Research Article

Intensive Language Action Therapy in Chronic Aphasia: A Randomized Clinical Trial Examining Guidance by Constraint

Jacquie Kurland,^a Edward J. Stanek, III,^b Polly Stokes,^a
Minming Li,^b and Mary Andrianopoulos^a

Purpose: Intensive language action therapy (ILAT) can be effective in overcoming learned nonuse in chronic aphasia. It is suggested that all three guiding principles (constraint, communication embedding, massed practice) are essential to ILAT's success. We examined whether one of these, guidance by constraint, is critical.

Method: Twenty-four participants with aphasia (PWAs) were assigned to ILAT or a modified version of promoting aphasic communicative effectiveness (PACE) in a randomized block, single-blind, parallel-group treatment study. Blocking was by severity (mild/moderate, moderate to severe, severe). Both groups received intensive treatment in the context of therapeutic language action games. Whereas the ILAT group was guided toward

spoken responses, the PACE group could choose any response modality.

Results: All participants, whether assigned to ILAT or PACE groups, improved on the primary outcome measure, picture naming. There was a Severity $\times$ Treatment interaction, with the largest effects estimated for PWAs with mild/moderate and moderate to severe aphasia. Regardless of severity, the ILAT group outperformed the PACE group on untrained pictures, suggesting some benefit of ILAT to generalization. However, this difference was not statistically significant.

Conclusion: Although the groups differed in subtle ways, including better generalization to untrained pictures for ILAT, the study was inconclusive on the influence of guidance by constraint.

Although the majority of the approximately 800,000 persons experiencing a new or recurrent stroke each year in the United States survive the incident, stroke remains a leading cause of serious and chronic disability (Go et al., 2014). Approximately 100,000 persons are discharged annually from U.S. hospitals with a diagnosis of aphasia (Ellis, Dismuke, & Edwards, 2010), leaving them with chronic, persistent poststroke language impairments that significantly restrict their ability to participate in routine activities associated with a meaningful life.

Whereas initial severity is the single most important factor in predicting language recovery in aphasia (Pedersen, Jorgensen, Nakayama, Raaschou, & Olsen, 1995; Pedersen, Vinter, & Olsen, 2004), the influence of the many intrinsic

and extrinsic variables contributing to individual variability in treatment outcomes remains unknown. Evidence has suggested that therapeutic gains are greater, and perhaps longer lasting, with intensive therapy schedules (Basso & Caporali, 2001; Bhogal, Teasell, & Speechley, 2003). Until very recently, the trend has favored this "more is better" framework for treatment intensity (Cherney, Patterson, Raymer, Frymark, & Schooling, 2008). This framework has recently been called into question, however, with one systematic review finding equivocal results in studies directly comparing intensive with nonintensive schedules of treatment delivery (Cherney, Patterson, & Raymer, 2011). It should be noted, however, that the review only included five studies, most of which had low numbers of participants. This undoubtedly contributed to the authors' conclusion that the more is better framework warrants "cautious reexamination."

One means of providing more treatment is through the use of massed practice, one of three fundamental principles believed to support the beneficial effects of a family of aphasia treatment techniques collectively referred to as *intensive language action therapy* (ILAT; Pulvermüller & Berthier, 2008). ILAT has roots in two disparate branches of poststroke intervention: (a) pragmatic communication

^aDepartment of Communication Disorders, University of Massachusetts Amherst

^bDepartment of Biostatistics and Epidemiology, University of Massachusetts Amherst

Correspondence to Jacquie Kurland: jkurland@comdis.umass.edu

Editor: Anastasia Raymer

Associate Editor: Janet Patterson

Received September 14, 2015

Revision received March 12, 2016

Accepted July 19, 2016

DOI: 10.1044/2016_AJSLP-15-0135

Disclosure: The authors have declared that no competing interests existed at the time of publication.

therapy (Aten, Caligiuri, & Holland, 1982; Davis & Wilcox, 1985; Pulvermüller & Roth, 1991) and (b) constraint-induced motor therapy (Taub et al., 1994; Taub, Crago, & Uswatte, 1998; Taub, Uswatte, & Elbert, 2002). First conceived by Pulvermüller et al. (2001) as *constraint-induced language therapy* (CILT), ILAT's fundamental principles derive from basic tenets of neuroscience research, such as Hebbian learning (i.e., "neurons that fire together, wire together"). Three basic principles constitute the core philosophy of ILAT aimed at overcoming the deleterious effects of learned nonuse (Taub et al., 1994). These three principles as applied to overcoming learned nonuse of speech are: (a) *massed practice*: administer therapy in high, concentrated doses; (b) *communication embedding*: encourage use of the speech modality in the context of relevant communicative exchanges; and (c) *guidance by constraint*: focus patients on their remaining language abilities, in particular those they have been avoiding using (i.e., speech). In practice, this last principle has evolved from the highly prescribed use of "constraint" in which pointing and gesturing were "not permitted" (Pulvermüller et al., 2001, p. 1622) to the notion of *guidance by constraint* (also known as *focusing*), in which compensatory means for communicating (e.g., drawing, writing, or gesturing) are permitted, as long as they are used to help elicit a spoken response (Berthier & Pulvermüller, 2011; Difrancesco, Pulvermüller, & Mohr, 2012; Pulvermüller & Berthier, 2008). Thus, in the last decade, the conceptualization of constraint has changed considerably, from one that was interpreted as focused solely on a response modality to one that focuses on the manner in which language action games, stimulus materials, explicit turn-taking rules, and behavioral techniques such as modeling, shaping, and positive reinforcement support overcoming learned nonuse. During this time, a growing number of investigators has replicated positive language outcomes first described by Pulvermüller et al. (2001), following as little as 2 weeks of intensive therapy incorporating these three basic principles (Breier, Maher, Novak, & Papanicolaou, 2006; Kurland, Baldwin, & Tauer, 2010; Kurland, Pulvermüller, Silva, Burke, & Andrianopoulos, 2012; Maher et al., 2006; Meinzer et al., 2006, 2008).

Whether and to what degree massed practice, communication embedding, and/or guidance by constraint are responsible for such improvement is still—and should be—under investigation (Cherney et al., 2008, 2011; Cherney, Patterson, Raymer, Frymark, & Schooling, 2010). All three principles are based on fundamental principles of aphasia therapy originating in the neuroscience of learning (Berthier & Pulvermüller, 2011; Pulvermüller & Berthier, 2008), and hence these authors have recommended applying them together. Nonetheless, it is useful to understand their relative contributions, particularly because there may be barriers (e.g., insurance company limits on the provision of reimbursable speech-language pathology services) that impede the implementation of massed practice in the delivery of ILAT in the real world.

Of the three fundamental principles, massed practice has been examined more than the others. It was one of the variables under study in the original investigation suggesting

that motor rehabilitation-based constraint-induced motor therapy techniques could be applied to language recovery (Pulvermüller et al., 2001). In their seminal study, treatment outcomes in a CILT group were compared with those of a "standard aphasia therapy" group. The latter included service delivery over a 3- to 5-week period, compared with the schedule of delivery for CILT that included massed practice over a shorter period (3 hr/day over 10 days). Only the CILT group made overall language gains and increased the amount of day-to-day communication as assessed by a communication activity log (Pulvermüller et al., 2001). It is important to note that their study was not designed to tease apart the contribution of massed practice.

Other investigators have attempted to compare different treatment protocols that use massed practice. Barthel, Meinzer, Djundja, and Rockstroh (2008) compared treatment outcomes following application of model-oriented aphasia therapy (MOAT, a model-based approach that was based on Nickels, 2002) in 12 participants with chronic aphasia (PWAs) to outcomes previously described for *constraint-induced aphasia therapy* (CIAT; Meinzer, Djundja, Barthel, Elbert, & Rockstroh, 2005). In the CIAT study, 12 of 27 participants had been treated with CIAT (also known as CILT), whereas the other 15 underwent CIATplus (i.e., CIAT with additional reading and other training modules) and involvement of family members. Both CIAT/CIATplus and MOAT treatments were administered using massed practice (i.e., 3 hr/day for 10 consecutive workdays) and both featured shaping as a principal component of constraint. Only the CIAT treatment, however, utilized constraint of nonverbal communication. The MOAT treatment was individually administered and included an emphasis on written language, with no constraint to spoken responses. Although the significant improvements by 11 of 12 PWAs undergoing MOAT were encouraging, the authors noted that varying so many elements between treatment approaches was suboptimal for evaluating the influence of any particular element.

One study that tightly controlled two of the three guiding principles of CILT (massed practice and communication embedding) and focused only on the contribution of constraint was that of Maher et al. (2006). They studied two groups, those receiving CILT ($n = 4$) and a second group ($n = 5$) receiving a modified version of *promoting aphasic communicative effectiveness* (PACE; Davis & Wilcox, 1985). Both groups engaged in massed practice (3 hr/day, 4 days/week for 2 weeks). Both groups worked on similar therapy tasks in a game-like environment (mostly dual-card matching tasks). The one difference between groups was in the availability, cuing, shaping, and reinforcement of methods of communication. Whereas the PACE participants could choose from any modality that might lead to a successful conclusion of a turn, the CILT participants were strictly limited to spoken verbal responses. This stipulation was enforced by the use of visual barriers, reminders to attempt only speech in therapy (and at home), and even a suggestion that participants "sit on their hands" if necessary. Following treatment, both groups appeared to benefit from treatment, although more of the participants in the CILT

group showed more meaningful change on measures of the Western Aphasia Battery (WAB; Kertesz, 1982), Boston Naming Test (BNT; Kaplan, Goodglass, & Weintraub, 2001), and Action Naming Test (Nicholas et al., 1985; as cited in Maher et al., 2006). The strongest responder in the PACE group (P5) refused to engage in any modality other than spoken responses, thus muddying the contrast between PACE and CILT. There were no statistically significant differences in treatment outcomes between the groups, although P5 certainly contributed to this null effect. Moreover, the authors noted that this was a small sample of convenience and they recommended further investigation with a larger sample in a randomized clinical trial.

There has been a dearth of randomized clinical trials comparing CIAT/CILT/ILAT with other approaches to speech-language therapy that might illustrate which of the three principle components contribute most to language recovery. One exception is the recent single-blind, randomized clinical trial comparing CIAT with conventional aphasia therapy and using a modified CIAT schedule in patients with subacute stroke (Sickert, Anders, Munte, & Sailer, 2014). Participants ($n = 100$) were randomly assigned to either a CIAT group or a standard aphasia treatment group and were treated 2 hr/day over 15 days to accommodate a typical subacute therapeutic setting. In addition to modifying the schedule of massed practice from 30 hr over 2 weeks to 30 hr over 3 weeks in both groups, the modified CIAT training included some writing exercises, although this group did not receive as much writing practice as the standard group. Rather than engaging in language action games, patients randomly assigned to receive conventional aphasia therapy participated in exercises focusing on training specific deficits (e.g., sentence completion and other word retrieval tasks), repeating and following instructions. Both groups made significant gains in at least one subtest of a standardized test of aphasia (Aachener Aphasia Test [AAT]; Huber, Poeck, Weniger, & Willmes, 1983). There were no significant differences between the CIAT and conventional groups on either language improvement or on the communication activity log. Thus, Sickert et al. suggest that massed practice is a critical component but that the intensity could be varied (i.e., from 3 hr/day over 2 weeks to 2 hr/day over 3 weeks) to accommodate real conditions in a rehabilitation center. Their results could not directly assess the relative contributions of guidance by constraint or communication embedding (i.e., how either or both might be critical to the success of CIAT in the subacute population). Moreover, the authors noted that it is impossible to account for the contribution of spontaneous recovery in these patients.

The aim of the current study was to evaluate whether one of the three principal components of ILAT (i.e., guidance by constraint) is a critical component of ILAT in chronic aphasia. Thus, we aimed to determine whether gains in naming following ILAT would exceed those observed following intensive language therapy that is not constrained to overt speech practice (i.e., unconstrained language therapy; hereinafter referred to as PACE). Whereas Maher et al. (2006) asked a similar question and

found no difference between groups, their sample size ($n = 9$) was small, and their definition of *constraint* was influenced by the original constraint-induced aphasia treatment study (Pulvermüller et al., 2001). As noted earlier, in this study, the authors explicitly stated that “all communication had to be performed by use of spoken words or sentences; pointing or gesturing was not permitted” (p. 1622). This instruction led to a number of early attempts to replicate CIAT/CILT in which this strong form of constraint was included in the methodology. However, Pulvermüller and others have since refined the principle to guidance by constraint (i.e., focusing), noting that the aim of avoiding learned nonuse is not to prevent the use of gestures altogether but rather to make it more difficult for persons with aphasia to perform communicative actions without using spoken language (Berthier & Pulvermüller, 2011; Difrancesco et al., 2012; Pulvermüller & Berthier, 2008). Moreover, as Meinzer et al. (2012) noted, many recent studies have allowed gesturing during constraint-induced treatment (e.g., Meinzer, Elbert, Djundja, Taub, & Rockstroh, 2007) due to the interdependence of language, action, and perception (Pulvermüller & Fadiga, 2010) and the potential for gesture to facilitate language processing (Rose, 2012).

In the current study, we thus aimed to examine the influence of one fundamental principle of ILAT, namely guidance by constraint, on treatment outcomes by comparing two treatment groups, ILAT and PACE. On the basis of pilot data (Kurland et al., 2012), our working hypothesis was that all participants, even those with persistent severe nonfluent aphasia and comorbid apraxia of speech (AOS) would show gains from either treatment but that those randomly assigned to ILAT would show greater gains, especially on the primary outcome measure, confrontation naming of actions and objects.

Method

Participants

Twenty-seven PWAs were recruited through advertising or referrals from local speech-language pathologists. The institutional review board of the University of Massachusetts Amherst approved the study, and signed informed consent was obtained from all participants before commencing assessment and treatment. Three participants dropped out of the study for personal reasons prior to completing the baseline testing.

Twenty-four individuals (nine women, 15 men), mean age = 66.8 years ($SD = 8.2$; range = 47.3–81 years) participated in this study (see Table 1 for individual demographic/clinical characteristics). All participants were chronically aphasic, and most ($n = 21$) reported a medical history that included a single, unilateral, left hemisphere, middle cerebral artery stroke. Most participants ($n = 22$) were premorbidly right-handed according to self-report on the Edinburgh Inventory (Oldfield, 1971). The average time poststroke was 26.7 months ($SD = 35.3$ months; range = 6–142 months). Most participants ($n = 21$) were monolingual English speakers.

Table 1. Demographic and clinical characteristics of participants.

ID	Treatment Arm	Dyad	Age (years)	Gender	Premorbid handedness	Time postonset (months)	Etiology (CVA)	Aphasia classification	Aphasia severity block	Presence, absence, severity of AOS	Education (years)	Primary site of lesion
P1	ILAT	A	71.00	F	R	13	Isch	Anomic	mod/sev	mild	16	F-T
P2	ILAT	A	55.33	M	R	23	Isch	Mixed TC	mod/sev	absent	14	F-T-P
P3	ILAT	B	64.08	M	R	24	Isch	Anomic	mild/mod	absent	16	—
P4	ILAT	B	77.75	F	R	7	Isch	Anomic	mild/mod	mild	16	F-T-P
P5	ILAT	C	74.33	M	R	57	Isch	Wernicke	severe	absent	16	F-T-P
P6	ILAT	C	70.00	M	R	6	Isch	Wernicke	severe	absent	12	T-P
P7	ILAT	D	70.50	F	R	108	Hem	Broca	mod/sev	mod-marked	12	F-T-P
P8	ILAT	D	74.67	M	R	142	Isch	TCM	mod/sev	mild/mod	20	F-T
P9	ILAT	E	74.92	F	R	25	Isch	Anomic	mild/mod	absent	16	T-P
P10	ILAT	E	67.50	M	R	8	Isch	Conduction	mild/mod	mild	17	T
P11	ILAT	F	62.50	M	R	7	Hem	Anomic	mild/mod	absent	16	T-P-O
P12	ILAT	F	60.83	M	R	7	Hem	Broca	mod/sev	mod-marked	11	F-T-P
P13	PACE	A	63.00	M	R	9	Isch	CrWernicke	mild/mod	BL mild	16	RH T-P
P14	PACE	A	59.67	M	R	82	Isch	Wernicke	mod/sev	mild/mod	16	T-P
P15	PACE	B	67.83	F	R	25	Isch	Anomic	mild/mod	mild/mod	14	F-T-P
P16	PACE	B	81.00	F	R	16	Isch	TCS	mod/sev	mild	12	—
P17	PACE	C	68.42	M	R	10	Isch	Global	severe	DNT	5	T-P
P18	PACE	C	61.75	M	L	11	Isch	Global	severe	severe	18	F-T
P19	PACE	D	64.50	F	R	12	Isch	TCM	mod/sev	mild	16	F
P20	PACE	D	79.83	M	R	12	Hem	Broca	severe	mod/sev	8	F
P21	PACE	E	68.83	M	L	9	Isch	TCS	mild/mod	absent	20	T-P-O
P22	PACE	E	47.33	M	R	13	Hem	Anomic	mod/sev	absent	16	T
P23	PACE	F	63.00	F	R	8	Isch	Anomic	mild/mod	absent	18	—
P24	PACE	-	55.00	F	R	7	Hem	Optic	mild/mod	absent	16	T-O

Note. CVA = cerebrovascular accident; AOS = apraxia of speech; ILAT = intensive language action therapy; R = right; Isch = ischemic; mod = moderate; sev = severe; F = frontal; T = temporal; Mixed TC = mixed transcortical aphasia; P = parietal; Hem = hemorrhagic; TCM = transcortical motor aphasia; O = occipital; PACE = promoting aphasic communicative effectiveness; CrWernicke = crossed Wernicke aphasia; BL = borderline; RH = right hemisphere; TCS = transcortical sensory aphasia; DNT = did not test; L = left. Em dashes indicate data not available.

Most participants ($n = 16$) had at least a bachelor's degree, but three did not finish high school. One participant (P20) reported a history of developmental dyslexia. None of the other participants reported any history of developmental learning disabilities. The presence and severity of AOS were determined by an experienced speech-language pathologist's evaluation of each participant's results on the Apraxia Battery for Adults–Second Edition (Dabul, 2000) and video-recorded speech samples, which were analyzed using the clinical criteria and cardinal signs of AOS published by Darley, Aronson, and Brown (1975); Wambaugh, Duffy, McNeil, Robin, and Rogers (2006); McNeil, Robin, and Schmidt (2009); and Duffy (2012). The acoustic-perceptual features of each participant's speech samples were also evaluated and quantified using the Consensus Auditory-Perceptual Evaluation of Voice (Kempster, Gerratt, Verdolini Abbott, Barkmeier-Kraemer, & Hillman, 2009). Last, each participant was assigned an overall rating for the presence/absence and severity of AOS on the basis of the outcomes obtained from all test procedures and protocols using the following 5-point scale (0 = *normal–no evidence of AOS*; 1 = *mild AOS*; 2 = *moderate AOS*; 3 = *marked AOS*; and 4 = *severe AOS*). The Apraxia Battery for Adults–Second Edition and speech samples for each participant were also scored and analyzed independently by a trained, second judge to assess reliability. Any discrepancies between the

two judges' ratings for severity and/or diagnosis were resolved by reviewing the video recordings and achieving consensus. Interrater reliability on 17.4% of the data revealed what Landis and Koch identify as “substantial” agreement at .78 (.61–.80; Landis & Koch, 1977).

The study was a randomized block, single-blind, parallel-group study. Participants were blocked by aphasia severity (mild/moderate [mild/mod], moderate to severe [mod/sev], severe [sev]), with severity determined by the WAB–Revised Aphasia Quotient (WAB-R AQ; Kertesz, 2006). Participants in the mild/mod block (WAB AQ > 75) scored in the range of 2–4 on the Boston Diagnostic Aphasia Examination (BDAE) Aphasia Severity scale (Goodglass, Kaplan, & Barresi, 2001) with mild/moderate word retrieval impairments as demonstrated by their mean pretreatment object-naming scores (group $M = 60\%$ correct; $SD = 0.15\%$); those in the mod/sev block ($50 < \text{WAB AQ} < 75$) scored in the range of 1.5–2 on the BDAE Aphasia Severity scale with moderate to severe word retrieval impairments ($M = 44\%$; $SD = 0.22\%$); and participants in the severe block were in the range of 1 to 1.5 on the BDAE Aphasia Severity scale with severe word retrieval impairments ($M = 10\%$; $SD = 0.04\%$).

Participants in a block were randomly assigned to either of two treatment groups: ILAT (Pulvermüller & Berthier, 2008) or PACE (Davis & Wilcox, 1985). Participants assigned

to the ILAT group were explicitly guided toward the speech modality in accordance with treatment procedures outlined in the following paragraphs; participants assigned to the PACE group were exposed to alternate modalities to speech via explicit modeling and were free to use any modality, provided they successfully communicated a message. Apart from this difference in focusing or not focusing on the speech modality, all other aspects of the treatment, as described in detail below, were identical between groups. Treatment was administered to dyads (A, B, C, and so on; see Table 1) formed by pairs of participants assigned to the same treatment group. A total of 10 consecutive workdays of treatment consisting of 3 hr/day (morning or afternoon) was scheduled for each participant. One ILAT and four PACE participants each missed one treatment day, and one PACE participant missed three treatment days due to illness. All other participants completed all 10 treatment sessions. There were two exceptions to the customary assignment of dyads: P23 and P24 both had scheduled dyadic partners drop out of the study just prior to the start of treatment. P23 was treated in a dyad with a volunteer with aphasia (volunteer's results were not examined and are not reported here). P24 was treated individually.

Participants were enrolled in the study sequentially. The randomization automatically alternated assignment of participants from each block to the next treatment group, which alternated between PACE and ILAT. One participant (P7, mod/sev) was unable to arrange transportation for the next assigned PACE treatment group, which was to meet in the morning. Because she was able to attend therapy in the afternoon, when an ILAT treatment group would meet, this resulted in the one exception to pure random assignment using blocking by severity. Given the sequential presentation of participants in the mod/severe and sev blocks, there was also some imbalance in treatment assignment by severity (Table 1), with five versus four participants with moderate to severe aphasia and two versus three participants with severe aphasia randomly assigned to the ILAT and PACE groups, respectively.

There were no statistically significant differences between the two treatment groups on age, gender, stroke etiology, time post onset, aphasia severity, or education (Table 2). However, we noted that there was considerable heterogeneity between study participants, as is typical for persons with aphasia.

Procedure

Testing

Pre- and posttreatment testing included the following standardized tests of aphasia: the Boston Diagnostic Aphasia Examination–Third Edition (BDAE-3; Goodglass et al., 2001) including the Boston Naming Test–Second Edition (BNT-2; Kaplan et al., 2001) and the Cookie Theft Picture Description; and the Porch Index of Communicative Ability (PICA; Porch, 1981). Changes in content units (CUs; Yorkston & Beukelman, 1980) were analyzed for the Cookie Theft Picture Description. In addition to standardized tests

Table 2. Participant characteristics by treatment group.

Parameter	ILAT (constrained), <i>n</i> = 12	PACE (unconstrained), <i>n</i> = 12
Age (years) ^a		
<i>M</i>	68.6	65.0
Range	55–78	47–81
Gender ^b		
Female (<i>n</i>)	4	5
Male (<i>n</i>)	8	7
Etiology ^b		
Ischemic (<i>n</i>)	9	9
Hemorrhagic (<i>n</i>)	3	3
Time postonset ^a		
<i>M</i>	35.6	17.8
Range	6–142	7–82
Severity (BDAE) ^a		
<i>M</i>	2.00	1.96
<i>SD</i>	0.83	0.86
Handedness ^b		
Right (<i>n</i>)	12	10
Left (<i>n</i>)	0	2
Education ^a		
<i>M</i>	15.2	14.6
Range	11–20	5–20

Note. ILAT = intensive language action therapy; PACE = promoting aphasic communicative effectiveness; BDAE = Boston Diagnostic Aphasia Examination.

^aTwo-sample *t* test. ^bChi-square test.

of aphasia, participants were tested on three occasions in order to select sets of stimuli to be trained (or not). The results of this testing also serve as the pretreatment measure for the primary outcome measure, confrontation naming of objects and actions. Three sets of 40 black-and-white line drawings (half objects, half actions) were selected from a subset (*n* = 218 objects) of the Snodgrass and Vanderwart (1980) normed set of objects and from An Object and Action Naming Battery (Masterson & Druks, 1998; *n* = 100 actions). One set of 20 objects and 20 actions was chosen from the pictures that were consistently correctly named (CORR), if possible, on three of three pretreatment tests. The purpose of the CORR set of pictures was to have a set that each participant could successfully name during a pretreatment functional neuroimaging protocol (not reported here). It also served as a control set, given that the CORR pictures were probed but not treated. In the case of the more participants with more severe aphasia and/or verbal apraxia (P5, P6, P17, P18, P20, and P22), it was necessary to include a smaller set of CORR pictures that was repeated in the probes, pictures with responses that were approximately correct (e.g., “bush” for “push”), or pictures that were correct on two of three pretreatment tests, including the last test. The other two sets of 20 objects and 20 actions were chosen on the basis of poor performance during pretreatment testing, and, if possible, were chosen from among pictures that were not named correctly (accuracy = 0/3). For seven of 24 participants (P3, P4, P9, P15, P19, P21, P24), it was necessary to include some pictures that were named on one

of three (but not the last) pretreatment tests, in order to establish a large enough pool of overlapping pictures that both participants in a dyad missed on most, if not all, pretreatment tests. In this way, the dyad could share one set of to-be-trained pictures.

Of the 40 objects and 40 actions, two sets of 20 each were then matched on a number of psycholinguistic variables known to affect word retrieval, including word and syllable length, word frequency, familiarity, age of acquisition, and visual complexity. Matching of sets was initially accomplished manually, but in the second year of the study, a simulation program based on random sampling was implemented that balanced the average score of picture attributes between sets (Kurland et al., XXXX). One set of 20 objects and 20 actions was randomly assigned for treatment (TR) and the other, matched set, was untreated (UNTR).

Each participant's set of 120 pictures (40 CORR, 40 TR, and 40 UNTR) was probed daily at the beginning of each treatment session and on three occasions within the first month posttreatment (see Table 3 for timeline). Confrontation naming probes were presented via computer presentation software (PowerPoint, Microsoft Office Tools) with an intertrial interval of 3 s. Pictures were alternated by treatment condition (e.g., CORR, TR, UNTR, CORR, TR, UNTR) and presented in blocks of, for example, 15 objects, 15 actions, 15 objects. All probes were videotaped for reliability rating. The probes were scored online by an experienced speech-language pathologist and also independently by a trained second judge to assess reliability. Any discrepancies between the two judges' ratings for accuracy were resolved by reviewing the video recordings and achieving consensus. Due to equipment failure, eight probes were not successfully videotaped. Interrater reliability on 94.5% of the data revealed what Landis and Koch identify as "almost perfect" agreement at 98.66% (.81–1.0, Landis & Koch, 1977).

Treatment

In both ILAT and PACE groups, 40 TR pictures (half objects, half actions) were trained in the context of therapeutic language action games (LAGs; Difrancesco, Pulvermüller, & Mohr, 2012). LAGs included the traditional matching card game (i.e., "Go Fish"; Pulvermüller et al., 2001), a matching game in which players have to remember locations of picture and/or word cards (i.e., "Concentration"), a barrier game in which players try to get each other to match the order of cards on their card holders on the basis of referring as a collaborative process (Clarke & Wilkes-Gibbs, 1986), various versions of a guessing game in which players give hints or guess to reveal their own or others' cards, and games in which one or more cards are used to prompt stories, again requiring many hints and guesses and usually coconstruction of the message.

Two certified speech-language pathology clinicians facilitated massed practice retrieval of the target words or concepts during all of the games. They took turns assisting each of the partners in a dyad or facilitated communication

between both partners, depending on the LAG. Accurate responses—spoken or otherwise—were shaped using fading cues over time (Taub et al., 1994) and techniques to reduce errors similar to methods of errorless learning (Fillingham, Sage, & Lambon Ralph, 2006). For example, during the first round of the card-matching game, when ILAT participants typically needed the most support in accurately retrieving the names of pictures during a requesting turn, a clinician would assist a participant to avoid error-filled struggle. The clinician might use associative cuing, offer a cloze sentence, provide an initial phoneme or letter, show the word card, or even model the word for repetition for some participants with severe aphasia. The level of cuing was individually geared toward participants to provide just enough support to help them be successful. These supports were faded as participants gradually improved their independent productions of the words. Players in the ILAT group were also encouraged to retrieve the words in accordance with the methods of shaping, reinforcement, and increasing complexity as described by Pulvermüller and colleagues (2001, 2008). For participants with severe aphasia, including comprehension deficits, each game was first modeled by the two clinicians requesting and responding to requests to demonstrate the game rules and the interactive nature of the games. For all ILAT participants, the clinicians modeled different possibilities for responding, depending on the evolving ability of each player (e.g., "No, no fishing"; "No, I don't have fishing"; "No, I'm sorry, Jane, I don't have the card with the gentleman who is fishing"). The use of varied utterance types, including politeness formulations, and specific grammatical constructions for requesting, proposing, accepting, denying, and so on were continuously modeled to encourage ILAT participants to attempt different utterance forms (e.g., "Do you have . . . ?"; "Can I have . . . ?"; "I would like . . ."). Use of increasing complexity was also modeled—that is, as players became competent at making simple requests, the level of difficulty increased (e.g., "Do you have the apple?"; "John, can I please have the apple?"; "John, can you please give me the three red apples?").

PACE pictures were trained in a similar manner, except that participants' requests and responses were not constrained to (i.e., focused on) the speech modality. Instead, the clinicians modeled and encouraged the most effective modality using techniques described in detail by Carlomagno (1994). For participants with severe aphasia and/or AOS, this might include modeling alternative methods of communication, such as pointing and gesturing a request, writing, drawing, mimicking, or making sounds related to an object or action, without any spoken language. For example, in one exchange, a clinician requested if P20 had the "chisel" by pointing to him and then gesturing carving wood with a chisel, making "ch, ch, ch" sounds with every hammer-on-chisel gesture. The PWA responded in kind, reproducing the gesture while shaking his head no. In the role of "listener," the clinicians' responses emphasized the effectiveness of alternative modalities. As Carlomagno notes, it is especially beneficial to adopt strategies spontaneously produced by PWAs. For example, in one exchange, P20 requested the

Table 3. Individual pre- to posttreatment (Tx) scores on training (TR) pictures ($n = 40$).

PWA	Tx	Pre-Tx TR targets			Days between Pre-Tx probes		Daily probes during treatment period										Post-Tx TR targets			Days between Post-Tx probes			Tx effect size ^a
		T1	T2	T3	T1 to T2	T2 to T3	T4	T5	T6	T7	T8	T9	T10	T11	T12	T13	T14	T15	T16	T13 to T14	T14 to T15	T15 to T16	
P1	ILAT	4	5	0	6	6	14	16	22	25	29	28	27	30	32	30	25	28	26	10	6	7	8.82
P2	ILAT	0	0	1	6	12	0	2	3	4	5	3	11	2	8	4	6	9	14	10	12	9	16.17
P3	ILAT	19	9	12	6	6	n/a	25	27	28	33	33	33	35	38	34	32	32	32	6	2	7	3.64
P4	ILAT	14	18	16	6	6	21	30	28	32	30	31	34	33	33	30	27	28	30	11	7	6	6.17
P5	ILAT	0	0	0	13	6	1	5	7	9	8	5	13	12	8	8	12	7	12	4	1	10	20.17
P6	ILAT	0	0	0	13	13	2	6	5	2	4	1	1	5	5	3	4	2	5	3	1	4	3.17
P7	ILAT	0	0	0	6	19	n/a	13	17	19	21	21	22	26	27	27	28	30	25	3	2	12	3.76
P8	ILAT	0	3	0	13	13	n/a	12	14	19	24	23	20	21	24	24	25	21	22	3	1	1	12.51
P9	ILAT	32	33	28	4	6	27	25	24	20	28	32	33	31	38	34	32	31	34	13	4	1	0.50
P10	ILAT	3	7	3	34	34	13	20	19	24	21	26	29	32	31	34	30	35	31	11	1	5	11.98
P11	ILAT	4	4	3	21	1	9	13	15	18	14	18	17	21	22	21	25	18	17	4	5	9	28.29
P12	ILAT	3	5	8	13	1	17	20	21	28	23	25	32	NA	31	28	29	23	28	10	1	14	8.48
P13	PACE	5	5	0	5	6	13	18	22	21	21	23	25	33	28	NA	29	27	27	3	6	6	8.43
P14	PACE	0	0	0	13	6	4	7	12	10	NA	12	15	15	10	15	14	15	17	4	12	6	7.17
P15	PACE	17	10	9	6	9	11	13	26	30	28	30	28	NA	32	33	33	35	30	7	16	16	4.74
P16	PACE	0	4	2	6	9	9	17	15	28	29	27	28	31	29	32	31	29	29	7	16	16	13.83
P17	PACE	0	0	0	13	22	0	4	4	10	10	7	5	6	9	11	9	7	NA	2	6	NA	3.50
P18	PACE	0	0	0	13	13	0	0	3	3	5	5	5	3	2	6	2	6	3	4	1	4	1.94
P19	PACE	14	13	16	25	20	27	35	36	36	37	37	36	36	34	34	36	36	33	4	5	1	13.53
P20	PACE	2	2	2	20	41	1	1	10	6	7	9	10	9	8	13	8	6	7	4	14	6	10.50
P21	PACE	7	6	3	6	1	12	23	22	28	29	21	25	23	24	32	22	28	25	13	1	1	9.45
P22	PACE	7	4	5	6	6	4	4	5	10	7	NA	12	13	11	18	15	15	15	10	1	1	6.33
P23	PACE	6	3	10	2	20	13	15	12	15	13	18	22	20	20	14	19	17	25	12	6	1	3.99
P24	PACE	12	12	12	32	13	18	NA	13	17	NA	24	24	21	25	NA	22	23	25	5	4	1	3.28

Note. Days between pre- and posttreatment probes are shown in bold. PWA = participant with aphasia; T = time; ILAT = intensive language action therapy; n/a = data not available due to technical issues; PACE = promoting aphasic communicative effectiveness; NA = data not available due to participant missing session.

^aTx effect size = Busk & Serlin's d score, where $d = (\text{Mean} [\text{post-Tx}] - \text{Mean} [\text{pre-Tx}]) / SD [\text{pre-Tx}]$; in cases where there was no variability in pre-Tx scores (italicized), the first available probe score (e.g., T4) was added for calculating M and SD pre-Tx.

“marching” card from P19 by making an upper body (arms swinging) gesture. P19 did not understand the gesture, and responded, “pushing?” to which P20 responded no and gestured holding a gun. When P19 did not reply, a clinician suggested that P20 demonstrate how the legs also moved, after which he gestured marching with arms and legs. Then P19 replied, “oh, marching.” Clinicians responded to participants’ attempts to request, no matter how vague, with interpretive guesses, thus providing inferential feedback, as Davis and Wilcox (1985) recommend. Thus, a vague up-and-down arm motion might be responded to with a more explicit gesture of “hammering,” accompanied by a verbal response such as “I think I know what you’re asking for. Do you want the hammer?” In all cases, regardless of modality, feedback was used to provide informative adequacy of participants’ messages.

Analysis

We summarized the study results via estimates of the mean, standard deviation, and 95% confidence interval response (i.e., proportion of named pictures) for each treatment group’s (ILAT or PACE) performance prior to treatment (pre), after treatment (post), and gain (post – pre). This summary includes estimates for TR and UNTR pictures at each point in time. Paired *t* tests are used to compare the response between TR and UNTR pictures at each time point within each treatment. Two-sample *t* tests are used to compare treatments at each time point. Finally, paired *t* tests are used to test the null hypothesis of no gain in response over time for each treatment, with two-sample *t* tests used to compare gain between treatment groups. The *p* values are presented for statistically significant tests (with *p* < .05).

Additional analyses are based on fitting a linear mixed model (Fitzmaurice, Laird, & Ware, 2011) using SAS (Littell, Milliken, Stroup, Wolfinger, & Schabenberger, 2006) for each type of target (TR, UNTR) pre- and post-treatment for each intervention group (ILAT and PACE). Participants are considered to be random effects. The mixed model properly accounts for repeated measures on a participant, including the proportion of named TR and UNTR pictures at each time. Fixed effects in the mixed model correspond to the expected response under each treatment at each time, as illustrated in Figures 1 and 2.

At pretest, μ represents the expected number of named pictures for study participants. In practice, there may be a small difference in picture naming between the TR and UNTR pictures due to an imperfect balance of the random assignment of pictures to the groups. Such an effect is included in the model but for simplicity is not illustrated in Figure 1. The change in the expected response that occurs over time in naming UNTR pictures is represented by δ , the period effect. This effect may be due to natural improvements over time (although all participants had chronic aphasia and were not expected to demonstrate spontaneous recovery), to co-occurring effects due to repeated evaluation of the pictures during the intervention and posttreatment probes, or to diffusion effects to the UNTR pictures

Figure 1. Parameters for participants receiving the intensive language action therapy (ILAT) treatment.

as a result of treatment. The final parameter, γ , represents the additional gain in naming of TR pictures over and above the period effect. Period effects (δ) and treatment effects (γ) could vary between participants. We accounted for such variability by including random effects for each of these terms in the mixed model.

Two additional parameters are added to the model for the PACE treatment (Figure 2). The first parameter, δ^* , is an additional period effect for the PACE treatment. This may occur if diffusion of treatment effects to UNTR pictures occurs to a different extent for ILAT and PACE treatments. Finally, the model for the expected response posttreatment for PACE includes the parameter γ^* , which represents the additional treatment effect for PACE over and above the treatment effect for ILAT. Random effects for δ^* and γ^* were also included in the model to account for variability among participants.

Additional mixed models were fit that include a variable for participant severity (using mild/mod, mod/sev, and sev blocks of participants) so that interactions between

Figure 2. Parameters for participants receiving the promoting aphasic communicative effectiveness (PACE) treatment.

each of the effects μ , δ , δ^* , γ , and γ^* and severity could be examined.

Results

We first conducted an analysis of participants in each treatment group over time, observations of which are noted below in the *t* Tests Results section. Results of the linear mixed model are presented below in the Mixed Model Analyses section. Although the current study is focused on the difference between groups randomly assigned to ILAT versus PACE treatment, individual performance on TR targets over time is presented in Table 3 to demonstrate participant variability as well as the timing of probes. Eleven of 12 ILAT and 10 of 12 PACE participants demonstrated at least a small effect size (4.0), and eight of 12 ILAT and seven of 12 PACE participants showed at least a medium effect size (7.0), according to Robey and Beeson (2005; as cited in Beeson & Robey, 2006) for lexical treatment.

t Test Results

Pretreatment, there were no statistically significant differences between naming of TR and UNTR pictures in either the ILAT ($p = .782$) or PACE ($p = .150$) group. This is to be expected because TR and UNTR pictures were randomly assigned to balance the percentage of named pictures pretreatment. Approximately 16% of the pictures were named pretreatment (see Table 4). Figure 3 illustrates the percentage of named pictures among participants with mild/moderate aphasia ($n = 5$ in ILAT and $n = 5$ in PACE), moderate to severe aphasia ($n = 5$ in ILAT and $n = 4$ in PACE), and severe aphasia ($n = 2$ in ILAT and $n = 3$ in PACE) illustrates that most of the correctly named pictures pretreatment occurred among the participants with mild/moderate or moderate to severe aphasia.

The percentage of named pictures posttreatment, as well as the gain in naming, are summarized in Table 4. The results indicate large gains in naming for TR pictures, with estimates of 40.4% ($p < .0001$) for ILAT and 37.4% ($p < .0001$) for PACE. Differences in the naming gain for TR pictures between treatment groups were not statistically significant ($p = .720$). Significant positive gains were

Figure 3. Mean percentage trained (TR) and untrained (UNTR) targets named pretreatment.

also reported for UNTR pictures, with somewhat larger gains (23.4%) reported for the ILAT group ($p < .0015$) versus 15.2% ($p < .001$) for PACE. Differences in naming gain for UNTR pictures between treatment groups were not statistically significant ($p = .221$).

Both ILAT and PACE were effective in improving participants' performance on the primary outcome measure, confrontation naming of action and object pictures. All participants showed gains in naming TR pictures, and nearly all participants (23/24) improved on naming UNTR pictures. However, as expected after this short period of treatment, improvements in performance on standardized measures of aphasia were small. Nonetheless, most participants demonstrated some gains in performance, and nearly all scores improved for both groups. One exception was in CUs (Yorkston & Beukelman, 1980) produced for the Cookie Theft Picture Description (Goodglass et al., 2001): The ILAT participants produced marginally more CUs, whereas the PACE group produced marginally fewer (Table 5).

Mixed-Model Analyses

The data were analyzed via a mixed model considering participants as random effects, with additional random effects

Table 4. Percentage named targets pre- and posttreatment (Tx) and gain (Post – Pre) by treatment condition.

Tx arm	Pre-Tx		Post-Tx		Post – Pre ^a			
	M	SD	M	SD	M	SE	95% CI	p value
ILAT ($n = 12$)								
Primary outcome (TR pictures)	16.3	23.3	56.7	24.1	40.4	6.29	[26.6, 54.3]	<.0001
Period effect (UNTR pictures) ^b	16.5	23.7	39.9	26.2	23.4	5.59	[11.1, 35.7]	.0015
PACE ($n = 12$)								
Primary outcome (TR pictures)	13.1	12.7	50.4	26.4	37.4	5.58	[25.1, 49.7]	<.0001
Period effect (UNTR pictures) ^b	16.5	15.2	31.6	23.1	15.2	3.4	[7.7, 22.7]	.001

Note. Post-Tx data are shown in bold. SE = standard error; CI = confidence interval; ILAT = intensive language action therapy; TR = trained; UNTR = untrained; PACE = promoting aphasic communicative effectiveness.

^aPaired *t* tests. ^bPeriod effect may represent generalization to UNTR condition.

Table 5. Means (SD) for secondary outcomes measures of speech and language characteristics pre- and post-treatment (Tx).

Measure	ILAT treatment group (<i>n</i> = 12)		PACE treatment group (<i>n</i> = 12)	
	Pre-Tx (SD)	Post-Tx (SD)	Pre-Tx (SD)	Post-Tx (SD)
BDAE aphasia severity (0–6)	2.0 (0.8)	2.3 (0.9)	2.0 (0.9)	2.1 (1.0)
BDAE aud comprehension (%tile)	52.5 (27.5)	54.9 (27.3)	40.8 (25.8)	43.2 (27.5)
BDAE sentence repetition (%tile)	44.6 (18.3)	48.5 (22.7)	41.8 (30.9)	46.9 (28.1)
BNT (% correct)	56.5 (16.1)	58.0 (16.2)	61.6 (16.1)	63.8 (20.0)
Cookie Theft Description–Content Units	8.7 (5.3)	10 (6.5)	8.5 (5.5)	8.25 (4.9)
PICA overall (1–16)	12.3 (1.1)	12.5 (1.3)	12.2 (1.2)	12.6 (1.2)

Note. Post-Tx data are shown in bold. ILAT = intensive language action therapy; PACE = promoting aphasic communicative effectiveness; BDAE = Boston Diagnostic Aphasia Examination; aud = auditory; BNT = Boston Naming Test; PICA = Porch Index of Communicative Ability.

for period and treatment, and Participants $\times$ Period $\times$ Treatment interactions. Estimates of the variance in treatment effects between participants were zero. As a result, only random effects associated with participants, periods, and Period $\times$ Treatment were included in subsequent models.

There was a significant period effect, $\hat{\delta} = 22.9\%$ $p < .0001$, and a significant treatment effect, $\hat{\gamma} = 18.1\%$, $p < .0001$, for ILAT, but additional period and treatment effects (i.e., $\hat{\delta}^* = -7.18\%$ and $\hat{\gamma}^* = 3.09\%$, representing a difference between ILAT and PACE treatments) were not significant ($p = .2822$ and $p = .5539$, respectively). For this reason, the model was simplified by excluding these effects.

Additional models were fit to include severity (classified as mild/moderate, moderate to severe, and severe), with the moderate to severe group considered as the reference group in the model. We report results on the basis of the linear mixed model adjusted for severity. All models included random effects for participants, for period effects (for participants), and for treatment effects (for participants). Severity $\times$ Period; Severity $\times$ Period $\times$ Type of Treatment; Severity $\times$ Treatment; and Severity $\times$ Treatment $\times$ Type of Treatment interactions were added as fixed effects to the model. The Severity $\times$ Period $\times$ Type of Treatment ($p = .8114$); Severity $\times$ Treatment ($p = .4252$); and Severity $\times$ Treatment $\times$ Type of Treatment ($p = .5888$) interactions were not statistically significant and were dropped from the model.

We summarized the results in Table 6 and Figures 4 and 5. The percentage of named pictures was higher for participants with mild or moderate aphasia compared with severe aphasia. This difference was nearly statistically significant ($p = .0553$). The difference in the percentage of the named UNTR versus TR pictures pretreatment, though not statistically significant ($p = .0641$), also differed somewhat by severity. Of particular note is the difference in period effect by severity ($p = .0083$), where among UNTR pictures, much larger effects were evident for participants with mild aphasia and moderate to severe aphasia. Although a slightly lower percentage of correctly named pictures was reported for participants in the PACE group, this difference (estimated as -6.12% ; standard error [SE] = 5.61%) was not statistically significant ($p = .2763$). Finally, we note that significant gains were made for all participants among TR pictures, as noted in the treatment effect in Figure 5, whereas

the difference between ILAT and PACE treatments (2.92% [$SE = 5.96\%$] $p = .5557$) was not statistically significant.¹

Discussion

ILAT (Berthier & Pulvermüller, 2011; Pulvermüller & Berthier, 2008) is an effective aphasia treatment program founded on three basic principles of neuroscience: (a) guidance by constraint; (b) communication embedding; and (c) massed practice. Although the authors assert that all three principles are fundamental to overcoming learned nonuse in chronic aphasia, this has not been empirically established. We examined whether one of the three principles of ILAT, guidance by constraint, is critical to its success. Twenty-four PWAs were randomly assigned to ILAT or a modified version of PACE. Both groups received massed practice in the context of therapeutic LAGs in which picture cards representing trained sets of actions and objects were used to stimulate communicative actions such as requesting, accepting, and denying. Whereas the ILAT group was guided toward spoken responses, even when gestures sufficed to communicate a message, the PACE group was guided toward multiple response modalities and participants were free to choose any one or more.

All participants, even those with moderate to severe aphasia and comorbid AOS, improved on the primary outcome measure, picture naming. Contrary to our hypothesis that participants who were randomly assigned to ILAT would demonstrate stronger treatment effects than PACE participants, there were no significant differences between the two groups on confrontation naming of TR and UNTR objects and actions. Improvements on standardized measures of aphasia and changes in discourse over the short 2-week treatment period were not robust. Nonetheless, most participants demonstrated some gains in performance, and nearly all scores improved for both groups. However, like

¹In an abundance of caution and at the request of a reviewer, we reran the mixed model after excluding P23 and P24, who may not have been exposed to precisely the same treatment intensity as other participants in the typical dyadic treatment protocol. The results were similar to those reported herein.

Table 6. Model estimates (standard error [SE]) of pretreatment % and gain in % of named targets overall and by severity.

Parameter	Overall		Mild		Mod/Sev		Severe	
	Estimates	SE	Estimates	SE	Estimates	SE	Estimates	SE
Pretreatment								
UNTR								
ILAT	16.5	(7.2)	24.8	(5.6)	15.4	(5.9)	1.6	(7.9)
PACE	16.5	(5.6)	24.8	(5.6)	15.4	(5.9)	1.6	(7.9)
TR								
ILAT	14.6	(3.9)	20.3	(5.6)	16.3	(5.9)	0.3	(8.0)
PACE	14.6	(3.9)	20.3	(5.6)	16.3	(5.9)	0.3	(8.0)
TR/UNTR period								
ILAT	22.9	(4.7)	31.4	(5.2)	20.0	(5.2)	8.3	(7.0)
PACE	15.7	(4.7)	25.3	(5.2)	13.9	(5.5)	2.2	(6.6)
Gain, TR training								
ILAT	18.1	(3.8)	18.2	(3.6)	18.2	(3.6)	18.2	(3.6)
PACE	21.2	(3.8)	21.1	(3.6)	21.1	(3.6)	21.1	(3.6)

Note. Mod/Sev = moderate to severe; UNTR = untreated; ILAT = intensive language action therapy; PACE = promoting aphasic communicative effectiveness; TR = trained.

the primary outcome measure, there were no differences between the two treatment groups.

It is intuitively appealing to expect that participants who effectively had more opportunities to practice speaking should perform better than participants with fewer opportunities. According to Hebbian theory, the more often neuronal circuits are activated synchronously, the more their connections ought to be strengthened. Thus, as Berthier and Pulvermüller (2011) suggest, ILAT should lead to synaptic strengthening and hence, better outcomes than therapy in which spoken responses are only one choice of many. However, our results seem to suggest that ILAT's guidance by constraint added little beyond communication embedding and massed practice, both of which were infused in both treatments.

One caveat, however, is that most PACE participants often chose to attempt spoken responses. This is important because, in spite of clinician modeling of alternative communicative behaviors that are assumed to be most effective for each participant during PACE (including gesture,

onomatopoeia, pantomime, facial expressions, drawings, written words), ultimately, the participant has freedom to choose a preferred communicative modality. Moreover, PACE principles suggest that clinicians should honor whichever communicative modality a participant chooses. For the more severe participants whose aphasia and/or AOS initially precluded effective spoken responses, the lack of constraint to the speech modality in the PACE group was helpful in supporting communicative success early in therapy. Even in these cases, however, feedback that is based on communicative adequacy often involves the spoken target word. Thus, even the PACE group, although not explicitly guided by constraint to speech, nonetheless had plenty of opportunity to hear and repeat picture names. Moreover, many of the PACE participants actively chose speech, a problem that has surfaced in some participants in previous comparisons between CILT and PACE (Kurland et al., 2010; Maher et al., 2006).

Another factor that may have influenced the not statistically significant differences between treatment groups

Figure 4. Model parameter estimates for pretreatment and period effects overall and by aphasia severity for untrained (UNTR) pictures.

Figure 5. Model parameter estimates for pretreatment, period, and treatment effects overall and by aphasia severity for trained (TR) pictures.

may be related to the fact that both groups were probed daily on the CORR, TR, and UNTR picture sets just prior to treatment. During probes, all participants were expected to attempt to name the pictures (i.e., not to communicate by any means what pictures they were viewing). In other words, all participants were aware of a daily probe testing their naming ability and seemed to be motivated to demonstrate improvement over their prior day's performance. Had the probes been more reflective of the differences in treatment with respect to focusing on speech, it is possible that PACE participants would not have been as self-motivated to practice the names of the pictures during treatment. Of course, in such a scenario, the comparison between treatment groups would not have been possible. Finally, the lack of a statistically significant difference in naming between ILAT and PACE is in part a function of the number of participants in the study and the study's power. This latter problem was also observed by Maher et al. (2006), who piloted the first exploratory treatment study comparing CILT ($n = 4$) and PACE ($n = 5$).

Our results are similar to those of Maher et al. (2006). Not only did one of their PACE participants refuse any modality other than spoken responses, blurring the distinction between PACE and CILT, they too noted improvements in both groups, more so for CILT than PACE, even for one CILT participant with severe AOS. Whereas their results also did not reach statistical significance for a difference between groups (albeit small groups), they noted that the data suggested some additional benefit conferred by the CILT approach.

A reasonable alternative hypothesis to explain this apparent additional benefit might stem from a higher dosage of treatment in the CILT group. In the current study, we attempted to control for treatment intensity (i.e., massed practice was defined as 3 hr/day, 5 days/week for 2 weeks for both the CILT and PACE groups). However, a recent line of inquiry into treatment intensity suggests that not all so-called intensive treatments are alike and that the relationship between intensity and outcome is not always linear. Baker (2012a), although acknowledging that the establishment of optimal intensity for a given intervention is no trivial matter, suggested that clinical researchers report dosage parameters as conceptualized by Warren, Fey, and Yoder (2012). These parameters include: *dose* (number or rate of correctly administered teaching episodes per session), *dose form* (typical task or activity during which the dose is administered), *dose frequency* (frequency of administering teaching episodes per unit of time), *total intervention duration* (period of time during which the treatment is provided), and *cumulative intervention intensity* (Dose $\times$ Dose Frequency $\times$ Intervention Duration). Baker's expansion on the framework of Warren et al. includes identification of the "active ingredients" constituting the teaching episodes (e.g., the timing, modeling, shaping, direct instructions).

The goal of identifying these dose parameters has merit, and in theory, might be able to tease apart differences between two intensive treatment protocols that may have delivered different doses and/or different active ingredients

to participants. The problem, however, is that precisely identifying the dosage parameters can be challenging and complicated (Baker, 2012b; Kamhi, 2012). It has obvious application to computer-delivered aphasia treatment wherein the dose and dose form are fairly well prescribed (e.g., Cherney, 2012; Harnish et al., 2014; Off, Griffin, Spencer, & Rogers, 2015). For example, Harnish et al. (2014) easily characterized the active ingredients (e.g., semantic, phonemic, and orthographic cues) and dose (50 pictures $\times$ 8 cues or presentations = 400 teaching episodes) within their dose form (computer-assisted picture naming). In such an environment, the dosing parameters are consistent across participants, and it makes absolute sense to use this framework for reporting dosage. Moreover, computer-based treatment has the added advantage of automatically keeping track of time spent on various active ingredients (e.g., Cherney, 2012).

As Harnish et al. (2014) noted, however, "some treatments will lend themselves better to systematically evaluating dosage than others" (p. S297). Given the highly interactive and game-like environment of the therapeutic LAGs in which treatment was delivered by two clinicians to two PWAs during ILAT and PACE in the current study, it would have been impractical, if not impossible, to record precise dose parameters. Moreover, it would have detracted from the communicative embedding that is essential to both treatments. In retrospect, it is unlikely that dose and dose frequency were consistent across groups, although the therapeutic intensity ratio was designed to be the same for both: 0.375 (15 hr/40-hr work week). Some of the individual differences in dose stem from the fact that not all participants attended all 10 sessions (see Table 3). Perhaps more germane to the question of differences in dose between the ILAT and PACE groups is that with the currently available data, we cannot quantify precisely how many more spoken requests or cues the participants in ILAT were exposed to compared with those in PACE, or how many more times the ILAT participants provided spoken responses or initiated spoken requests compared with the PACE participants. Future investigations into this question might benefit from videotaping all treatment sessions and attempting to rate them along the various dose parameters. Scoring, however, would likely be a challenging task and reliability difficult to obtain, given the complex multivariate nature of the active ingredients inherent in LAGs. As Baker (2012b) noted, reduction of dosage in speech-language pathology interventions to a pharmacological framework may not always be appropriate. Some interventions might not be easily reduced to a summation of teaching episodes.

For the current study, we have to assume that massed practice was controlled—that is, both groups received similar total intervention durations (approximately 30 hr per PWA, not including absences). We also controlled for communication embedding between the two groups but did not find a statistically significant difference between the groups. Given this, we cannot conclude that guidance by constraint either is or is not a critical active ingredient in ILAT. We can only conclude that both treatments led to significant improvements on the primary outcome measure (i.e., picture

naming) and that there were subtle differences that we were unable to conclusively identify using our sample of observations.

Rapid gains in naming and picture description and modest improvements on standardized tests of aphasia over a 2-week period, whether due to ILAT or PACE therapy, may be a promising outcome for individuals whose chronic aphasia interferes with their ability or desire to participate in routine activities associated with a meaningful life. Often, the treatment effects boost participants' confidence, effectively "jump-starting" a system that may have settled into a state of "learned nonuse" (Taub et al., 1994). For durable language improvements, however, it may be critical for the 2-week intensive treatment programs to be followed by other, more long-term opportunities to strengthen what will likely be persistently compromised language networks.

In their evidence-based systematic review of the effects of intensity and constraint on aphasia therapy, Cherney et al. (2010) recommended that research extend beyond the exploratory phase to determine the efficacy of ILAT. The goal of the present study was to determine whether guidance by constraint to speech was a critical aspect of ILAT by comparing two similar intensive treatment programs, one guided by constraint (ILAT) and the other unconstrained (PACE), in a randomized controlled clinical trial. Although not finding a statistically significant difference between the two groups, the results do indicate that larger gains in naming of UNTR pictures occurred for ILAT relative to PACE, whereas slightly better naming occurred for TR pictures with PACE. The advantage for ILAT on UNTR pictures suggests better generalizability of treatment, which may be the elusive additional benefit conferred by the ILAT approach.

The current study was conducted with 24 heterogeneous participants with aphasia. Although participants were randomly assigned to treatment, by chance the PACE group included a greater proportion of participants whose demographic and/or clinical characteristics may have unduly handicapped the outcomes of that group. For example, severe aphasia, severe AOS, and low education occurred more frequently among PACE participants. It could be argued, however, that the ILAT group had a disadvantage in the frequency of older participants with longer time post-stroke. Such lack of absolute balance in prognostic attributes is unfortunately an unavoidable consequence of randomization in small clinical trials, especially when there are so many endogenous variables (e.g., site and size of lesion, aphasia classification and severity, time postonset, presence and severity of AOS, education) that can potentially affect treatment outcomes. One plausible solution is a larger controlled clinical trial, such as described by Robey and Schultz (1998), a solution that might also have a better chance of conclusively answering the question of the role of guidance by constraint. It is important to note that for the current study, there were no statistically significant differences between the groups along these demographic and clinical characteristics.

The current study, although not demonstrating a statistically significant difference between the groups on the basis of guidance by constraint, nonetheless provided sufficient data to model effects of time, of general treatment, of particular treatments, and of the interactions between severity and these other variables. One surprising result was the strength of the period effect in the mixed model—or what is commonly referred to as "generalization to untrained exemplars" in treatment studies that include an untrained control set (e.g., Boyle, 2004; Kiran, 2007, 2008). This effect, strongest in persons with mild and moderate aphasia and weakest among those with severe aphasia, could be due to a naturally occurring improvement in naming ability over time, although this outcome is unlikely given the chronic nature of participants' aphasia. Another interpretation is that the improvement is due in part to repeated exposure following pretreatment testing (daily probe testing during the 2-week treatment period and in post-treatment testing). Off et al. (2015) recently demonstrated that repeated practice, without feedback, improved naming in participants with a variety of aphasia severity levels.

A third, traditional interpretation suggests that something about the treatment leads participants to generalize the treatment to UNTR pictures. If the relatively large period effect is a result of the first two interpretations, we would expect similar effects between treatment groups. The larger (but not statistically significant) period effect for ILAT—even among the PWAs with severe aphasia—suggests that participants in that group were better able to generalize the treatment to UNTR exemplars. Here again, we cautiously concur with Maher et al. (2006) that there is "some aspect of the [CILT] approach that confers additional benefit" (p. 850). The distinction between these interpretations and the influence of severity on both the period and treatment effects may be an important area for future research.

Acknowledgment

This research was supported by funding from National Institute on Deafness and Other Communication Disorders (NIDCD) Grant R01DC011526 awarded to Jacquie Kurland. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

- Aten, J. L., Caligiuri, M. P., & Holland, A. L. (1982). The efficacy of functional communication therapy for chronic aphasic patients. *Journal of Speech and Hearing Disorders*, 47, 93–96.
- Baker, E. (2012a). Optimal intervention intensity. *International Journal of Speech-Language Pathology*, 14, 401–409.
- Baker, E. (2012b). Optimal intervention intensity in speech-language pathology: Discoveries, challenges, and uncharted territories. *International Journal of Speech-Language Pathology*, 14, 478–485.
- Barthel, G., Meinzer, M., Djundja, D., & Rockstroh, B. (2008). Intensive language therapy in chronic aphasia: Which aspects contribute most? *Aphasiology*, 22, 408–421.

- Basso, A., & Caporali, A. (2001). Aphasia therapy or the importance of being earnest. *Aphasiology*, 15, 307–332.
- Beeson, P. M., & Robey, R. R. (2006). Evaluating single-subject treatment research: Lessons learned from the aphasia literature. *Neuropsychology Review*, 16, 161–169.
- Berthier, M. L., & Pulvermüller, F. (2011). Neuroscience insights improve neurorehabilitation of poststroke aphasia. *Nature Reviews Neurology*, 7, 86–97.
- Bhogal, S. K., Teasell, R., & Speechley, M. (2003). Intensity of aphasia therapy, impact on recovery. *Stroke*, 34, 987–993.
- Boyle, M. (2004). Semantic feature analysis treatment for anomia in two fluent aphasia syndromes. *American Journal of Speech-Language Pathology*, 13, 236–249.
- Breier, J. I., Maher, L. M., Novak, B., & Papanicolaou, A. C. (2006). Functional imaging before and after constraint-induced language therapy for aphasia using MEG. *Neurocase*, 12, 322–331.
- Carlomagno, S. (1994). *Pragmatic approaches to aphasia therapy*. London, United Kingdom: Whurr.
- Cherney, L. R. (2012). Aphasia treatment: Intensity, dose parameters, and script training. *International Journal of Speech-Language Pathology*, 14, 424–431.
- Cherney, L. R., Patterson, J. P., & Raymer, A. M. (2011). Intensity of aphasia therapy: Evidence and efficacy. *Current Neurology and Neuroscience Reports*, 11, 560–569.
- Cherney, L. R., Patterson, J. P., Raymer, A., Frymark, T., & Schooling, T. (2008). Evidence-based systematic review: Effects of intensity of treatment and constraint-induced language therapy for individuals with stroke-induced aphasia. *Journal of Speech, Language, and Hearing Research*, 51, 1282–1299.
- Cherney, L. R., Patterson, J. P., Raymer, A. M., Frymark, T., & Schooling, T. (2010, October). *Updated evidence-based systematic review: Effects of intensity of treatment and constraint-induced language therapy for individuals with stroke-induced aphasia*. Rockville, MD: American Speech-Language-Hearing Association.
- Clarke, H. H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition*, 22, 1–39.
- Dabul, B. L. (2000). *Apraxia Battery for Adults—Second Edition* (ABA-2). Austin, TX: Pro-Ed.
- Darley, F. L., Aronson, A. E., & Brown, J. R. (1975). *Motor speech disorders* (3rd ed.). Philadelphia, PA: Saunders.
- Davis, A., & Wilcox, J. (1985). *Adult aphasia rehabilitation: Applied pragmatics*. San Diego, CA: Singular.
- Difrancesco, S., Pulvermüller, F., & Mohr, B. (2012). Intensive language action therapy: The methods. *Aphasiology*, 26, 1317–1351.
- Duffy, J. R. (2012). *Motor speech disorders: Substrates, differential diagnosis, and management* (3rd ed.). St. Louis, MO: Elsevier Mosby.
- Ellis, C., Dismuke, C., & Edwards, K. K. (2010). Longitudinal trends in aphasia in the United States. *NeuroRehabilitation*, 27, 327–333.
- Fillingham, J., Sage, K., & Lambon Ralph, M. A. (2006). The treatment of anomia using errorless learning. *Neuropsychological Rehabilitation*, 16, 129–154.
- Fitzmaurice, G. M., Laird, N. M., & Ware, J. H. (2011). *Applied longitudinal analysis* (2nd ed.). Hoboken, NJ: Wiley.
- Go, A. S., Mozzafarian, D., Roger, V. L., Benjamin, E. J., Berry, J. D., Blaha, M. J., . . . Turner, M. B. (2014). Heart disease and stroke statistics—2014 update: A report from the American Heart Association. *Circulation*. <http://dx.doi.org/10.1161/01.cir.0000441139.02102.80>
- Goodglass, H., Kaplan, E., & Barresi, B. (2001). *The assessment of aphasia and related disorders* (3rd ed.). Austin, TX: Pro-Ed.
- Harnish, S. M., Morgan, J., Lundine, J. P., Bauer, A., Singletary, F., Benjamin, M. L., . . . Crosson, B. (2014). Dosing of a cued picture-naming treatment for anomia. *American Journal of Speech-Language Pathology*, 23, S285–S299.
- Huber, W., Poeck, K., Weniger, D., & Willmes, K. (1983). *Aachener Aphasia Test* (AAT). Göttingen, Germany: Hogrefe.
- Kamhi, A. G. (2012). Pharmacological dosage concepts: How useful are they for educators and speech-language pathologists? *International Journal of Speech-Language Pathology*, 14, 447–450.
- Kaplan, E., Goodglass, H., & Weintraub, S. (2001). *Boston Naming Test—Second Edition*. Philadelphia, PA: Lea & Febiger.
- Kempster, G. B., Gerratt, B. R., Verdolini Abbott, K., Barkmeier-Kraemer, J., & Hillman, R. (2009). Consensus auditory-perceptual evaluation of voice: Development of a standardized clinical protocol. *American Journal of Speech-Language Pathology*, 18, 124–132.
- Kertesz, A. (1982). *Western Aphasia Battery*. San Antonio, TX: The Psychological Corporation.
- Kertesz, A. (2006). *Western Aphasia Battery—Revised*. San Antonio, TX: Pearson.
- Kiran, S. (2007). Semantic complexity in the treatment of naming deficits. *American Journal of Speech-Language Pathology*, 16, 18–29.
- Kiran, S. (2008). Typicality of inanimate category exemplars in aphasia treatment: Further evidence for semantic complexity. *Journal of Speech, Language, and Hearing Research*, 51, 1550–1568.
- Kurland, J., Baldwin, K., & Tauer, C. (2010). Treatment-induced neuroplasticity following intensive naming therapy in a case of chronic Wernicke's aphasia. *Aphasiology*, 24, 737–751.
- Kurland, J., Pulvermüller, F., Silva, N., Burke, K., & Andrianopoulos, M. (2012). Constrained vs. unconstrained intensive language therapy in two individuals with chronic, moderate-to-severe aphasia and apraxia of speech: Behavioral and fMRI outcomes. *American Journal of Speech-Language Pathology*, 21, S65–S87.
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33, 159–174.
- Littell, R. C., Milliken, G. A., Stroup, W. W., Wolfinger, R. D., & Schabenberger, O. (2006). *SAS for mixed models* (2nd ed.). Cary, NC: SAS Institute.
- Maher, L. M., Kendall, D., Swarengin, J. A., Rodriguez, A., Leon, S. A., Pingel, K., . . . Rothi, L. J. (2006). A pilot study of use-dependent learning in the context of constraint induced language therapy. *Journal of the International Neuropsychological Society*, 12, 843–852.
- Masterson, J., & Druks, J. (1998). Description of a set of 164 nouns and 102 verbs matched for printed word frequency, familiarity and age of acquisition. *Journal of Neurolinguistics*, 11, 331–354.
- McNeil, M. R., Robin, D. A., & Schmidt, R. A. (2009). Apraxia of speech: Definition and differential diagnosis. In M. R. McNeil (Ed.), *Clinical management of sensorimotor speech disorders* (2nd ed.). New York, NY: Thieme.
- Meinzer, M., Djundja, D., Barthel, G., Elbert, T., & Rockstroh, B. (2005). Long-term stability of improved language functions in chronic aphasia after constraint-induced aphasia therapy. *Stroke*, 36, 1462–1466.
- Meinzer, M., Elbert, T., Djundja, D., Taub, E., & Rockstroh, B. (2007). Extending the constraint-induced movement therapy (CIMT) approach to cognitive functions: Constraint-induced aphasia therapy (CIAT) of chronic aphasia. *NeuroRehabilitation*, 22, 311–318.

- Meinzer, M., Flaisch, T., Breitenstein, C., Wienbruch, C., Elbert, T., & Rockstroh, B. (2008). Functional re-recruitment of dysfunctional brain areas predicts language recovery in chronic aphasia. *NeuroImage*, 39, 2038–2046.
- Meinzer, M., Flaisch, T., Obleser, J., Assadollahi, R., Djundja, D., Barthel, G., & Rockstroh, B. (2006). Brain regions essential for improved lexical access in an aged aphasic patient: A case report. *BMC Neurology*, 6(28), 1–10.
- Meinzer, M., Rodriguez, A. D., & Gonzalez-Roth, L. J. (2012). First decade of research on constrained-induced treatment approaches for aphasia rehabilitation. *Archives of Physical Medicine and Rehabilitation*, 93, S35–S45.
- Nickels, L. (2002). Therapy for naming disorders: Revisiting, revising, and reviewing. *Aphasiology*, 16, 935–979.
- Off, C. A., Griffin, J. R., Spencer, K. A., & Rogers, M. A. (2015). The impact of dose on naming accuracy with persons with aphasia. *Aphasiology*, 30, 983–1011. doi:10.1080/02687038.2015.1100705
- Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh Inventory. *Neuropsychologia*, 9, 97–113.
- Pedersen, P. M., Jorgensen, H. S., Nakayama, H., Raaschou, H. O., & Olsen, T. S. (1995). Aphasia in acute stroke: Incidence, determinants, and recovery. *Annals of Neurology*, 38, 659–666.
- Pedersen, P. M., Vinter, K., & Olsen, T. S. (2004). Aphasia after stroke: Type, severity and prognosis. The Copenhagen aphasia study. *Cerebrovascular Diseases*, 17, 35–43.
- Porch, B. E. (1981). *Porch Index of Communicative Ability—Third Edition*. Austin, TX: Pro-Ed.
- Pulvermüller, F., & Berthier, M. L. (2008). Aphasia therapy on a neuroscience basis. *Aphasiology*, 22, 563–599.
- Pulvermüller, F., & Fadiga, L. (2010). Active perception: Sensorimotor circuits as a cortical basis for language. *Nature Reviews Neuroscience*, 11, 351–360.
- Pulvermüller, F., Neininger, B., Elbert, T., Mohr, B., Rockstroh, B., Koebbel, P., & Taub, E. (2001). Constraint-induced therapy of chronic aphasia after stroke. *Stroke*, 32, 1621–1626.
- Pulvermüller, F., & Roth, V. M. (1991). Communicative aphasia treatment as a further development of PACE therapy. *Aphasiology*, 5, 39–50.
- Robey, R. R., & Schultz, M. C. (1998). A model for conducting clinical-outcome research: An adaptation of the standard protocol for use in aphasiology. *Aphasiology*, 12, 787–810.
- Rose, M. L. (2012). Releasing the constraints on aphasia therapy: The positive impact of gesture and multimodality treatments. *American Journal of Speech-Language Pathology*, 22, S227–S239.
- Sickert, A., Anders, L.-C., Munte, T. F., & Sailer, M. (2014). Constraint-induced aphasia therapy following sub-acute stroke: A single-blind, randomized clinical trial of a modified therapy schedule. *Journal of Neurology, Neurosurgery, and Psychiatry*, 85, 51–55.
- Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, 6, 174–215.
- Taub, E., Crago, J. E., Burdick, L. D., Groomes, T. E., Cook, E. W., DeLuca, S. C., & Miller, N. E. (1994). An operant approach to rehabilitation medicine: Overcoming learned non-use by shaping. *Journal of the Experimental Analysis of Behavior*, 61, 281–293.
- Taub, E., Crago, J. E., & Uswatte, G. (1998). Constraint-induced movement therapy: A new approach to treatment in physical rehabilitation. *Rehabilitation Psychology*, 43, 152–170.
- Taub, E., Uswatte, G., & Elbert, T. (2002). New treatments in neurorehabilitation founded on basic research. *Nature Reviews Neuroscience*, 3, 228–236.
- Wambaugh, J. L., Duffy, J. R., McNeil, M. R., Robin, D. A., & Rogers, M. (2006). Treatment guidelines for acquired apraxia of speech: Treatment descriptions and recommendations. *Journal of Medical Speech Language Pathology*, 14, 35–66.
- Warren, S. F., Fey, M. E., & Yoder, P. J. (2012). Differential treatment intensity research: A missing link to creating optimally effective communication interventions. *Mental Retardation and Developmental Disabilities Research Reviews*, 13, 70–77.
- Yorkston, K. M., & Beukelman, D. R. (1980). An analysis of connected speech samples of aphasic and normal speakers. *Journal of Speech and Hearing Disorders*, 45, 27–36.