompkins, Baumgaertner, Lehman, & Fassbinder, 2000). Metaphor processing requires that the most relevant of the associations held in common between two nouns is identified as the ground of the metaphor. The correct intended meaning (that is, the ground of a metaphor) will often emerge from potential alternatives without directed processing by the listener (e.g., Glucksberg et al., 1982). However, if automatic processing fails to favour an interpretation, understanding will require strategic effort on the part of the listener to evaluate candidate interpretations. The anatomical regions most important for a strategic selection component include left and right prefrontal regions, especially dorsolateral areas. Fletcher, Shallice, Frith, Frackowiak, and Dolan (1996), for example, have reported increased activation in the left and right prefrontal areas during cued recall as demands for response selection increased. More generally, executive function, which subsumes selection among alternatives, is often associated with prefrontal regions (Chiappe & Chiappe, 2007; Crosson et al., 1999).

For assessments of patients' selection ability, we have relied on work by Tompkins and her colleagues that connects metaphor comprehension to working memory, one component of the executive system (e.g., Tompkins, 1990; Tompkins, Bloise, Timko, & Baumgaertner, 1994; Tompkins, Boada, & McGarry, 1992). The Tompkins and colleagues' programme of research suggests that some patients with RHD show apparently normal appreciation of metaphoric alternative meanings when the experimental context is designed to minimise any requirement for strategic planning or conscious effort by participants. However, the same patients with RHD show deficits in apprehending metaphoric alternative meanings when the task requirements are altered to maximise strategic requirements. Additionally, Tompkins et al. (1994) have reported that capacity of working memory correlates with performance in discourse comprehension in patients with RHD (Baddeley, 1986; Baddeley & Hitch, 1994; Caplan & Waters, 1999; Just, Carpenter, & Keller, 1996). Specifically, Tompkins et al. (1994) have adapted the classic reading (listening) working memory test of Daneman and Carpenter (1980) for use with patients who have sustained brain damage. This task requires active processing of information in addition to simple retention: participants judge the truth value of sentences while retaining the final word in each of a set of sentences for later recall.

Generating semantic associations and selecting appropriate shared semantic associations are the core abilities that are targeted in our Metaphor Training Program (Lundgren, Brownell, Roy, & Cayer-Meade, 2006). The foundation of the training is using pictorial displays to make explicit and concrete the semantic ingredients for a potential metaphor so that a patient can view the semantic relationships that he or she may not be consciously aware of, or able to process without aid. Explicit provision of missing or weak links in a task has worked in other domains. Wellman et al. (2002), for example, have used thought bubbles (balloons used to convey thoughts of cartoon characters) in order to help some children with autism perform theory of mind tasks with greater success. The purpose of the present investigation was to determine whether a similar approach would improve success with metaphor for patients with RHD.

Our approach to graphic representation of verbal information in our training program is based on a modification of two out of eight Thinking Maps® originally designed for use with children (Hyerle, 1993, 1995). The single bubble map consists of a main circle (bubble) and lines connecting this main circle to outside circles (bubbles). The format of the double bubble map consists of two main circles and lines connecting these two circles to a number of outside circles. For our purposes, main circles are used to represent noun concepts, which eventually will be the topic and vehicle of a metaphor (see Appendix). Outside circles are used to represent the semantic features or associations to a concept. An outside circle (association) that connects to the two main circles thus represents a possible ground for a metaphor, which may or may not be the dominant association of either noun concept.

We assessed metaphor processing using a comprehension task at the start and end of the training. However, during baseline assessments and the training protocol, we administered an oral interpretation task. We acknowledge that an explanatory task is more challenging and qualitatively different than a purely receptive task for the purpose of assessing what individuals know about a metaphor (Gardner, 1974). The oral interpretation task forces a person to engage strategic processes that are probably not the same as those implicated in most naturally occurring discourse. One motivation for the more explicit oral interpretation task was that, by hypothesis, it required similar operations as the tasks used in the training protocol. In addition, patients' oral interpretation responses provide a rich source of information about fine-grained differences in responses. (See Table 1 and the Discussion section for examples.) A practical motivation in the early stages of this research was tied to measurement: the more difficult interpretation task was less likely to be limited by unreliability or ceiling effects in performance, even with just a few items, than a multiple choice or reaction time assessment task. Patients' success in on-line comprehension of metaphor in natural settings without the added metalinguistic demands of oral interpretation is, of course, the ultimate goal of this research programme.

TABLE 1. Appropriateness of Metaphor Interpretation Scale and examples of scoring

	Father is a steam kettle
Rating description	**Sample interpretations**
6. *Correct.* Complete and appropriate interpretation, characterised by rich, descriptive language (only rarely used)	It appears that this father is under stress, and, given the prolonged stress and the nature of the situation, he will become angry. He may display this anger by yelling at his children for a period of time
5. *Correct.* Complete and appropriate (non-concrete) but basic/simple	Sounds like somebody that has a short temper. He's ready to demonstrate this by yelling real loud
4. *Correct.* Complete and appropriate but (a) delayed (longer than 5 seconds required for response initiation) but eventually gets to the correct response (b) may contain self corrections, or false starts but eventually gets to the correct response, or (c) may include some tangential comments and/or personalisation but eventually gets to the correct response	I can totally relate to that one. I get all mad and then I just scream. I did that last night with the grandkids. I'd say he's ready to blow his mind, I guess
3. *Incorrect.* Close substitution using appropriate alternative metaphor to close substitution/with some clear elements of abstraction but not the complete, correct/appropriate interpretation.	Well the father when he gets angry and he's gonna pop his lid … like a steam-kettle
2. *Incorrect.* Literal associate may be partially correct with some partial or inappropriate non-concrete extension	He's a got a hot head
1. *Incorrect.* Literal associate without any non-concrete extension	Hot
0. *Incorrect.* No response, “I don't know” response, off-topic comments unrelated to the metaphor	I have no idea

METHOD

Development of novel metaphors

The large pool of novel metaphors needed for pre- and post-training metaphor probes and training tasks was constructed using a set of 500 words with at least 10 associations. These words were taken from association data found in texts and websites (Palermo, 1964; Postman & Keppel, 1970; http://www.eat.rl.ac.uk/). A set of 250 metaphors was generated according to the following guidelines. All follow the schema “x is a y”, in which x is the topic and y is the vehicle (Lakoff, 1994). Word pairs were identified for which five associations apply only to the semantic domain of the topic of the metaphor, five associations apply only to the domain of the vehicle, and five relate to the semantic domains of both the topic and vehicle. Frozen metaphors, i.e. metaphors that have become common in the language, were avoided. From this pool we assembled 25 sets of 10 items each to be used for baseline assessments, assessments after each task in the training protocol, and in follow-up testing. The baseline metaphors were never used as part of the training programme; however, some of the metaphors were repeated over sessions to provide an index of repetition effects and untrained learning. There is variation in difficulty among items, but the items in each set of 10 were selected such that overall difficulty within a set was approximately equal based on pilot testing with participants with no brain damage. Out of the 250 metaphors generated, 40 metaphors were reserved for testing purposes only. Out of the 40, 10 were always used for initial assessment, 10 were incorporated into the baseline testing, 10 were used for the immediate post-training assessment, and 10 were used for the 3-month post-training follow-up testing.

Scoring of oral metaphor interpretations

Quality of metaphor interpretation was based on orally produced (later transcribed) interpretations. An oral interpretation task was chosen rather than a multiple choice comprehension task as the major assessment tool in order to demonstrate best the quality and variety of the response and better detect the change in performance over time. Oral interpretations were scored using the Appropriateness of Metaphor Interpretation Scale (AMIS; Lundgren et al., 2007). The AMIS extends from 0 to 6. A 7-point scoring system was used instead of a plus–minus system in order to capture the subtleties of verbal performance (delay, self correction, confabulation) and to provide to provide a fine-grained analysis of the changes in a patient's performance over time. The advantages of this type of scoring system have been recognized in the field of aphasiology (Porch, 2007). Table 1 provides a complete description of the scoring system and examples.

Two judges independently rated each baseline oral metaphor interpretation probe. Discrepancies were resolved via discussion among three judges, all of whom were experienced speech-language pathologists who received extensive training in the use of the metaphor scoring system. Oral metaphor interpretations from different patients and different phases of the protocol were intermixed, which meant that two out of the three judges were blind to the identity of the patient, the training condition (pre- or post-training), and session. This scale has been used reliably (independent inter-rater reliability ranged from 80% to 99% for different patients and pairs of judges).

PARTICIPANTS

Five patients with RHD were recruited through the Harold Goodglass Aphasia Research Center (HGARC) at the Boston VA Healthcare System and Boston University School of Medicine. Patients were excluded on the basis of any of the following: a history of significant alcohol or drug abuse, significant history of prior psychiatric or neurological disorder or learning disability, or significantly impaired (i.e., uncorrected) hearing or vision. All five patients grew up speaking predominantly American English and completed high school. Each patient received a comprehensive speech and language evaluation. Table 2 presents summary information.

TABLE 2. Patient background information and results

Patient	**S-1**	**S-3**	**S-4**	**S-5**	**S-9**
Age	79	60	75	68	68
Years of education	12	12	12	12	16
Time post-onset (years)	25	5	4	2	2
Lesion and aetiology	Right parietal AVM, surgically repaired	Right MCA infarct	Right MCA infarct	Large right hemisphere lesion	Right frontal lesion
Communication/cognitive features	Verbose, tangential, failure to read social cues	Slight residual neglect	Failure to read social cues, concrete thought	Tangential, verbose, inconsistent at reading social cues	Verbose, disinhibited, failure to read social cues
Working Memory Measure (Tompkins et al., 1994) (max = 42), *m* = 12.4 (errors), *SD* = 5.9					
Pre-training (total errors)	10	18	12	15	11
Post-training (total	10	13	11	11	8

errors)					
FANL-C (Kempler & Van Lanker, 1996) (max = 20, RHD*m* = 10.2, *SD* = 4. 4)					
Pre training (Formulaic)	7	10	5	14	15
Post training (Formulaic)	15	8	4	14	16
Total number of sessions	21	31	26	36	19
RESULTS					
Oral Metaphor Interpretation Task					
Baseline mean (*SD*)	23.7 (5.3)	31.0 (3.1)	26.7 (2.7)	34.4 (2.9)	34.5 (3.0)
Post initiation of training mean	36.5	36.8	32.5	40.8	39.7
Effect size: change/*SD*	+2.4	+1.9	+2.2	+2.2	+1.7
3-month follow-up	n/a	30	31	40	43
Simple *r* for training	.802, $p = .011$	.705, $p = 0.002$	.638, $p = .010$	.692, $p = .000$	.702, $p = .030$
Simple *r* for session	.646	.628	.585	.682	.546
The Benton Line Orientation Task short form					
Baseline mean (*SD*)	22.1(2.9)	16.4 (3.9)	1.6 (1.0)	12.9 (0.9)	19.8 (3.0)
Post initiation of training means	22.5	18.7	1.8	13.7	20.9

Effect size: change/*SD*	+0.2	+0.6	+0.2	+0.9	+0.4
3-month follow-up	n/a	19	0	14	20
Simple *r* for training	.091, *p* =.655	.339, *p* = .060	.066, *p* = .655	.431, *p* = .034	.214, *p* = .569
Simple *r* for session	.180	.410	.172	.392	.129

Study design

The structure of the study is based on principles of single-subject experimental design (Kearns, 2000; Olswang, Thompson, Warren, & Minghetti, 1990). Our replicated time series design included pre-training assessment, pre-training baseline probes, training probes, post-training assessment, post-training baseline probes, and 3-month follow up. Within-participant experimental control was obtained by monitoring performance on a control line orientation task throughout each phase of the investigation.

All testing was carried out by licensed speech-language pathologists (SLPs). Each patient was assessed and trained by a single SLP. Two SLPs working on this project learned and practised the protocol under the guidance of the first author, also a SLP, thereby supporting consistency across patients. Weekly meetings regarding training and periodic field observations of the training were initiated in order to be sure that administration of the training protocol remained consistent across clinicians. Patients also performed on an untrained line orientation task to provide a comparison for any changes in oral metaphor interpretation observed over the course of the study. This visuospatial task, while very different in processing demands from oral metaphor interpretation task, is sensitive to right hemisphere injury and, therefore, provides information on non-specific improvement on symptoms associated with RHD.

Finally, two of the five patients performed an extra 10 baseline assessments prior to initiation of training. They were selected on the basis of scheduling availability—whether a patient seemed willing and able to perform over a longer period of time. These patients were not matched with other individual patients. Nonetheless, the extensive amount of additional experience with the oral metaphor interpretation task provided a way to examine the importance of practice effects apart from the training per se.

Pre-training assessment (1 session)

The initial assessment of cognitive and linguistic functioning included, among other tests, The Formulaic and Novel Language Comprehension Test (FANL-C; Kempler & Van Lancker, 1996); The Cognitive and Linguistic Quick Test (CLQT; Helm-Estabrooks,2001); The Benton Line Orientation Task short form (Qualls, Bliwise, & Stringer, 2000); and a working memory span test (Tompkins et al., 1994). These tests were chosen in order to assess the patients' knowledge of figurative language, cognitive-linguistic abilities, visuospatial performance, and working memory.

Baseline phase (10 or 20 sessions, two per week)

The following tests were administered each session: untrained oral metaphor interpretation probes, Benton line orientation, and a modified quality of life scale (SAQOL-39; Hilari, Byng, Lamping, & Smith, 2003) that includes measures of mood and energy. Training was initiated immediately following baseline for three of the patients with RHD. For the other two of the patients, baseline phase was doubled in length for an additional 10 baseline assessments (twice per week for 5 weeks) to provide additional information on the effects of additional time and (untrained) practice with the tasks.

Training phase (twice per week)

In each session a patient performed on untrained metaphors, Benton line orientation, and a quality of life assessment. The training programme comprises five tasks (see Appendix) ordered to move a patient along a performance continuum from easier to more difficult as a way to keep patients engaged in the program. Also, some tasks build directly on other tasks. The easiest tasks require evaluation of connotative meaning of single words (Task I) or word associations (Task II). Tasks III and IV require more effort and call on a number of skills often at risk in brain-damaged patients: generation of word associations (Task III) and evaluation of patient-generated associations linking two nouns (Task IV). Finally, Task V requires selection of a metaphor ground from candidate shared associations.

Post-training assessment (1 session)

The purpose of the post-training assessment was to gauge generalisation and maintenance of gains achieved during training, and also to provide a preliminary basis for interpreting any gains. Novel oral metaphor interpretation, line orientation, and quality of life were evaluated, as well as (among others) working memory span (Tompkins et al., 1994) and non-literal language comprehension (FANL-C; Kempler & Van Lancker, 1996).

Long-term follow-up (1 session)

Three to four months after the end of training, patients were seen once more for a brief session that included untrained oral metaphor interpretation and Benton line orientation to provide an indication of the permanence of gains.

RESULTS

We tested whether initiation of training coincided with a detectable change in patients' oral metaphor interpretation that could be distinguished from any gradual change over time or repeated practice with the task. Also, as a comparison, we tested for corresponding change in a patient's visuospatial performance as measured by Benton line orientation scores. The conventional use of visual-graphic data analysis for single-participant studies provides a useful interpretation aid. However, traditional presentation of graphic data has been shown to be unreliable and subject to interpreter bias (e.g., Fisher, Kelly, & Lomas, 2003; Robey, Schultz, Crawford, & Sinner, 1999). Our data analysis approach, although novel to aphasiology research, objectively manages the most serious form of this bias inherent in visual-graphic data analysis, serial dependence, and autocorrelation.

Statistical tests of hypotheses were based on a recent application of bootstrapping/simulation developed by Borckardt et al. (2008, http://clinicalresearcher.org). In brief, one starts with a null hypothesis of no overall change in performance from pre- to post-training due to training. The program calculates a Pearson's *r* value predicting the dependent measure using pre- versus post-initiation of training (that is, using an X variable consisting of 0s and 1s). The program also computes the autocorrelation (lag 1) for the entire set of observations. Assuming the null hypothesis of no effect of training, the programme draws (from a normal population) a very large number (e.g., 10,000) of random samples of pre- and post-treatment data with the identified level of autocorrelation (that is, gradual improvement over sessions and non-independence) calculated from the available data. For each sample, the program calculates a Pearson's *r* value predicting pre- versus post-initiation scores. The resulting *r* values provide a distribution of what to expect if the null hypothesis were true. It is then straightforward to evaluate the Pearson *r* value calculated from a patient's data to see if the *r* is a stronger than 95% of the null hypothesis values. One strength of this approach is that it avoids the difficult statistical issue of how to carry out inferential analysis on a single participant's data, which typically violate the assumption of independence, and it also has greater statistical power for small sample sizes than other procedures (e.g., ITSACOR) (Borckardt et al., 2008; Crosbie, 1993).

Table 2 lists results for oral metaphor interpretation and line orientation for all five patients: baseline mean and standard deviation, post initiation of training mean, improvement effect size, and the Pearson's *r* value along with its statistical significance. Effect size was defined as the difference in average performance divided by the standard deviation of the baseline assessments. The size of the *r* value includes both the effect of training and also any gradual change over sessions that would be anticipated without any training. A higher positive value of *r* reflects greater improvement from baseline to post-baseline assessments. However, the probability value listed next to each *r* value reflects the statistical significance of the training effect; that is, how likely a result that large or larger would happen by change under the null hypothesis of no training effect over and above other change over sessions. Table 2 also provides 3-month follow up performances (with one exception), and pre- and post-training scores on the working memory span test and the FANL-C test for non-literal language comprehension.

All five patients with RHD tolerated the training and showed significant gains in oral metaphor interpretation scores. The Pearson*r* values for metaphor interpretation ranged from +.6 to +.8, and their effect sizes ranged from +1.7 to +2.4. Three patients maintained their gains at the 3-month follow-up testing, and only one patient (S-3) returned to baseline levels. (One patient was not available for the 3-month follow-up testing due to health issues.) By comparison, there was less improvement on the untrained Benton line orientation task. The *r* values for the line orientation task were smaller, +.1 to +.4, and the effect sizes were also smaller, +0.2 to +0.9.

S1 obtained the largest *r* value for metaphor interpretation ($r = +.80$, $p = .011$) and the largest effect size (+2.4). Moreover, S1 was the only patient to show marked improvement on the FANL-C test of non-literal language comprehension: initial score = 7 (below average for stroke patients), post-training score = 15 (above average). S1 was the most severely impaired of the five patients on the FANL-C at the initial assessment, but scored well on the Tompkins et al. test of working memory (initial score = 10 errors, which was slightly better than average for stroke patients). S1's *r* value (+.09) and effect size (+.2) for the Benton line orientation task suggest that the improvement observed on the oral metaphor interpretation task was selective.

There was no consistent connection between improvement in working memory performance and oral metaphor interpretation: For example, S1 who showed the greatest improvement in metaphor performances showed no change in working memory score.

Patient S-4 is noteworthy for showing a reliable effect of training on metaphor interpretation despite very low performance on the Benton line orientation task.

A final observation is that the two patients (S3, S5) who performed on 20 rather than just 10 baseline assessments were the only patients to show significant or nearly significant effects on the Benton line orientation measure. The *r* values were +.3 and +.4, and the effect sizes were +.6 and +.9 for S3 and S5, respectively. These were the largest changes observed in the set of five patients for the Benton line orientation task. These same two patients both obtained clearly greater improvement on metaphor interpretation. The *r* values were +.7 for both S3 and S5, and the effect sizes were 1.9 and 2.2 for S3 and S5, respectively.

DISCUSSION

This structured training protocol was designed to remediate one type of cognitive-linguistic deficit in patients with RHD, processing a common form of non-literal language. One foundation of the protocol was graphically supported practice with the types of word associations that play a major role in many instances of metaphor according to our guiding theoretical analysis. The other foundation was selecting a good completion for a metaphor from the available associations. Patients reliably completed the tasks and were willing to engage in both the assessment and training components of the programme.

Overall the results of this study suggest that individuals with RHD, even years post stroke, are capable of completing this intervention and can show varying amounts of improvements in oral metaphor interpretation. For example, prior to training, when asked to interpret the metaphor “the game is a nightmare”, S-5 responded with the following:

> The game is a nightmare. Well the Red Sox are losing 10 to 2 in the last of the eighth. It's a monster. It's gonna swallow you up. You're not gonna win this one. (4 second delay) Now if you write all that down, make something up it'll make it shorter. ...You're making it harder on yourself.

The response was given an AMIS score of 2. This score indicates that the interpretation was incorrect. As outlined in Table 1, a literal associate may be partially correct with some partial or inappropriate non-concrete extension.

Following training when given the metaphor “The project is a monster”, S-5 responded with “It's too big. Everybody's afraid to tackle it.” This response was given a score of 5, which indicates a response that is correct, complete, and appropriate. S-5 was able to find the shared associates for “project” and “monster” (large, not easy to handle, scary) and use these commonalities to apprehend the meaning of the metaphor.

All of the patients with RHD demonstrated success with this programme, despite the difficulty one patient (S-4) had with visuospatial processing. Patient S-4 exhibited a reliable effect of training on metaphor interpretation despite very low performance on the Benton task. This discrepancy between metaphor interpretation and visuospatial skill raises the important questions of what specific perceptual skills are needed for success on the training tasks and how those skills can be assessed.

Our approach to evaluating the efficacy of the training programme yields necessary though not sufficient evidence for success. We have shown a statistically significant improvement linked to initiation of training for five patients. That improvement was selective for three patients: the oral metaphor interpretation improved but not line judgements. In addition, the strong Pearson *r*values for each patient indicate effect sizes that express the degree of performance change pre versus post initiation of training that takes into account variability from session to session. Furthermore, the 3-month follow-up testing revealed that three out of four patients for whom data were available maintained their gains overall time. While longer-term follow-up results would strengthen claims about the value of training, the initial results make a compelling case for further investigation to explore which patient characteristics predict good response to training. This conclusion is further supported by the fact that our results included generalisation to new metaphors.

There are several unanswered questions growing out of our preliminary results. Because we used a very structured protocol, we have not determined the optimal number of sessions or duration of session to maximise performance, nor have we identified the optimal type or amount of self-study a patient should carry out at home after the session with the SLP. These are empirical issues to be investigated.

While our choice of an interpretation task for assessment makes sense for an initial effort, we found that only one patient, S-1 who showed the largest effect size, also showed marked improvement on the FANL-C, which uses a sentence–picture matching, multiple choice format. Clearly, not all assessments will be as sensitive as the oral interpretation task; nor will all assessments measure exactly the same ability. We have not assessed whether gains associated with training can be detected using more natural metaphor comprehension tasks than oral interpretation. We anticipate that sensitive reaction time paradigms will be needed to detect subtle improvements. We also look forward to testing generalisation of training gains to other aspects of non-literal language processing and to discourse performance more generally. Most importantly, we have yet to assess whether the gains can be shown to have noticeable impact for patients and their communication partners interacting in natural settings.

Some intriguing possibilities stem from the impact of the extended baseline used with two patients. The additional assessments prior to training were intended to provide preliminary information on patients' ability to improve on either the Benton line orientation task or the oral metaphor interpretation task with simple practice. The extra baseline sessions for two patients do not support direct comparison between patients given 10 as opposed to 20 baseline sessions because patients were not paired on the basis of severity of impairment or any other characteristics. Nonetheless, 20 sessions prior to training represents a great deal of practice and, for that reason, offers a useful look at practice effects. The two patients who performed on the extended baseline sessions exhibited some marginal improvement in the untreated Benton line orientation task that coincided with initiation of metaphor training. One speculation is that the observed improvement in these two patients could be in part due to novelty effects: after an extra long period of baseline assessment, just starting training could provide a strong boost in attitude and effort in tasks that leads to improvement across domains. While the present study was not designed to examine novelty effects, this issue is central to the efficacy of many training programmes and deserves prospective investigation.

CONCLUSIONS

The findings of this study demonstrate that structured semantic training to support oral metaphor interpretation had significant effects on performance for five patients with varied lesions in the right hemisphere due to stroke. Importantly, for three out of five patients the effects of training appeared to be distinct from non-specific effects of practice, particularly when oral metaphor interpretation performances were contrasted against those for the Benton line orientation task. The patient who showed the most improvement both on oral metaphor interpretation and also on a more traditional metaphor comprehension test scored relatively poorly on initial metaphor assessments but relatively well on a working memory test. There are several questions that remain open at this time: precisely which visuospatial skills must be intact for a patient to benefit from the graphic displays used in this programme, the durability of gains, the precise scope of generalisation of improvement documented on the oral interpretation task, the "real-world" impact of observed gains, and, finally, the linguistic and cognitive symptom profile of patients who will benefit the most from this protocol. This study succeeds in demonstrating the usefulness of this type of programme to remediating metaphor processing in individuals with RHD. It provides a flexible model for interventions that target this or other communication deficits exhibited by patients with RHD due to stroke or other aetiologies.

Acknowledgments

This work was supported by NIH: NIDCD grants R21 DC007165, P30 DC05432, and RO1 DC0520702 as well as the Research Service of the Boston VA Healthcare System. We wish to acknowledge Douglas Katz MD and T. Joy DePiero MD at Braintree Rehabilitation Hospital, Carol Abbott MS, CCC-SLP at New England Rehabilitation Hospital, and William Milberg PhD at the VA Boston Healthcare System for referring these patients, and Marissa Fond, Soma Roy, Michelle Nichols, Jacque Spitzer, Kate Caddick, and Andy Stringfellow for their contributions to this project. We thank the anonymous reviewers for their many useful comments that improved the paper substantially, and finally the participants and their families for their time and willingness to participate.

APPENDIX

Metaphor training programme: Step-by-step description of the training and scoring procedures, rationale, and criteria for advancement for each task

Assessments of oral metaphor interpretation, line orientation, and mood and energy (SAQOL-39) will be administered twice weekly during Baseline, Training, and Post-Training phases, administered again at the end of the training protocol, and finally after a 3-month delay.

Baseline phase (5 weeks)

To assess the patient's ability to interpret metaphor, he or she is given 10 novel metaphors and asked to provide an oral interpretation for each. Oral interpretations are recorded for later transcription.

Task I: Judgements of single word connotative meaning

This task aims to illustrate connotative meaning and to provide practice in thinking about connotative meaning. Stimulus words are presented one at a time in a bubble (with no links) to familiarise the patient with the computer display. The patient answers yes/no questions about the value of 10 single words (presented one at a time) along several connotative scales: Think of the word “father” (boy, window, butterfly, doctor, carpet, fruit, lamp, needle, girl, star), is this word typically considered “beautiful”, is this word typically considered “strong”?

Scoring procedure: Items are not scored as correct or incorrect at this stage except in terms of time to respond. An item is “correct” if the patient responds in less than five seconds. This task allows the patient to begin to think about words in terms of a variety of dimensions. Criterion for Completion of Task I: Completion of 5 sets of 10 words each.

Task II: Judgements of word associations

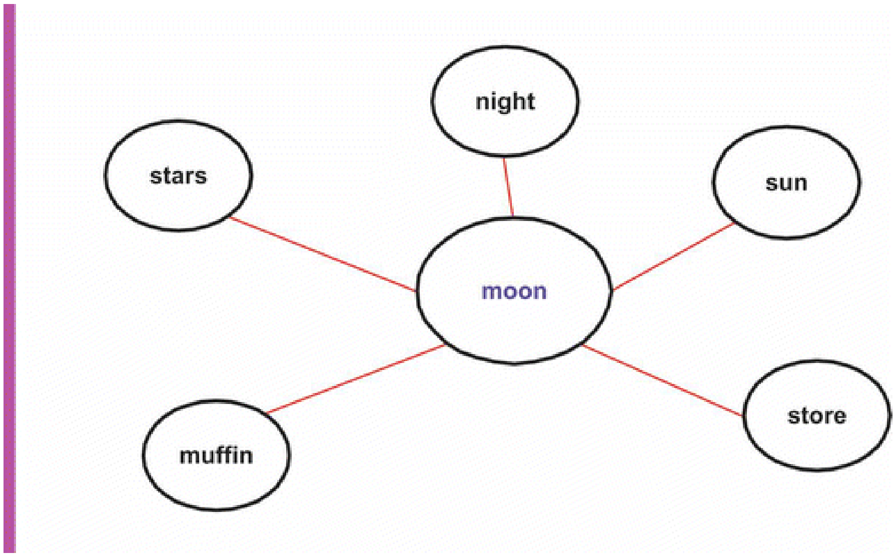
This task aims to illustrate typical concrete associations to target stimuli and to provide practice assessing associations. A single bubble map is used for this task. The examiner points to the target word (enclosed within the centre bubble), and to the five lines attached to this centre bubble which link the target word to the associated words enclosed in five outside bubbles.

For example, the patient is given a target word (e.g., "moon") and then answers yes/no questions about five single words (presented one at a time) associated with the target word. Is (sun, muffin, store, stars, night) typically associated with this word?

Scoring procedure: In order to receive full credit (1 point) on an item, a patient must respond promptly and accurately. A half point is taken off if the patient takes longer than 5 seconds to initiate a response. Criterion for completion of Task II: 90% "correct" × 3 sets of 10 target words each, or completion of 5 sets of 10 target words.

Task III: Generation of word associations

This task aims to provide practice in generating at least five word associations related to a target word, which appears in a single bubble map. The examiner begins a trial by placing a single target word in a bubble that has five links drawn from it to five empty bubbles. The examiner points to an empty bubble and asks the patient to provide a *typical* associate. If the patient is unable to generate associations, or if he or she generates personalised associations, the examiner will provide cues and will redirect as needed by returning to the set of 10 questions listed under Task I.



Scoring procedure: A patient achieves one point for each correct typical association. (Personalised responses do not count as correct.) A half point is deducted for a significantly delayed response. Criterion for completion of Task III: 90% “correct” × 3 sets of 10 target words each or completion of 5 sets of 10 target words.

Task IV: Judgement of patient-generated associations to link two words

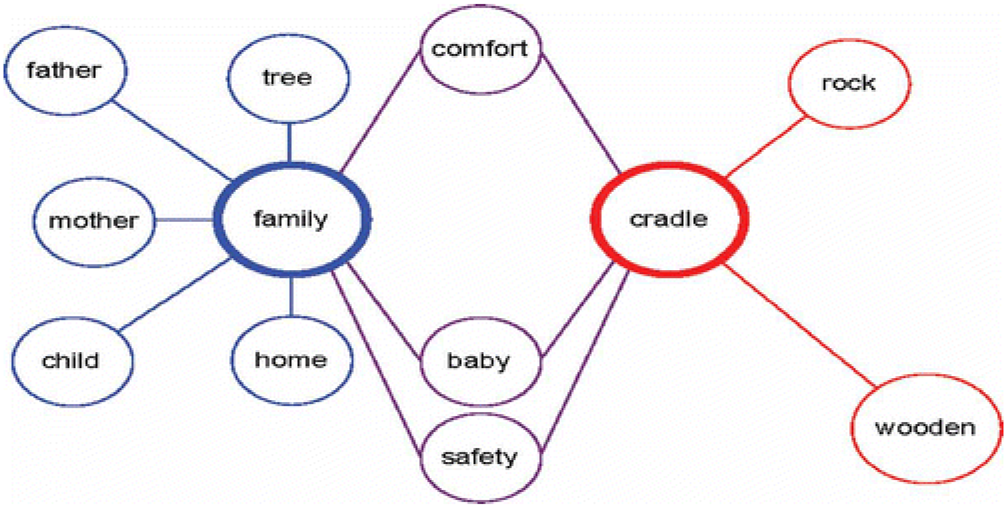
This task aims to provide additional practice in generating associations and, more important, to provide practice in evaluating appropriateness of associations to (i.e., between) two words. The examiner starts a trial by providing a target word (Word 1) inside a single bubble of the double bubble map, with five links drawn into five empty bubbles. The patient accurately generates five associations related to the target word (Word 1). The examiner cues or re-directs if any of the associations are inappropriate for any reason. The examiner then shows the patient a new word (Word 2) in the adjacent bubble. The patient answers yes/no for each association whether that association generated for Word 1 can also be associated with Word 2.

Scoring procedure: A patient will receive one point if able to generate two overlapping associations and a half point for one overlapping association. Criterion for Completion of Task IV: 90% “correct” × 3 sets of 10 trials each or completion of 5 sets of 10 trials each.

Task V: Selection of appropriate metaphor ground from candidate dual associations

This task aims to provide practice selecting the basis for a metaphor from a set of candidates. The patient views a metaphor placed within a double bubble map (topic in bubble one and vehicle in bubble two). The patient selects the appropriate interpretation of this metaphor appearing in the double bubble map from the field of three written interpretations (correct, literal, close substitution using another metaphor).

Scoring procedure: A patient's score is the number of correct responses, one point for prompt, correct response and a half point if the response is correct but delayed (longer than 5 seconds). Criterion for Completion of Task V: 90% correct × 3 sets of 10 trials each (a patient's score will be the number correct out of 10) or completion of 5 sets of 10 trials each.



Cueing provided

Each task begins with instructions and an example. If the patient is having difficulty with a response the instructions and the example are repeated. If the patient continues to have difficulty with an item in a task, an example of a correct response for that particular item is given and discussed. All clinicians used this particular cueing hierarchy during training.

REFERENCES

1. Aitchison, J. 1987. *Words in the mind*, Oxford, , UK: Basil Blackwell.

2. Baddeley, A. D. 1986. *Working memory*, Oxford, , UK: Oxford University Press.

3. Baddeley, A. D. and Hitch, G. J. 1994. Developments in the concept of working memory. *Neuropsychology*, 8: 485–493.

4. Baum, S. R. and Dwivedi, V. D. 2003. Sensitivity to prosodic structure in left- and right-hemisphere-damaged individuals.*Brain and Language*, 87: 278–289.

5. Beeman, M. 1998. "Coarse semantic coding and discourse comprehension". In *Right hemisphere language comprehension. Perspectives from cognitive neuroscience*, Edited by: Beeman, M. and Chiarello, C. 255–284. Mahwah, NJ: Lawrence Erlbaum Associates Inc.

6. Borckardt, J. J., Nash, M. R., Murphy, M. D., Moore, M., Shaw, D. and O'Neil, P. 2008. Clinical practice as natural laboratory for psychotherapy research. A guide to case-based time-series analysis. *American Psychologist*, 63(2): 77–93.

7. Brown, R. 1958. *Words and things*, Glencoe, IL: Free Press.

8. Brownell, H. and Stringfellow, A. 1999. Making requests: Illustrations of how right hemisphere brain damage can affect discourse production. *Brain and Language*, 68: 442–465.

9. Brownell, H. H., Potter, H. H., Michelow, D. and Gardner, H. 1984. Sensitivity to lexical denotation and connotation in brain damaged patients: A double dissociation?. *Brain and Language*, 22: 253–262.

10. Brownell, H. H., Simpson, T. L., Bihrle, A. M., Potter, H. H. and Gardner, H. 1990. Appreciation of metaphoric alternative word meaning by left and right brain-damaged patients. *Neuropsychologia*, 28: 375–383.

11. Burgess, C. and Simpson, G. B. 1988. Cerebral hemispheric mechanisms in the retrieval of ambiguous word meanings. *Brain and Language*, 33: 86–103.

12. Caplan, D. and Waters, G. S. 1999. Working memory and sentence comprehension. *Behavioral and Brain Sciences*, 22: 77–126.

13. Chiappe, D. L. and Chiappe, P. 2007. The role of working memory in metaphor production and comprehension. *Journal of Memory and Language*, 56: 172–188.

14. Chiarello, C. 1998. "On codes of meaning and the meaning of codes: Semantic access and retrieval with and between hemispheres". In *Right hemisphere language comprehension. Perspectives from cognitive neuroscience*, Edited by: Beeman, M. and Chiarello, C. 141–160. Mahwah, NJ: Lawrence Erlbaum Associates Inc.

15. Crosbie, J. 1993. Interrupted time-series analysis with brief single-subject data. *Journal of Consulting and Clinical Psychology*, 61: 966–974.

16. Crosson, B. R., Rao, S. M., Woodley, S. J., Rosen, A. C., Bobholz, J. A.Mayer, A. 1999. Mapping of semantic, phonological, and orthographic verbal working memory in normal adults with functional magnetic resonance imaging. *Neuropsychology*, 13: 171–187.

17. Daneman, M. and Carpenter, P. A. 1980. Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, 19: 450–466.

18. Faust, M. E. and Gernsbacher, E. 1996. Cerebral mechanisms for suppression of inappropriate information during sentence comprehension. *Brain and Language*, 53: 234–259.

19. Fisher, W. W., Kelley, M. E. and Lomas, J. E. 2003. Visual aids and structural criteria for improving visual inspection and interpretation of single-case designs. *Journal of Applied Behavioral Analysis*, 36: 387–406.

20. Fletcher, P. C., Shallice, T., Frith, C. D., Frackowiak, R. S. J. and Dolan, R. J. 1996. Brain activity during memory retrieval: The influence of imagery and semantic cueing. *Brain*, 119: 1587–1596.

21. Gardner, H. 1974. Metaphors and modalities: How children project polar adjectives onto diverse domains. *Child Development*, 45: 84–91.

22. Gardner, H. and Denes, G. 1973. Connotative judgements by aphasic patients on a pictorial adaptation of the semantic differential. *Cortex*, 9: 183–196.

23. Glucksberg, S., Gildea, P. and Bookin, H. A. 1982. On understanding nonliteral speech: Can people ignore metaphors?.*Journal of Verbal Learning and Verbal Behavior*, 21: 85–98.

24. Happe, F., Brownell, H. and Winner, E. 1999. Acquired ‘theory of mind’ impairments following stroke. *Cognition*, 70: 211–240.

25. Hilari, K., Byng, S., Lamping, D. L. and Smith, S. C. 2003. Stroke and Aphasia Quality of Life Scale-39 (SAQOL-39): Evaluation of acceptability, reliability, and validity. *Stroke*, 34: 1944–1950.

26. Hillson, T. R. and Martin, R. A. 1994. What's so funny about that?: The domains-interaction approach as a model of incongruity and resolution in humor. *Motivation and Emotion*, 18: 1–29.

27. Helm-Estabrooks, N. 2001. *CLQT Cognitive Linguistic Quick Test*, New York: The Psychological Corporation.

28. Hyerle, D. 1993. *Thinking Maps for Multiple Modes of Understanding*, Unpublished doctoral dissertation, University of California Berkeley.

29. Hyerle, D. 1995. *Thinking Maps®: Tools for learning*, Cary, NC: Innovative Learning Group, Innovative Sciences, Inc.

30. Joanette, Y., Goulet, P. and Hannequin, D. 1990. *Right hemisphere and verbal communication*, New York: Springer-Verlag.

31. Jung-Beeman, M. 2005. Bilateral brain processes for comprehending natural language. *Trends in Cognitive Sciences*, 9(11): 512–518.

32. Just, M. A., Carpenter, P. A. and Keller, T. A. 1996. The capacity theory of comprehension: New frontiers of evidence and arguments. *Psychological Review*, 103: 773–780.

33. Kearns, K. P. 2000. “Single-subject experimental designs in aphasia”. In *Aphasia and language*, Edited by: Nadeau, S. E., Gonzalez Rothi, L. J. and Crosson, B. 421–441. New York: Guilford Press.

34. Kempler, D. 2005. *Neurocognitive disorders in aging*, Thousand Oaks, CA: Sage.

35. Kempler, D. and Van Lancker, D. R. 1996. *The Formulaic and Novel Language Comprehension Test* unpublished manuscript

36. Lakoff, G. (1994). Conceptual metaphor home page of the University of California at Berkeley http://cogsci.berkeley.edu/metaphors/ (http://cogsci.berkeley.edu/metaphors/)

37. Lundgren, K., Brownell, H., Cayer-Meade, C., Spitzer, J., Nichols, M. and Caddick, K. 2007. *The Appropriateness of Metaphor Interpretation Scale (AMIS)*, Boston, MA: Poster presentation at the annual convention of the American Speech-Language Hearing Association.

38. Lundgren, K., Brownell, H., Roy, S. and Cayer-Meade, C. 2006. A metaphor comprehension intervention for patients with right hemisphere brain damage: A pilot study. *Brain and Language*, 99: 69–70.

39. Lundgren, K., Moya, K.L. and Bennowitz, L.I. Perception of nonverbal cues after right brain damage [abstract]. Clinical Aphasiology: Conference Proceedings. Minneapolis. Edited by: Brookshire, R. H. Vol. 282, BRK Publishers.

40. Lyons, J. 1968. *Introduction to theoretical linguistics*, Cambridge, , UK: Cambridge University Press.

41. McDonald, S. 1993. Viewing the brain sideways? Frontal versus right hemisphere explanations of non-aphasic language disorders. *Aphasiology*, 7: 535–549.

42. McDonald, S. 1999. Exploring the process of inference generation in sarcasm: A review of normal and clinical studies. *Brain and Language*, 68: 486–506.

43. Myers, P. S. 1999. *Right hemisphere damage: Disorders of communication and cognition*, San Diego, CA: Singular Publishing Group.

44. Myers, P. S. and Linebaugh, C. W. 1981. "Comprehension of idiomatic expressions by right-hemisphere-damaged adults". In*Clinical Aphasiology: Clinical proceedings*, Edited by: Brookshire, R. H. 254–261. Minneapolis: BRK Publishers.

45. Olswang, L. B., Thompson, C. K., Warren, S. F. and Minghetti, N. J., eds. 1990. *Treatment efficacy research in communication disorders*, Rockville, MD: American Speech-Language-Hearing Foundation.

46. Osgood, C. E., Suci, G. J. and Tannenbaum, P. H. 1957. *The measurement of meaning*, Urbana, IL: University of Illinois Press.

47. Palermo, D. 1964. *Word association norms*, Minneapolis: University of Minnesota Press.

48. Porch, B. 2007. Comments on "Comparison of alternatives to multidimensional scoring in the assessment of language comprehension in aphasia" by Odekar and Hallowell (2005). *American Journal of Speech-Language Pathology*, 16: 84–86.

49. Postman, L. and Keppel, G. 1970. *Norms of word associations*, New York: Academic Press.

50. Qualls, C. E., Bliwise, N. G. and Stringer, A. Y. 2000. Short forms of Benton Judgement of Line Orientation Test: Development and psychometric properties. *Archives of Clinical Neuropsychology*, 15: 159–163.

51. Robey, R. R., Schultz, M. C., Crawford, A. B. and Sinner, C. A. 1999. Single-subject clinical-outcome research: Designs, data, effect sizes, and analyses. *Aphasiology*, 13: 445–473.

52. Shammi, P. and Stuss, D. T. 1999. Humor appreciation: A role of the right frontal lobe. *Brain*, 122: 657–666.

53. Shibata, M., Abe, J., Terao, A. and Miyamoto, T. 2007. Neural mechanisms involved in the comprehension of metaphoric and literal sentences: An fMRI study. *Brain Research*, 1166: 92–102.

54. Smith, E. E. and Jonides, J. 1999. Storage and executive processes in the frontal lobes. *Science*, 283: 1657–1661.

55. Stemmer, B., Giroux, F. and Joanette, Y. 1994. Production and evaluation of requests by right hemisphere brain-damaged individuals. *Brain and Language*, 47: 1–31.

56. Stringaris, A. K., Medford, N., Giampietro, V. C., Brammer, M. J. and David, A. 2007. Deriving meaning: Distinct neural mechanisms for metaphoric, literal, and non-meaningful sentences. *Brain and Language*, 100: 150–162.

57. Tompkins, C. A. 1990. Knowledge and strategies for processing lexical metaphor after right or left hemisphere brain damage.*Journal of Speech and Hearing Research*, 33: 307–316.

58. Tompkins, C. A. 1995. *Right hemisphere communication disorders: Theory and management*, San Diego, CA: Singular Publishing Group.

59. Tompkins, C. A., Baumgaertner, A., Lehman, M. T. and Fassbinder, W. 2000. Mechanisms of discourse comprehension impairment after right hemisphere brain damage: Suppression in lexical ambiguity resolution. *Journal of Speech Language & Hearing Research*, 43(1): 62–78.

60. Tompkins, C. A., Bloise, C. G. R., Timko, M. L. and Baumgaertner, A. 1994. Working memory and inference revision in brain-damaged and normally aging adults. *Journal of Speech and Hearing Research*, 37: 896–912.

61. Tompkins, C. A., Boada, R and McGarry, K. 1992. The access and processing of familiar idioms by brain-damaged and normally aging adults. *Journal of Speech and Hearing Research*, 35: 626–637.

62. Tourangeau, R. and Sternberg, R. J. 1981. Aptness in metaphor. *Cognitive Psychology*, 13: 27–55.

63. Van Lancker, D. R. 1990. The neurology of proverbs. *Behavioral Neurology*, 3: 169–187.

64. Van Lancker, D. R. and Kempler, D. 1987. Comprehension of familiar phrases by left but not by right hemisphere damaged patients. *Brain and Language*, 32: 265–277.

65. Wellman, H. M., Baron-Cohen, S., Caswell, R., Gomez, J. C., Swettenham, J.Toyer, E. 2002. Thought-bubbles help children with autism acquire an alternative to a theory of mind. *Autism*, 6: 343–363.

66. Weylman, S. T., Brownell, H. H., Roman, M. and Gardner, H. 1989. Appreciation of indirect requests by left and right brain-damaged patients: The effects of verbal context and conventionality of wording. *Brain and Language*, 36: 580–591.

67. Winner, E., Brownell, H., Happe, F., Blum, A. and Pincus, D. 1998. Distinguishing lies from jokes: Theory of mind deficits in right hemisphere brain-damaged patients. *Brain and Language*, 62: 89–106.

68. Winner, E. and Gardner, H. 1977. The comprehension of metaphor in brain-damaged patients. *Brain*, 100: 717–729.