Top Lang Disorders
Vol. 36, No. 1, pp. 80–96

An Integrated Approach for Treating Discourse in Aphasia

Bridging the Gap Between Language Impairment and Functional Communication

Lisa Milman

Purpose: A primary goal of aphasia intervention is to improve everyday communication. Although a large body of research focuses on treatment generalization, transfer of learning to real-world interactions involving discourse does not always occur. The goal of an integrated discourse treatment for aphasia (IDTA) approach is to facilitate such generalization. This article reviews generalization data from a series of four closely related IDTA studies. **Method:** Treatment in all studies (two case reports and two single subject experimental designs) followed a problem-based learning model targeting word retrieval, sentence processing, and discourse production. **Results:** Seven of eight participants acquired the target vocabulary and sentence structures. In addition, generalization was evident for most participants on related linguistic structures, standardized tests of language and cognition, and/or measures of spoken discourse production. **Conclusions:** Findings add to previous research supporting an IDTA approach to improving discourse in persons with aphasia. Participant characteristics and properties of outcome measures associated with these results are also discussed. **Key words:** *aphasia*, *discourse*, *disorder*, *generalization*, *integrated*, *language*, *lexical*, *neurogenic*, *syntax*, *treatment*

INTEGRATED DISCOURSE TREATMENT FOR APHASIA (IDTA) is an approach to intervention that focuses on promoting transfer of impairment-based language learning to everyday communicative use. Although a large body of aphasia research is dedicated to understanding treatment generalization (Raymer et al., 2008; Thompson, Shapiro, Kiran, & Sobecks, 2003; Wambaugh, Mauszycki, Cameron, Wright, & Nessler, 2013), the theoretical foundation for much of this work has focused on *horizontal* generalization within a single level of language processing. For example, several treatments are based on the premise that training a set of words sharing a semantic relation (e.g., *apple, peach, kiwi*) or a phonological relation (e.g., *garden, garbage, garage*) will result in generalization of treatment effects from trained to untrained words (Boyle, 2004; Greenwood, Grassly, Hickin, & Best, 2010; Kiran & Thompson, 2003). Similarly, many morphosyntactic sentence-level treatments (Thompson et al., 2003) are based on the premise that training will

Author Affiliation: *Department of Communicative Disorders & Deaf Education, Utah State University, Logan.*

This paper summarizes and discusses the findings of four related studies, which were conducted/presented/published by the author. A portion of this work was supported by a private endowment to advance aphasia treatment research at Ohio State University. The author was an employee at Ohio State University and at Utah State University when this research was conducted. There are no other financial disclosures associated with publication of this work.

The author has indicated that she has no additional financial or nonfinancial relationships to disclose.

Corresponding Author: *Lisa Milman, PhD, Department of Communicative Disorders & Deaf Education, Utah State University, 1000 Old Main Hill, Logan, UT 84322 (lisa.milman@usu.edu).*

DOI: 10.1097/TLD.0000000000000076

activate broad morphosyntactic processes resulting in generalized effects that extend from trained to untrained sentence structures. A further assumption of many of these treatments is that the theoretically motivated horizontal generalization will automatically extend vertically across levels of processing and transfer to everyday function.

Whereas research has shown relatively robust horizontal (word-to-word or sentence-to-sentence) generalization (Kiran & Thompson, 2003; Thompson et al., 2003), transfer of learning across levels has been less consistent (see discussions in Boyle, 2011; Savage, Donovan, & Hoffman, 2014; Wilkinson & Wielaert, 2012). These findings have prompted researchers to investigate alternate approaches for promoting change in daily communication (Boyle, 2011; Greenwood et al., 2010; Holland & Hinckley, 2002; Life-Participation Approach to Aphasia [LPAA] Project Group, 2008; Webster, Whitworth, & Morris, 2015; Whitworth et al., 2015).

This article examines an IDTA approach for promoting such generalization. The underlying hypothesis of the work presented here is that generalization of clinical training to real-world use can be facilitated by application of principles borrowed from a problem-based learning model that is often used in higher education.

PROBLEM-BASED LEARNING AND GENERALIZATION TO REAL-WORLD CONTEXTS

Problem-based learning promotes generalization of training to real-world contexts through the application of three basic principles: (1) centering learning tasks on everyday real-world problems/challenges; (2) maximizing personal relevance; and (3) systematically training complex component skills and integration of these skills into whole task performance (Hung, 2013; Schmidt, 1993). For example, with respect to communication, a real-world problem/challenge might be talking with a spouse about family and friends. Personal relevance could be enhanced by including names of family and friends in treatment. Training sessions would focus not only on producing these names in isolation (as in a picture-naming task) but also producing them in the context of authentic conversations. For many speech-language pathologists, these will be familiar principles that are routinely used in their clinical practice. In spite of the intuitive common sense appeal of these principles, the majority of treatment research has focused on impairment-based training in which component skills, such as naming, are trained in isolation from authentic communicative contexts.

PROBLEM-BASED LEARNING AND APHASIA REHABILITATION

Both the LPAA and the IDTA approach share principles with problem-based learning. The LPAA shifts the emphasis of intervention from language impairment to functional communication (LPAA Project Group, 2008). Examples of LPAA interventions include use of alternative and augmentative communication systems (Beukelman, Fager, Ball, & Dietz, 2007), script training (Holland, Halper, & Cherney, 2010), group therapy (Elman & Bernstein-Ellis, 1999), and environmental manipulations such as communication partner training (Boles, 2011; Kagan, 1998). All of these treatments directly target complex real-life problems (reducing communication barriers and/or social isolation). In addition, personal relevance is maximized by involving the client in selection of treatment goals, procedures, and materials (see discussions in Kimbarow, 2007; LPAA Project Group, 2008; Sorin-Peters, 2003).

Whereas the LPAA shifts the focus from language impairment to functional communication, an IDTA approach *bridges* impairment-based and communication-focused interventions by training linguistic structures in isolation *and* in functional discourse contexts. Applications of integrated discourse treatments in aphasia have focused on training word retrieval and discourse (Antonucci, 2009; Boyle, 2004; Cameron, Wambaugh, Wright, & Nessler, 2006; Falconer & Antonucci, 2012; Greenwood et al., 2010; Peach & Reuter, 2010); sentence processing and discourse (Murray, Timberlake, & Eberle, 2007); or word retrieval and sentence processing and discourse (Gaddie, Kearns, & Yedor, 1991; Goral & Kempler, 2009; Peach & Wong, 2004; Whitworth et al., 2015). Notably, in addition to showing gains in the trained structures (words and/or sentences), the 11 studies cited previously have also reported general improvements in discourse production in the majority of participants.

CURRENT RESEARCH

This article adds to this literature by reviewing a series of four related phase II treatment studies (Milman et al., 2014a; Milman & Bruner, 2010; Milman, Clendenen, & Vega-Mendoza, 2014b; Milman, Vega-Mendoza, & Clendenen, 2014c) that apply an IDTA approach. In keeping with the problem-based learning approach, treatment in all four studies incorporated the following three characteristics: the therapy focused on solving a real-world problem (e.g., conversing about a particular topic); participants took an active role in treatment planning; and isolated task components (e.g., word retrieval and sentence production) were trained as well as whole task performance (i.e., discourse). Consistent with the definition of phase II treatment studies (Robey, 2004), the broad purpose of this research was to identify the target population for this treatment, refine treatment procedures, and estimate treatment effect size. As such, the four studies varied along these domains. General and study-specific information about participants, experimental design, treatment, and outcomes is summarized in the sections that follow. An overview of this information is presented in Table 1. This review focuses on four specific questions related to treatment effect and generalization patterns:

1. Does an IDTA approach result in successful acquisition and maintenance of trained word and sentence structures?
2. Does an IDTA approach result in generalization to untrained but linguistically related word and sentence structures?
3. Is there evidence of generalization on measures of discourse production?
4. Is there evidence of generalization on standardized measures of cognition and communication (including measures of functional communication)?

METHODS

Participants

Seven persons with aphasia (five male, two female) participated in the four studies (see Table 2). One individual participated in two of the studies (P2 in Study 1 and P6 in Study 3), resulting in a total of eight replications. All participants were monolingual native English speakers. Participant age ranged from 35 to 68 years (mean = 58.6 years; *SD* = 10.1) and years of education ranged from 13 to 20 years (mean = 15.5 years; *SD* = 2.0). All of the participants were premorbidly right handed except for P7 who was ambidextrous. None of the participants had a history of language, learning, psychiatric, or neurological impairment. In addition, all participants passed a monaural pure tone audiometric screening (500; 1,000; and 2,000 Hz at 40 dB SPL). The diagnosis of aphasia was based on neurological history, results of standardized language testing, and independent clinical judgment by a certified speech-language pathologist specializing in adult neurogenic language disorders. Time postonset ranged from 1 to 7 years (mean = 3.25 years; *SD* = 2.0). All participants signed an informed consent form approved by the institutional review board at Ohio State University or Utah State University.

Participants in the four studies varied with respect to their neuropathology and behavioral profiles. Six of the participants had aphasia secondary to a left hemisphere cerebrovascular accident (CVA), one of these individuals (P4) most likely had an accompanying episode of anoxia at the time of the stroke, and one participant (P8) had aphasia

Table 1. Overview of four studies

	1. Milman, Vega-Mendoza, and Clendenen (2014c)	**2. Milman and Bruner (2010)**	**3. Milman, Clendenen, and Vega-Mendoza (2014b)**	**4. Milman et al. (2014a)**
Participants	Three participants with Broca's aphasia following LH CVA	One participant with Wernicke's aphasia following LH CVA	Three participants with Broca's aphasia following LH CVA	One participant with anomic aphasia following glioblastoma excision
Experimental design	Single subject experimental design with replication × participants and behaviors. Clinician-selected topics (food and hobbies), participant-selected stimuli.	Case study with replication × behaviors. Clinician-selected topics (food and hobbies), participant-selected stimuli.	Single subject experimental design with replication × participants and behaviors. Clinician-selected topics (describing people with adjective lists A and B), clinician- and participant-selected stimuli.	Case study with replication × behaviors. Clinician- and participant-selected language goals, participant-selected topics (sports, daughters' interests) and stimuli.
Treatment dosage	Three 1-hr sessions/week Duration contingent on performance and varied across participants (range: 8–20 weeks/tx block) × 2 tx blocks	Three 1-hr sessions/week Duration contingent on performance and varied (range: 8–16 weeks/treatment block) × 2 tx blocks	Three 1-hr sessions/week Duration contingent on performance and varied across participants (range: 4–16 weeks/tx block) × 2 treatment blocks	Four 1-hr sessions/week × 4 weeks/treatment block × 2 treatment blocks
Treatment				
Lexical task(s)	Fixed stimulus set (10 nouns in first tx block; 10 verbs in second tx block)	Fixed stimulus set (10 nouns in first tx block; 10 verbs in second tx block)	Fixed stimulus set (5 adjectives + 2 pronouns/tx block)	Fixed stimulus set (40 common or proper nouns/tx block)

(*continues*)

Table 1. Overview of four studies (*Continued*)

	1. Milman, Vega-Mendoza, and Clendenen (2014c)	**2. Milman and Bruner (2010)**	**3. Milman, Clendenen, and Vega-Mendoza (2014b)**	**4. Milman et al. (2014a)**
Sentence production tasks	Fixed stimulus set Ten sentences (Pronoun + V + Complement) with trained vocabulary/tx block. Cueing hierarchy included syntactic, reading, and repetition prompts.	Fixed stimulus set Ten sentences (Pronoun + V + Complement) with trained vocabulary/tx block. Cueing hierarchy included syntactic, reading, and repetition prompts.	Fixed stimulus set Ten sentences (Pronoun + V + Adjective) with trained vocabulary/tx block. Cueing hierarchy included syntactic, reading, and repetition prompts.	Open stimulus set Ten new topic-focused comment/questions each session with scaffolding to increase fluency and sentence complexity.
Whole task(s) (discourse)	1. Script training 2. Structured dialogue practice (Clinician asks question to elicit target vocabulary and sentence structure in novel real-world contexts e.g. Clinician: "What are you doing this weekend?" to elicit target response: I am ____") Modified RET protocol used to scaffold correct response	- 1. Structured dialogue practice (Clinician asks question to elicit target vocabulary and sentence structure in novel real-world contexts, e.g., Clinician: "What are you doing this weekend?" to elicit target response: I am ____") Modified RET protocol used to scaffold correct response	- 1. Descriptive discourse task (Clinician asks question about photo of person to elicit opinion with target vocabulary and sentence structure, e.g., Clinician: "What do you think about Willie Nelson?" to elicit target response: "He is ___") Modified RET protocol used to scaffold correct response	- 1. Structured topic-focused conversation (Clinician presents question/ comment about sports/daughters' interests, e.g.: "I hope the Seahawks win the Super bowl this year?" to elicit participant response) Modified RET protocol used to shape/expand fluent response

(*continues*)

Table 1. Overview of four studies (*Continued*)

	1. Milman, Vega-Mendoza, and Clendenen (2014c)	2. Milman and Bruner (2010)	3. Milman, Clendenen, and Vega-Mendoza (2014b)	4. Milman et al. (2014a)
	3. Structured dialogue practice (as above) but in group therapy with other persons with aphasia	-	2. Structured dialogue practice (as above) but in group therapy with other persons with aphasia	2. Topic-focused conversational practice with authentic conversational partners (family and friends)
Outcome and results	Three of three participants acquired the target structures and maintained posttreatment performance level at final posttreatment testing.	Participant acquired the target structures and maintained posttreatment performance level at final posttreatment testing.	Two of three participants acquired the target structures and maintained posttreatment performance level at final posttreatment testing	Participant acquired treatment target (increased CIUs/ utterance) but maintained posttreatment performance level with only one of four conversational partners.
Acquisition and maintenance horizontal generalization	None of the participants showed statistically significant evidence of generalization to related linguistic structures.	Participant showed evidence of generalization (based on visual inspection of graphs)	Two of three participants showed evidence of generalization to related linguistic structures.	Did not assess generalization to related structures.
Generalization to discourse	All three participants showed significant gains on the discourse measures	Participant did not improve on discourse measure	Two of three participants improved on at least one discourse measure.	Participant improved on two of three discourse measures.
Generalization to standardized tests	All three participants showed significant improvement on the WAB-R; only one participant improved on measures of functional communication.	Participant showed significant improvement on the WAB-R and on measure of functional communication (SCCAN).	Two of three participants improved on the WAB-R but not on measures of functional communication. The third participant showed the reverse pattern.	Improved on measures of language impairment and functional communication but not on the WAB-R.

Note. CIUs = Correct Information Units; LH CVA = left hemisphere cerebrovascular accident; RET = Response Elaboration Training; SCCAN = Scales of Cognitive & Communicative Ability for Neurorehabilitation; tx = treatment; WAB-R = Western Aphasia Battery-Revised.

Table 2. Participant characteristics

	Gender	Age	Ed (years)	TPO (years)	Etiology	Aphasia Type (WAB-R)	Aphasia Severity (WAB-R AQ; Maximum = 100)	General Cognitive-Communicative Function (SCCAN)	Accompanying Cognitive-Communicative Impairments
Study 1									
P1	F	66	16	2	LH CVA	Broca's	26.4 (severe)	48	Apraxia (severe)
P2	M	55	14	5	LH CVA	Broca's	36.3 (severe)	45	Apraxia (mild–moderate)
P3	F	66	16	3	LH CVA	Broca's	78.5 (moderate)	69	-
Study 2									
P4	M	58	13	1	LH CVA (with probable episode of anoxia)	Wernicke's	29.4 (severe)	27	General cognitive-communicative impairment affecting nonverbal attention, memory, and problem solving
Study 3									
P5	M	68	20	1	LH CVA	Broca's	35.8 (severe)	50	Impaired visuospatial reasoning
P6	M	56	14	7	LH CVA	Broca's	42.2 (moderate)	46	Apraxia (mild)
P7	M	65	16	2	LH CVA	Broca's	53.1 (moderate)	79	-
Study 4									
P8	M	35	14	5	L frontal lobe glioblastoma excision	Anomic	89.2 (mild)	83	Apraxia (mild–moderate), verbal memory impaired (mild)

Note. Ed = education; F = female; LH CVA = left hemisphere cerebrovascular accident; M = male; TPO = time post onset; WAB-R = Western Aphasia Battery-Revised (Kertesz, 2007); AQ = aphasia quotient; SCCAN = Scales of Cognitive & Communicative Ability for Neurorehabilitation (Milman & Holland, 2012).

secondary to excision of a left frontal lobe glioblastoma. Based on the Western Aphasia Battery-Revised (WAB-R, Kertesz, 2007) classification system, aphasia subtypes included Broca's aphasia (five participants), Wernicke's aphasia (one participant), and anomic aphasia (one participant). The severity of the aphasia ranged from severe (P1: WAB-R Aphasia Quotient = 26.4; mean length of utterance [MLU] = 0) to mild (P8: WAB-R Aphasia Quotient = 89.2; MLU = 9.1). Mild (P6), moderate (P2 and P8), and severe (P1) apraxia of speech were noted in four of the participants. Two of the participants (P4 and P5) had more general cognitive impairments affecting nonverbal attention, memory, and problem solving. As would be expected from these diverse clinical profiles, participants varied with respect to their spoken language production. For the majority of participants (P1, P2, P3, P6, P7), speech production was characterized by omissions, pauses, fillers, false starts, and occasional semantically or phonologically related paraphasias. Contrasting production patterns were observed, however, in P4 (fluent empty speech with frequent paraphasias including unrelated words, perseverations, and neologisms) and P5 (slow halting speech that was largely unintelligible due to frequent neologisms and paraphasic errors).

Experimental design

As indicated in Table 1, the four studies included two case studies with replication across behaviors (Studies 2 and 4) and two single subject multiple baseline design studies with replication across behaviors and participants (Studies 1 and 3). Balanced stimulus sets were developed to assess generalization of treatment effects on the experimental task to untrained stimuli that were semantically related (all four studies) and unrelated (Studies 1, 2, and 3) to trained language structures. Potential vertical (across context) generalization was also evaluated with standardized measures of language (all four studies), functional communication (Studies 1, 3, and 4), general cognition (all four studies), and discourse (all four studies). Regular probes (weekly or biweekly) were administered throughout baseline, treatment, and posttreatment phases. Probe data were collected on trained and untrained stimuli (Studies 1, 2, and 3) or only on untrained stimuli (Study 4). Scoring reliability for probe data was reported in three of the studies and ranged from 90% to 99%.

Standardized testing

The WAB-R (Kertesz, 2007), the Scales of Cognitive and Communicative Ability for Neurorehabilitation (SCCAN, Milman & Holland, 2012), and a discourse assessment (descriptive, narrative, and/or conversational samples) were administered in all four studies to evaluate pre- and posttreatment performance (see Table 1). Further language testing included evaluation of semantic processing (Pyramids and Palm Trees; Howard & Patterson, 1992; Study 3), naming (Boston Naming Test [BNT]; Kaplan, Goodglass, & Weintraub, 2001; Studies 3 and 4), motor-speech control (Apraxia Battery for Adults [ABA]; Dabul, 2000; Study 4), and functional communication (Communicative Effectiveness Index [CETI]; Lomas et al., 1989; Studies 1 and 4; American Speech-Language-Hearing Association Functional Assessment of Communication Skills for Adults (ASHA FACS); Frattali, Thompson, Holland, Wohl, & Ferketic, 1995; Study 3; or Assessment for Living with Aphasia [ALA]; Kagan et al., 2010; Study 4). Additional cognitive testing included assessment of visual-spatial reasoning (Raven's Colored Progressive Matrices [RCPM]; Raven & Court, 1998; Studies 3 and 4) and verbal memory (Digit Span Lezak, 2004; and California Verbal Learning Test; Delis, Kramer, Kaplan, & Ober, 2000; Study 4).

Treatment

Before initiating treatment, participants were interviewed informally (all four studies) and with standardized questionnaires (Studies 1, 3, and 4) to identify their personal communication goals. This information was combined with impairment data collected from the assessment battery to design patient-centered treatment protocols that were tailored to individual participant goals and impairment/ability profiles. Specifically, participants took an active role in designing the treatment by selecting stimulus materials (all four studies), conversational topics (Study 4), and language goals (Study 4).

Three to four 1-hr treatment sessions were administered per week in all four studies. For the first three studies, duration of treatment was contingent on acquisition of the target language structures and varied from 4 to 20 weeks per treatment block. In the fourth study, a predetermined interval (1 month) was assigned to each treatment block. Each therapy session included a core set of three topic-focused intervention tasks (see Table 1). The core therapy tasks included (1) word retrieval based on a semantic—phonological–orthographic cueing hierarchy (Boyle, 2004; Greenwood et al., 2010; Hashimoto & Frome, 2011); (2) sentence production training (see Table 1 for study-specific procedures); and (3) discourse training modeled after Response Elaboration Training (RET; Kearns, 1985; Wambaugh, Nessler, & Wright, 2013). Each of the three tasks was administered for approximately 15 min.

The three tasks administered in each session (and in each treatment block) focused on a specific semantic topic. Topics included hobbies and favorite foods (Studies 1 and 2), describing people (Study 3), and sports and daughters' interests (Study 4). Language goals were similar in focusing on word retrieval, sentence generation, and discourse. Specific goals varied, however, across participants and studies. Selected goals varied in terms of target word class (nouns, verbs, adjectives, or pronouns), sentence structure (simple canonical or multiple complex), or discourse type (descriptive, narrative, or conversational). All participants also completed approximately 30 min of daily homework to practice the target language structures. Additional therapy activities included script training (Study 1), group therapy (Studies 1 and 3), and conversational practice with family and friends (Study 4). More specific details about treatment procedures can be found in Table 1 and in the publications referenced in this table.

Analyses

Analyses included graphic displays of probe data (all four studies) and computation of treatment effect size (Studies 1, 3, and 4) using statistical procedures/interpretation specified in the study by Beeson and Robey (2006). Comparison of pre- and posttreatment performance on standardized tests and discourse analyses (all four studies) involved using interpretation data provided in accompanying test manuals.

RESULTS AND DISCUSSION

RQ1: Does an IDTA approach result in successful acquisition and maintenance of trained word and sentence structures?

Performance on vocabulary and sentence structures that were targeted and trained with individual participants is summarized in Table 3. The majority of participants (seven of eight) acquired the target structures ($\geq$80 on two consecutive probes) and five showed a statistically significant treatment effect. For these participants, treatment effect size varied from small to large. In general, the largest effect size was seen in individuals P1, P2, and P6, who had the most severe aphasia and the lowest WAB-R Aphasia Quotient scores. These participants began treatment at near floor levels, showed little variability in their baseline performance, and made the largest gains on trained tasks during intervention. In contrast, the smallest effect size was seen in individuals P3, P7, and P8, who had more mild aphasia and higher WAB-R aphasia quotient scores. These participants began treatment with higher scores on the experimental tasks, showed greater variability in their baseline performance, and made smaller gains on trained tasks during intervention.

All participants who acquired the target structures maintained a treatment effect, but to varying degrees, after treatment was discontinued. Maintenance was notably more robust in Studies 1, 2, and 3, in which treatment

Table 3. Acquisition, maintenance, and generalization to semantically related structures[a]

	Acquisition of Target Structures (≥ 80% on Two Consecutive Probes)	Treatment Effect Size	Maintenance (at Least 30% Above Baseline Level)	Generalization (at Least 30% Above Baseline Level)
Study 1				
P1	Acquired all target structures	Large	Maintained (final testing 2 months post-tx)	NS
P2	Acquired all target structures	Large	Maintained (final testing 2 months post-tx)	NS
P3	Acquired all target structures	Small	Maintained (final testing 2 months post-tx)	NS
Study 2				
P4	Acquired all target structures	–	Maintained (final testing 2 months post-tx)	Generalization reported (visual inspection only)
Study 3				
P5	No treatment effect	NS	NS	NS
P6	Acquired all target structures	Large	Maintained (final testing 1 week post-tx)	S
P7	Acquired all target structures	NS	Maintained (final testing 1 month post-tx)	S
Study 4				
P8	Acquired treatment target (increased CIUs/utterance with three of four conversational partners)	Small	Maintained (final testing 1 week post tx) but only with one conversational partner	–
Total	Seven of eight acquired the target structures	Five of seven showed a significant treatment effect	Seven of eight maintained immediate posttreatment performance effects at final testing	Three of seven showed generalization to related structures

Note. CIUs = Correct Information Units; NS = not significant; S = significant.
[a]Effect sizes computed using guidelines of Beeson and Robey (2006) for word retrieval, syntactic production, and mixed treatments.

duration was contingent on criterion performance (i.e., 2–3 consecutive sessions with probe scores >80%) than in Study 4, which used a fixed 1-month interval. However, it should be noted that the duration of treatment varied widely when it was contingent on participant performance (mean = 12 weeks; *SD* = 5 weeks, range: 4–20 weeks). As might be expected, individuals with lower initial performance on the probe task, who showed the greatest treatment effect, also required more sessions to reach criterion performance. As discussed elsewhere (Milman et al., 2014c; Robey, 1998), there appears to be a close relation between the number of treatment sessions and the strength (and longevity) of the treatment effect. Thus, shorter treatment durations are likely to result in reduced maintenance of a treatment effect.

One participant (P5) was unable to acquire the target structures. Although this participant was similar to P1, P2, and P6 (all of whom showed a large treatment effect) with respect to his aphasia classification and overall severity, he differed in the nature of his speech production errors. Specifically, the connected speech of P5 was characterized by a high preponderance of neologisms and unintelligible jargon. Although P5 did increase his overall speech output during intervention, this was accompanied by a parallel increase in neologistic errors. Overall, these results suggest that this intervention approach may be successful for individuals with a wide range of aphasia types (fluent and nonfluent) and severity levels (mild to severe). However, this intervention is less likely to be effective for individuals whose connected speech is characterized by frequent neologisms and/or far-from-target responses (see discussion in Milman et al., 2014b).

RQ2: Does an IDTA approach result in successful generalization and maintenance in untrained word and sentence structures?

Three of the studies examined generalization of treatment effects (defined as >30% above baseline level) to words that were semantically related to trained vocabulary. Only three of the seven participants showed evidence of generalization to untrained vocabulary. These results are consistent with findings reported for similar discourse-based treatments that have also shown successful vertical generalization but limited/absent horizontal generalization to untrained language forms (Boyle, 2011; Rider, Wright, Marshall, & Page, 2008). It may be that the focus on vertical generalization and particularly on training individuals to insert words and sentences into discourse directly impacts horizontal generalization effects. More specifically, it may be the case that horizontal (within level) and vertical (across level) language associations are in competition and that specific treatments may selectively stimulate one type of generalization over another.

RQ3: Is there evidence of generalization on measures of discourse production?

Table 4 summarizes results of the discourse analyses (measured as MLU) across the four studies. Six of the eight participants showed statistically significant (>1 standard error of measurement) improvement on at least one of the discourse measures. Although P8 did not improve his MLU score on the descriptive discourse task, he did decrease the number of dysfluencies/utterance. In contrast, two participants (P4 and P5) did not show any improvement on any of the discourse measures. These individuals differed from other participants in two respects. First, their speech output was characterized by frequent "far from target" paraphasic errors. Second, both individuals presented with more general cognitive impairments affecting nonverbal attention, memory, and reasoning. As evident from Table 4 there also appeared to be some variability in results across discourse measures (at least for the three participants assessed on multiple measures).

This is in line with previous research in the aphasia literature showing disparate outcomes across discourse tasks (Fergadiotis, Wright, & West, 2013). In general, narrative and conversational discourse tasks appeared to be more sensitive to change than descriptive discourse. This may have been due to the larger samples collected from narrative

Table 4. Pre- and posttreatment change on descriptive, narrative, and conversational discourse tasks (MLU)[a]

	Descriptive	Narrative	Conversational
	WAB-R Picnic Scene (MLU in morphemes)	**Cinderella Story (MLU in Morphemes[a])**	**Sports and Daughter's Interests (MLU)**
Study 1			
P1	S	–	–
P2	S	–	–
P3	S	–	–
Study 2			
P4	NS (↑ dysfluencies/ utterance)	–	–
Study 3			
P5	–	–	–
P6	NS	S	–
P7	S	S	–
Study 4			
P8	NS (↓ dysfluencies/ utterance)	S	S
Proportion of participants showing significant change in MLU	4/7 (57%)	3/3 (100%)	1/1 (100%)

Note. MLU = mean length of utterance; NS = not significant; WAB-R = Western Aphasia Battery-Revised.
[a]Significance for Descriptive and Narrative Discourse tasks = > 1 SEM (computed from "A system for the linguistic analysis of agrammatic language production," Thompson et al., 1995); Significance for Conversational Discourse based on treatment effect size (using guidelines of Beeson & Robey [2006]).

and conversational samples (see Boles & Bombard, 1998) or it may have been due to the fact that these discourse measures were more closely related to the trained experimental tasks.

RQ4: Is there evidence of generalization on standardized measures of cognition and communication?

Results of pre- and posttreatment testing on standardized measures are presented in Table 5. The majority of participants (six of eight) showed statistically significant pre-posttreatment change (>1 SEM unit) on the WAB-R. In addition, statistically significant increases were seen by three of eight participants on the SCCAN, two of five participants on measures of functional communication, three of four participants on the BNT, one of one participant on the ABA, and one of one participant on measures of verbal working memory. Greatest improvements on the WAB-R were seen on scores for Information Content (four of eight participants), Repetition (five of eight participants), and Naming (five of eight participants). None of the participants showed changes on measures of semantic processing (Pyramids & Palm Trees) or on visual spatial problem solving (RCPM).

These results are largely consistent with predicted outcomes in that changes were evident in isolated components of language production (e.g., word retrieval, information content, motor-speech control) and in broader measures of language functioning (WAB-R and SCCAN). The increase in verbal working memory (seen in one participant

Table 5. Pre- and posttreatment change on standardized measures (statistical significance determined by standardized procedures/data provided in accompanying test manuals)

			Functional Communication			Additional Speech and Language Tests			Additional Cognitive Tests		
	WAB-R AQ	**SCCAN**	**CETI**	**ASHA FACS**	**ALA**	**BNT**	**P&P**	**ABA**	**RCPM**	**Digit Span**	**CVLT**
Study 1											
P1	S	NS	NS	–	–	–	–	–	–	–	–
P2	S	NS	NS	–	–	–	–	–	–	–	–
P3	S	S	S	–	–	–	–	–	–	–	–
Study 2											
P4	S	S	–	–	–	–	–	–	–	–	–
Study 3											
P5	NS	S	–	–	–	S	NS	–	NS	–	–
P6	S	NS	–	–	–	S	NS	–	NS	–	–
P7	S	NS	–	NS	–	NS	NS	–	NS	–	–
Study 4											
P8	NS	NS	S	–	S	S	–	S	NS	S	S
Total participants showing significant gain	6 of 8 (75%)	3 of 8 (38%)	2 of 4 (50%)	0 of 1 (0%)	1 of 1 (100%)	3/4 (75%)	0/3 (0%)	1/1 (100%)	0/4 (0%)	1/1 (100%)	1/1 (100%)

Note. ABA = Apraxia Battery for Adults (Dabul, 2000); ALA = Assessment for Living with Aphasia (Kagan et al., 2010); ASHA FACS = ASHA Functional Assessment of Communication Skills for Adults (Frattali et al., 1995); BNT = Boston Naming Test (Kaplan et al., 2001); CETI = Communicative Effectiveness Index (Lomas et al., 1989); CVLT = California Verbal Learning Test (Delis et al., 2000); Digit Span (from Lezak, 2004); NS = not significant; P&P = Pyramids & Palm Trees (Howard & Patterson, 1992); RCPM = Raven's Colored Progressive Matrices (Raven & Court, 1998); S = significant; SCCAN = Scales of Cognitive & Communicative Ability for NeuroRehabilitation (Milman & Holland, 2012); WAB-R AQ = Western Aphasia Battery-Revised (Kertesz, 2007) Aphasia Quotient.

who was tested in this area) was not anticipated. Nonetheless, this outcome makes sense within the context of the broader literature that establishes a strong connection between language production and verbal working memory (Acheson, Hamidi, Binder, & Postle, 2011; Martin & Reilly, 2012; Martin & Slevc, 2014).

One disappointing finding was the limited change seen in functional communication measures (statistically significant changes seen in only two of five participants). These results suggest that the integrated treatment approach reviewed here did not fully transfer to everyday use. Interestingly, the two participants who showed statistically significant improvements on functional outcome measures (P3 and P8) also had the highest WAB-R AQ scores (mildest language impairment). It may be the case that participants with milder language impairments were better equipped to achieve the more demanding "far" transfer to functional communication than were participants with more severe language impairments.

P8 also differed from other participants in two important ways that may have contributed to his relatively strong performance on the functional outcome measures. First, the etiology of P8's aphasia (posttumor) differed from other participants (poststroke). Second, P8 participated in the fourth and final study. In this study, we intentionally increased participant involvement in treatment design (P8 selected not only the stimuli but also the communication goal and the discourse topic) with the hope of increasing transfer to everyday communicative tasks.

A final consideration centers on measurement properties of the tests. Whereas all three functional measures use a similar rating system, the scale end points differ for the CETI ("Not at all able" → "as able as before"), ALA ("Big problem → "No problem"), and ASHA FACS ("Does not" → "Does"). Notably, the ASHA FACS rating scale quantifies independence rather than success/ease/frequency of (independently) completing the task. As a result, this measure may be less sensitive to gains in performance for individuals who are already independent but change with respect to the success/ease/frequency of task completion. Since generalization of learning to everyday use is the primary goal of this and many other therapy approaches, understanding, improving, and measuring such generalization remains a central objective for future research in this area.

SUMMARY AND CONCLUSIONS

An Integrated Discourse Treatment approach aims to facilitate transfer of impairment-based language learning to real-world everyday communicative use. This study reviewed data for this approach from a series of four related studies. Most participants (seven of eight) acquired the target structures and maintained treatment effects for a period of time after intervention had ended. While approximately half of the participants (three of seven) showed evidence of *horizontal* generalization to related linguistic structures, almost all participants demonstrated evidence of *vertical* learning transfer to at least one standardized cognitive-communicative measure (all participants) and to measures of discourse production (six of eight participants). Collectively, the data reported above provide preliminary support for this treatment approach.

A broader goal of this phase II treatment research was to identify characteristics of individuals who are most likely to benefit from this intervention, refine treatment procedures, and identify the most sensitive outcome measures. Results of these studies suggest that this intervention is likely to benefit individuals who vary widely in aphasia subtype (fluent/nonfluent), severity (mild to severe), and etiology. This treatment was judged to be ineffective, however, in two individuals (P4 and P5) who presented with more diffuse (nonverbal) cognitive impairments, frequent neologisms, and far-from-target paraphasic errors. These behavioral characteristics, therefore, appear to be negative prognostic indicators for this intervention. Treatment parameters associated with greatest gains included (1) high personal relevance (participant-selected stimuli and communication/discourse goals); (2) duration/dosage based on successful acquisition of target word/sentence structures (≥80% on two consecutive probes); and (3) close relation between trained/practiced discourse task and target communication goal. The standardized measures that most consistently identified posttreatment change included the WAB-R, BNT, narrative and/or conversational discourse, the CETI, and ALA.

Finally, based on the data summarized previously and elsewhere, there appears to be considerable variability in the outcomes observed across individuals. Notably, individuals with relatively severe expressive language impairment tended to show a large treatment effect on component tasks (word and sentence production) but limited generalization on discourse production. Conversely, individuals with relatively mild expressive language impairment showed a smaller treatment effect on component tasks but greater generalization on discourse production. An important goal for future research will be to test these preliminary findings with a larger clinical population and greater experimental control of critical treatment variables.

REFERENCES

Acheson, D. J., Hamidi, M., Binder, J. R., & Postle, B. R. (2011). A common neural substrate for language production and verbal working memory. *Journal of Cognitive Neuroscience*, *23*(6), 1358–1367.

Antonucci, S. M. (2009). Use of semantic feature analysis in group aphasia treatment. *Aphasiology*, *23*(7–8), 854–866.

Beeson, P. M., & Robey, R. R. (2006). Evaluating single-subject treatment research: Lessons learned from the aphasia literature. *Neuropsychology Review*, *16*(4), 161–169.

Beukelman, D. R., Fager, S., Ball, L., & Dietz, A. (2007). AAC for adults with acquired neurological conditions: A review. *Augmentative and Alternative Communication*, *23*(3), 230–242.

Boles, L. (2011). A review of aphasia couples therapy. *Asia Pacific Journal of Speech, Language and Hearing*, *14*(3), 159–163.

Boles, L., & Bombard, T. (1998). Conversational discourse analysis: Appropriate and useful sample sizes. *Aphasiology*, *12*(7–8), 547–560.

Boyle, M. (2004). Semantic feature analysis treatment for anomia in two fluent aphasia syndromes. *American Journal of Speech-Language Pathology*, *13*, 236–249.

Boyle, M. (2011). Discourse treatment for word retrieval impairment in aphasia: The story so far. *Aphasiology*, *25*(11), 1308–1326.

Cameron, R. M., Wambaugh, J. L., Wright, S. M., & Nessler, C. L. (2006). Effects of a combined semantic/phonologic cueing treatment on word retrieval in discourse. *Aphasiology*, *20*(2–4), 269–285.

Dabul, B. L. (2000). *Apraxia battery for adults*. Austin, TX: Pro-ed.

Delis, D. C., Kramer, J. H., Kaplan, E., & Ober, B. A. (2000). *California verbal learning test* (2nd ed.). Bloomington, MN: Pearson.

Elman, R. J., & Bernstein-Ellis, E. (1999). The efficacy of group communication treatment in adults with chronic aphasia. *Journal of Speech, Language, and Hearing Research*, *42*, 411–419.

Falconer, C., & Antonucci, S. M. (2012). Use of semantic feature analysis in group discourse treatment for aphasia: Extension and expansion. *Aphasiology*, *26*(1), 64–82.

Fergadiotis, G., Wright, H. H., & West, T. M. (2013). Measuring lexical diversity in narrative discourse of people with aphasia. *American Journal of Speech-Language Pathology*, *22*(2), S397–S408.

Frattali, C. M., Thompson, C. K., Holland, A., Wohl, C. B., & Ferketic, M. M. (1995). *The American speech-language-hearing association functional assessment of communication skills for adults*. Rockville, MD: ASHA.

Gaddie, A., Kearns, K. P., & Yedor, K. (1991). A qualitative analysis of response elaboration training effects. *Clinical Aphasiology*, *19*, 171–183.

Goral, M., & Kempler, D. (2009). Training verb production in communicative context: Evidence from a person with chronic non-fluent aphasia. *Aphasiology*, *23*(12), 1383–1397.

Greenwood, A., Grassly, J., Hickin, J., & Best, W. (2010). Phonological and orthographic cueing therapy: A case of generalised improvement. *Aphasiology*, *24*(9), 991–1016.

Hashimoto, N., & Frome, A. (2011). The use of a modified semantic features analysis approach in aphasia. *Journal of Communication Disorders*, *44*(4), 459–469.

Holland, A. L., Halper, A. S., & Cherney, L. R. (2010). Tell me your story: Analysis of script topics selected by persons with aphasia. *American Journal of Speech-Language Pathology*, *19*(3), 198–203.

Holland, A. L., & Hinckley, J. J. (2002). Assessment and treatment of pragmatic aspects of communication in aphasia. InA. E. Hillis (Ed.), *The handbook of adult language disorders: Integrating cognitive neuropsychology, neurology, & rehabilitation* (pp. 413–427). New York, NY: Psychology Press.

Howard, D., & Patterson, K. E. (1992). *The Pyramids and Palm Trees Test: A test of semantic access from words and pictures*. London, Thames Valley Test Company.

Hung, W. (2013). Problem-based learning: A learning environment for enhancing learning transfer. *New Directions for Adult and Continuing Education*, *2013*(137), 27–38.

Kagan, A. (1998). Supported conversation for adults with aphasia: Methods and resources for training conversation partners. *Aphasiology*, *12*(9), 816–830.

Kagan, A., Simmons-Mackie, N., Rowland, A., Huijbregts, M., Shumway, E., McEwen, S., et al. (2010). *Assessment for living with aphasia*. Toronto, ON: Aphasia Institute.

Kaplan, E., Goodglass, H., & Weintraub, S. (2001). *Boston naming test*. Austin, Texas: Pro-ed.

Kearns, K. (1985). Response elaboration training for patient initiated utterances. *Clinical Aphasiology*, *15*, 196–204.

Kertesz, A. (2007). *Western aphasia battery-revised*. San Antonio, TX: PsychCorp.

Kimbarow, M. L. (2007). Integrating life participation approaches to aphasia treatment with adult learning theory: A synergistic approach. *Topics in Language Disorders*, *27*(4), 318–323.

Kiran, S., & Thompson, C. K. (2003). 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*, *46*, 773–787.

Lezak, M. D. (Ed.). (2004). *Neuropsychological assessment*. New York, NY: Oxford University Press.

Life-Participation Approach to Aphasia Project Group (Chapey, R., Duchan, J. F., Elman, R. J. L., Garcia, L. J., Kagan, A., Lyon, J. G., & Simmons-Mackie, N.) (2008). Life-participation approach to aphasia: A statement of values for the future. In R. Chapey (Ed.), *Language intervention strategies in aphasia and related neurogenic communication disorders* (5th ed.). Baltimore, MD: Lippincott, Williams, & Wilkins.

Lomas, J., Pickard, L., Bester, S., Elbard, H., Finlayson, A., & Zoghaib, C. (1989). The communicative effectiveness index: Development and psychometric evaluation of a functional communication measure for adult aphasia. *Journal of Speech and Hearing Disorders*, *54*, 113–124.

Martin, N., & Reilly, J. (2012). Short-term/working memory impairments in aphasia: Data, models, and their application to aphasia rehabilitation. *Aphasiology*, *26*(3–4), 253–257.

Martin, R. C.,, & Slevc, L. R. (2014). Language production and working memory. InM. Goldrick, V. Ferreira, & M. Miozzo (Eds.), *The Oxford handbook of language production* (pp. 437–450). New York, NY: Oxford University Press.

Milman, L., Anderson, E., Thatcher, K., Amundson, D., Christensen, C., Ellet, C., et al. (2014a). An integrated part-whole conversation therapy for aphasia: Comparing face-to-face and telerehabilitation contexts. Orlando, FL: American Speech-Language-Hearing Convention.

Milman, L. H., & Bruner, J. (2010, November). *Integrated therapy for aphasia: Application in fluent aphasia*. Paper presented at the American Speech-Language-Hearing Convention, Philadelphia, PA.

Milman, L., Clendenen, D., & Vega-Mendoza, M. (2014b). Production and integrated training of adjectives in three individuals with nonfluent aphasia. *Aphasiology*, *28*(10), 1198–1222.

Milman, L., & Holland, A. (2012). *The scales of cognitive and communicative ability for neurorehabilitation (SCCAN)*. Austin, TX: ProEd.

Milman, L., Vega-Mendoza, M., & Clendenen, D. (2014c). Integrated training for aphasia: An application of part–whole learning to treat lexical retrieval, sentence production, and discourse-level communications in three cases of nonfluent aphasia. *American Journal of Speech-Language Pathology*, *23*(2), 105–119.

Murray, L., Timberlake, A., & Eberle, R. (2007). Treatment of underlying forms in a discourse context. *Aphasiology*, *21*(2), 139–163.

Peach, R. K., & Reuter, K. A. (2010). A discourse-based approach to semantic feature analysis for the treatment of aphasic word retrieval failures. *Aphasiology*, *24*(9), 971–990.

Peach, R., & Wong, P. (2004). Integrating the message level into treatment for agrammatism using story retelling. *Aphasiology*, *18*(5–7), 429–441.

Raven, J. C., & Court, J. H. (1998). *Raven's progressive matrices and vocabulary scales*. San Antonio, Texas: Pro-ed.

Raymer, A. M., Beeson, P., Holland, A., Kendall, D., Maher, L. M., Martin, N., et al. (2008). Translational research in aphasia: From neuroscience to neurorehabilitation. *Journal of Speech, Language, and Hearing Research*, *51*(1), S259-S275.

Rider, J. D., Wright, H. H., Marshall, R. C., & Page, J. L. (2008). Using semantic feature analysis to improve contextual discourse in adults with aphasia. *American Journal of Speech-Language Pathology*, *17*(2), 161–172.

Robey, R. R. (1998). A meta-analysis of clinical outcomes in the treatment of aphasia. *Journal of Speech, Language, and Hearing Research*, *41*(1), 172–187.

Robey, R. R. (2004). A five-phase model for clinical-outcome research. *Journal of Communication Disorders*, *37*(5), 401–411.

Savage, M. C., Donovan, N. J., & Hoffman, P. R. (2014). Preliminary results from conversation therapy in two cases of aphasia. *Aphasiology*, *28*(5), 616–636.

Schmidt, H. G. (1993). Foundations of problem-based learning: Some explanatory notes. *Medical Education*, *27*(5): 422–432.

Sorin-Peters, R. (2003). Viewing couples living with aphasia as adult learners: Implications for promoting quality of life. *Aphasiology*, *17*(4), 405–416.

Thompson, C. K., Shapiro, L. P., Kiran, S., & Sobecks, J. (2003). 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*, *46*(3), 591–607.

Thompson, C. K., Shapiro, L. P., Tait, M. E., Jacobs, B. J., Schneider, S. L., & Ballard, K. J. (1995). A system for the linguistic analysis of agrammatic language production. *Brain and Language*, *51*, 124–129.

Wambaugh, J. L., Mauszycki, S., Cameron, R., Wright, S., & Nessler, C. (2013). Semantic feature analysis: Incorporating typicality treatment and mediating strategy training to promote generalization. *American Journal of Speech-Language Pathology*, *22*(2), S334–S369.

Wambaugh, J. L., Nessler, C., & Wright, S. (2013). Modified response elaboration training: Application to Procedural discourse and personal recounts. *American Journal of Speech-Language Pathology*, *22*(2), S409–S425.

Webster, J., Whitworth, A., & Morris, J. (2015). Is it time to stop "fishing"? A review of generalisation following aphasia intervention. *Aphasiology*, *29*(11), 1240–1264.

Whitworth, A., Leitão, S., Cartwright, J., Webster, J., Hankey, G. J., Zach, J., et al. (2015). NARNIA: A new twist to an old tale. A pilot RCT to evaluate a multilevel approach to improving discourse in aphasia. *Aphasiology*, *29*(11), 1345–1382.

Wilkinson, R., & Wielaert, S. (2012). Rehabilitation targeted at everyday communication: Can we change the talk of people with aphasia and their significant others within conversation? *Archives of Physical Medicine and Rehabilitation*, *93*(1), S70–S76.