Long-Term Stability of Short-Term Intensive Language–Action Therapy in Chronic Aphasia: A 1–2 year Follow-Up Study

Lea Doppelbauer, MSc[1,2,3], Bettina Mohr, PhD[4,5], Felix R. Dreyer, PhD[1,6],
Benjamin Stahl, PhD[7,8,9,10], Verena Büscher[1], and Friedemann Pulvermüller, PhD[1,2,3,6]

Abstract

Background. Intensive aphasia therapy can improve language functions in chronic aphasia over a short therapy interval of 2–4 weeks. For one intensive method, *intensive language–action therapy*, beneficial effects are well documented by a range of randomized controlled trials. However, it is unclear to date whether therapy-related improvements are maintained over years. *Objective.* The current study aimed at investigating long-term stability of ILAT treatment effects over circa 1–2 years (8–30 months). *Methods.* 38 patients with chronic aphasia participated in ILAT and were re-assessed at a follow-up assessment 8–30 months after treatment, which had been delivered 6–12.5 hours per week for 2–4 weeks. *Results.* A standardized clinical aphasia battery, the Aachen Aphasia Test, revealed significant improvements with ILAT that were maintained for up to 2.5 years. Improvements were relatively better preserved in comparatively young patients (<60 years). Measures of communicative efficacy confirmed improvements during intensive therapy but showed inconsistent long-term stability effects. *Conclusions.* The present data indicate that gains resulting from intensive speech–language therapy with ILAT are maintained up to 2.5 years after the end of treatment. We discuss this novel finding in light of a possible move from sparse to intensive therapy regimes in clinical practice.

Keywords

aphasia, rehabilitation, speech therapy, intensive language–action therapy

Introduction

A growing body of evidence highlights the effectiveness of *intensive* speech and language therapy (SLT) in chronic aphasia.[1-4] An intensity of 5–10 hours of training per week seems to be necessary to obtain significant improvements of language abilities in people with chronic aphasia (PWCA).[5] Massed-practice protocols may represent an alternative and even more efficient treatment application compared to conventional non-intensive (<5 hours/week) regimes as the same amount of therapy has been reported to yield relatively better outcomes when given in an intensive (rather than sparse temporally spread-out) format, as compared with sparse delivery.[3,6] However, short-term intensive regimes can only provide an advantage over non-intensive treatment if their effects persist over a subsequent therapy break of several months up to years. Although many studies so far demonstrated reliable improvements in language performance immediately after intensive SLT, only few examined whether any improvements obtained in chronic aphasia lasted for longer periods of time. For example, Breitenstein et al. reported improvements in functional communication of 158

[1]Brain Language Laboratory, Freie Universität Berlin, Berlin, Germany
[2]Einstein Center for Neurosciences, Berlin, Germany
[3]Berlin School of Mind and Brain, Humboldt Universität zu Berlin, Berlin, Germany
[4]Department of Psychiatry, Charité Universitätsmedizin, Berlin, Germany
[5]ZeNIS - Centre for Neuropsychology and Intensive Language Therapy, Berlin, Germany
[6]Cluster of Excellence 'Matters of Activity', Humboldt Universität zu Berlin, Berlin, Germany
[7]Department of Neurology, Charité Universitätsmedizin, Berlin, Germany
[8]Department of Neurology, Universitätsmedizin Greifswald, Greifswald, Germany
[9]Department of Neurophysics, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany
[10]Psychologische Hochschule Berlin, Berlin, Germany

Corresponding Author:
Friedemann Pulvermüller, PhD, Department of Philosophy and Humanities, Brain Language Laboratory, WE4 Freie Universität, Berlin 19145, Germany.
Email: friedemann.pulvermuller@fu-berlin.de

PWCA (on average, 35 months post stroke) immediately after intensive therapy, which were maintained 6 months later.[2] An earlier study by Meinzer et al. had previously shown stability of the effect of intensive constraint-induced aphasia therapy (CIAT) at a 6-month follow-up in 27 PWCA (on average, 47.9 months post stroke).[7]

When looking at treatment-related improvements and their stability, it is important to distinguish true intervention effects leading to generalization across linguistic materials from so-called "trivial" effects restricted to items directly trained in therapy. However, only very few studies reported explicitly that language materials trained in therapy were excluded from testing. Recent studies paying close attention to testing and treatment materials found that an intensive naming treatment (INT), given at least 10 hours per week for several weeks in 8[8] and 18[9] PWCA, led to maintenance of naming improvements for up to 8 months. However, maintenance effects were only present for practiced words, without any generalization to unpracticed words. Hence, at least for specific types of aphasia therapy, including naming treatment, general long-term therapy effects across practiced and not-practiced items seem difficult to achieve.[10]

In sum, it is unclear from most pre-existing work on stability of SLT effects in chronic aphasia whether generalized or item-specific therapy effects were measured. Generally, as outlined above, the majority of those few studies looking at maintenance effects[2,7-9,11,12] showed, if any, maintenance of treatment gains for about half a year. However, in order to plan intensive therapy regimes with optimal efficacy and cost-effectiveness, it is important to examine effect stability across longer time windows of one or even 2 years. Such long-term stability would offer novel perspectives for intervention regimes by which short intensive treatment intervals could be separated by long therapy breaks.

Therefore, we examined whether improvements in language performance brought about by a short (2–4 weeks) intensive aphasia treatment would last over a long time span of 8–30 months in 38 PWCA. To exclude the possibility of documenting "trivial" training effects, we used different materials for therapy and for testing. As in previous work,[7] we chose an established intensive SLT, *intensive language–action therapy* (ILAT), a communicative method extending an approach known as CIAT.[3,13] Its effectiveness in chronic aphasia has been confirmed by a range of RCTs[3,7,14-19] and, as mentioned, maintenance of its beneficial effects for 6 months has been reported before.[7] Based on this pre-existing evidence, and given that ILAT is based on neuroscientific principles delivered with high treatment intensity which is necessary to induce long-lasting neural reorganization,[20] we hypothesized that language improvements may even be maintained over a time window substantially exceeding 6 months, possibly up to 2.5 years. An additional aim of this work was to evaluate the impact of individual patient characteristics or clinical variables on long-term stability of language improvements.

Methods

Participants

A power analysis revealed that a minimum number of 27 PWCA was required to find an effect size of pre–post differences on standardized aphasia test batteries after 2–4 weeks of ILAT reported in previous studies (α = .05; $1 - \beta$ = .80; estimated Cohen's f = .625, d = 1.25[14,16]; groups: 1; repeated-measures: 3). Given the long follow-up interval, we assumed a drop-out rate of 30%. PWCAs were consecutively selected from 52 patients who had been recruited in the context of one of 3 RCTs at the Brain Language Laboratory, Freie Universität Berlin, Germany, between January 2014 and May 2019. These RCTs had been registered in the German Clinical Trials Register (identifier: DRKS00005482,[15] DRKS00007829,[16] and DRKS00018057). Participants in the present study met the following inclusion criteria: diagnosis of aphasia by a neurologist, chronic stage of the disease (i.e., duration ≥12 months before the onset of ILAT), completion of 2 language assessments with the primary outcome measure, the Aachen Aphasia Test (AAT)[21] (PRE and POST therapy), absence of significant change (deterioration) in their neurological condition during a period of 8–30 months after conclusion of ILAT, and availability for long-term follow-up testing. Furthermore, participants were right-handed according to the Edinburgh Handedness Inventory,[22] had normal or corrected-to-normal hearing and vision, and showed normal spatial short-term memory, as shown by the Corsi block-tapping test.[23] We excluded patients with aphasia resulting from neurodegenerative disease, pronounced non-verbal cognitive deficits, or severe additional neurological impairments. Out of the 52 available patients, 5 were not included because they could no longer be contacted (n = 3) or because their health condition had changed (n = 2). Of the remaining 47 patients, 9 declined the invitation for a single re-assessment session (most frequent reason: long traveling time); the remaining 38 were included and finished testing (see CONSORT diagram, Figure 1). Aphasia was caused by a single left-hemispheric lesion (Figure 2; see Online Supplemental Table S1 for details on individual lesions). All participants were native speakers of German, most of them monolinguals (n = 35). A majority (n = 29) were situated in an active social environment, living together with a partner, and/or frequently engaging in social interactions with friends or family. Informed consent was obtained from all patients. PWCA received at least 21 hours of ILAT within a period of 2–4 weeks: 14 patients received 10.5 hours per week for 2 weeks,[15] 12 patients received 6 or 12 hours per week for 4 weeks,[16] and 12 patients received 12.5 hours per week for 2 weeks. Detailed patient characteristics and therapy hours are listed in Table 1. All trials were

conducted in accordance with the Declaration of Helsinki; ethics permissions were obtained from the Charité University Hospital Ethics Committees, Campus Benjamin Franklin (EA4/122/12), and Campus Mitte (EA1/382/16), Berlin, Germany.

Intervention: ILAT Protocol

ILAT is characterized by high therapy intensity, action-embedding of language use into social communicative contexts and tailoring of the therapy setting to the patients' individual needs. Language training is embedded in everyday communicative interactions as defined by therapeutic language games. These games are played in groups of 2–3 PWCA and a therapist.[13,24] In each game, a set of 12 picture card pairs is distributed among the players, while barriers between participants prevent them from seeing each other's cards. In a "request game," cards depict objects and the participants' task is to obtain matching pictures by

making verbal requests (e.g., "Give me the [NOUN]"). In a "planning game," cards depict actions and participants have to propose joint performance of activities in order to obtain a matching action card (e.g., "Let's [VERB] together"). In response, co-players follow a request or agree to participate in a proposed action and hand over the corresponding card, if available in their set. If not, co-players reject the request or proposal. Players ask clarifying questions in case of a misunderstanding.[13,25] Therapists and patients are equal participants in the game, although therapists act as a model for appropriate communication, perform repairs, and provide positive feedback.[13] Individual therapeutic goals are defined and difficulty levels adjusted by person-specific constraints.[13] Verbal communication is emphasized, while gesturing is not suppressed. Therapy materials did not include any test items used in language assessments.

Assessment of Language Functions

Participants underwent language assessment prior to (PRE) and immediately after ILAT (POST). Long-term follow-up (LTFU) testing was conducted 8–30 months (*mean* = 15.6, *SD* = 5.7) after the intervention. Diagnostic sessions were conducted by a clinical linguist or neuropsychologist. Apart from the assessment, testers had no contact with patients. In most patients, PRE/POST and LTFU assessments were carried out by different testers. Testers were blind to the group assignment of patients in the respective RCTs; they were not told at which milestone tests were obtained. However, as blinding of testers to the time points of testing could not be guaranteed in every case, an independent reviewer re-evaluated the assessment and scoring by means of audio recordings, and assessment protocols were checked for coherence with the test instructions and manuals. Four subscales of a standardized, impairment-centered aphasia test battery, the AAT[21], were conducted: token test, repetition, naming, and comprehension. An average score, calculated as the average across age-corrected standard *t*-scores from the 4 subscales, was taken as primary outcome measure.[3] Similar to the Western Aphasia Battery's[26] Aphasia Quotient, this
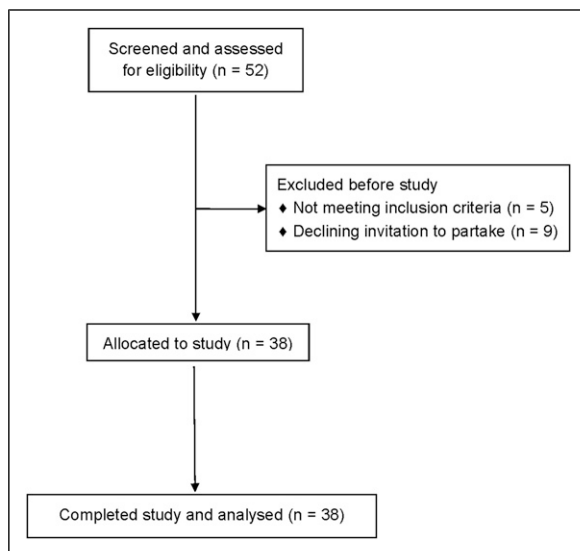
Figure 1. CONSORT diagram of the study of long-term effects of intensive language–action therapy.
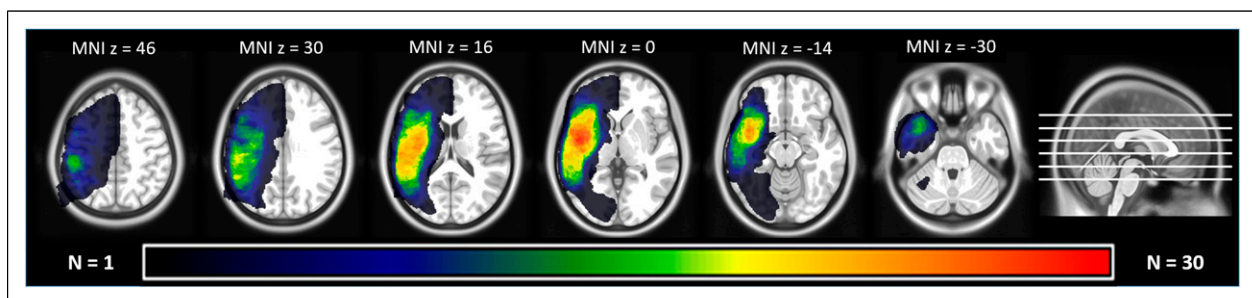
Figure 2. Lesion overlay map of 36 out of the 38 partaking patients (2 were missing imaging data). Different colors indicate the number of lesion overlaps. The largest overlap (n = 28) was found in the left insula/superior temporal cortex.

Table 1. Patient information.

ID	Sex	Age y	Educ y	Etiol	Sever TT	Post on-set, m	ILAT h/w	ILAT Total h	Int'j h/w	Int'j Total h	Lag POST-LTFU, m
1	M	63	16	HMR	66	45	10.5	21	3	89.7	30
2	M	73	10	ISCH	41	61	10.5	21	2	121.2	14
3	F	39	12	ISCH	44	78	10.5	21	2.5	184	17
4	F	49	13	TBI	48	149	10.5	21	0	0	29
5	M	63	13	ISCH	54	31	10.5	21	1.5	129.9	20
6	M	54	21	ISCH	47	49	10.5	21	4	433	25
7	F	47	12	ISCH	55	245	10.5	21	1	34.6	8
8	F	37	11	ISCH	54	30	10.5	21	4	259.8	15
9	F	35	12	ISCH	33	13	10.5	21	2.2	130.3	14
10	M	32	14	ISCH	56	40	10.5	21	4	242.5	14
11	M	62	17	ISCH	42	23	10.5	21	2.5	194.9	18
12	M	49	13	ISCH	39	41	10.5	21	2.4	263.7	25
13	F	45	21	TBI	41	49	10.5	21	2.5	313.9	29
14	M	51	12	ISCH	51	42	10.5	21	2	242.5	28
15	M	74	17	ISCH	42	29	12	48	2	112.6	13
16	M	70	13	ISCH	55	30	12	48	4	225.2	13
17	F	51	18	ISCH	45	23	6	24	3	155.9	12
18	F	77	9	ISCH	50	23	6	24	0	0	12
19	M	74	19	HMR	54	24	6	24	3	168.9	13
20	M	84	11	ISCH	60	53	6	24	1	56.3	13
21	F	82	13	ISCH	48	197	12	48	0	0	14
22	M	45	13	ISCH	37	23	6	24	4	242.5	14
23	F	56	21	HMR	45	12	12	48	3	168.9	13
24	M	48	13	ISCH	45	12	12	48	4	225.2	13
25	F	58	17	ISCH	43	12	6	24	2	103.9	12
26	M	81	13	ISCH	69	44	6	24	2	112.6	13
27	M	62	16	ISCH	50	24	12.5	25	3	142.9	11
28	F	71	17	ISCH	47	104	12.5	25	2	95.3	11
29	M	74	15.5	ISCH	58	25	12.5	25	2	95.3	11
30	M	77	19	ISCH	47	21	12.5	25	2	95.3	11
31	M	67	14	ISCH	45	183	12.5	25	1	56.3	13
32	M	33	13	ISCH	47	18	12.5	25	3	168.9	13
33	M	45	20	ISCH	53	31	12.5	25	3	181.9	14
34	M	58	15.5	ISCH	48	30	12.5	25	2	112.6	13
35	F	60	22	ISCH	55	84	12.5	25	3	194.9	15
36	M	59	17	ISCH	47	28	12.5	25	3	155.9	12
37	M	62	12	ISCH	50	47	12.5	25	4	242.5	14
38	F	76	14	ISCH	55	12	12.5	25	3	155.9	12
Mean		59.03	15.0		49	52	10.5	26.4	2.4	155.5	15.6
SD		14.6	3.4		7	54	2.3	8.7	1.1	90.4	5.7
Range		2–84	9–22		33–69	12–245	6–12.5	21–48	0–4	0–433	8–30

ID, patient ID; Educ-years of education; Etiol, etiology; Sever, severity as suggested by the result of t-scores of the TT; TT, token test (severe: < 43, moderate: 44–53, mild: > 53 impairment); h, hours; HMR, left intracerebral hemorrhage; ILAT, intensive language–action therapy; Int'j, interjacent therapy; ISCH, left middle cerebral artery ischemia; Lag POST-LTFU, time between the end of intensive LST and long-term follow-up testing; m, months; SD, standard deviation; TBI, left traumatic brain injury; y, years.

average score provides an estimate of overall language proficiency and the severity of verbal language deficits. As secondary, non-standardized, outcome measures assessing functional communication, we used the quantitative scale of the communicative activity log[3] (CAL) questionnaire answered by close relatives of PWCA and the Action Communication Test[27] (ACT) that consists of 2 subscales assessing verbal naming or requesting performance in game-like interactions similar to ILAT. At LTFU, patients and, if available, their accompanying person were asked for estimates of the weekly amount of "interjacent" therapy received between POST and LTFU testing milestones.

Statistical Analysis

Analyses were based on mean *t*-scores of the AAT and raw values of ACT and CAL. To assess overall effects, we conducted repeated-measures analyses of variance (ANOVA) with *time* as within-subject factor (3 levels: PRE, POST, and LTFU). Paired-sample two-tailed *t*-tests were used to perform pairwise comparisons of clinical language performance between the following time points: ΔPOST–PRE—evaluating short-term treatment effects immediately after ILAT, ΔLTFU–POST—monitoring performance changes during the 8–30–month follow-up interval, and ΔLTFU–PRE—examining whether any improvement in language performance was maintained over the LTFU. Repeated-measures linear mixed models (LMM) were used to evaluate whether stability of treatment effects was influenced by non–treatment-specific factors: *age*, *education*, *sex*, *aphasia severity* (as defined by *t*-score points of the AAT token test at baseline; severe: <43, moderate: 44–53, mild: >53 impairment), *time since stroke*, *total hours of ILAT*, *duration of the LTFU*, and the *estimated weekly* and *total* (i.e., the *weekly interjacent SLT* multiplied by the duration of the follow-up period in weeks) *amount of interjacent SLT* during the LTFU. All factors entered a full model as two-way interactions (*time**factor). A main effect of *aphasia severity* controlling for baseline performance differences and random intercepts for *participants* was included. The significant factors contributing to long-term stability were selected by using a stepwise procedure of removing interaction effects from the full model. Likelihood-ratio tests (LRTs) and Akaike (AIC) and Bayesian (BIC) information criterion were considered in model comparisons. To fit and analyze LMMs, the "lme4" statistical package (version 1.1-21) in R (version 3.5.3) was used.[28] Model assumptions were evaluated via diagnostic plots. For all statistical tests, two-tailed *P*-values and alpha levels of .05 were applied. Bonferroni corrections were applied in the case of multiple comparisons. The Greenhouse–Geisser correction was used where applicable. See online Supplemental Material for additional control analyses addressing potential variability introduced by the different underlying RCTs and varying aphasia etiologies.

## Results

Primary Outcome Measure: AAT

A repeated-measures ANOVA revealed a significant main effect of *time* (PRE, POST, and LTFU) [$F_{(2,74)} = 16.91$, $P < .001$, $\eta^2_G = .02$]. Paired-sample *t*-tests demonstrated a significant improvement of AAT scores during ILAT [$t(37) = 5.45$, $P < .001$], which was preserved at LTFU relative to baseline [$t(37) = 4.66$, $P < .001$]. There was no significant change of AAT overall scores during the follow-up period [$t(37) = -.51$, $P = 1.0$] (Figure 3, Table 2).

Secondary Outcome Measures: ACT and CAL

Missing data reduced sample sizes of ACT (n = 34) and CAL data (n = 33). An ANOVA performed on ACT results showed a main effect of *time* across all time points after Greenhouse–Geisser correction [Mauchly's test: $\chi^2(2) = 22$, $P < .001$; $F(2, 66) = 4.97$, $\varepsilon = .67$, $P = .022$]. Pairwise comparisons showed a significant improvement after treatment [$t(33) = 5.05$, $P < .001$], but no maintained improvement at LTFU relative to baseline [$t(33) = 2.39$, $P = .07$] (Table 2). However, a stability effect was present on the ACT requesting subscale [$t(33) = 2.51$,
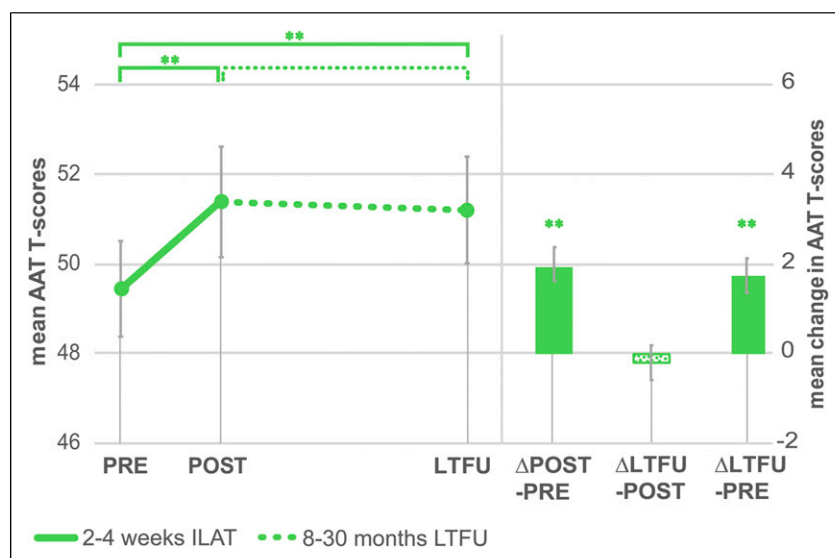
Figure 3. Average *t*-scores on the clinical language test (AAT) before (PRE) and immediately after (POST) ILAT and at the long-term follow-up (LTFU) test 8–30 months later. Significant changes are indicated for the pairwise comparisons (*P < .05, **P < .01). Error bars represent standard errors of the mean. AAT, Aachen Aphasia Test; LTFU, long-term follow-up; ILAT, intensive language–action therapy.

Table 2. Language test results.

	PRE	POST	LTFU	ΔPOST–PRE	ΔLTFU–POST	ΔLTFU–PRE
AAT (n = 38)	49.44	51.38	51.19	1.94**	−.19 ns	1.75**
	(6.67)	(7.59)	(7.37)	(2.19)	(2.29)	(2.32)
ACT (n = 34)	21.43	24.46	24.28	3.03**	−.18 ns	2.85 ns
	(13.96)	(13.54)	(14.33)	(3.50)	(7.61)	(6.95)
CAL (n = 33)	7.97	86.83	83.71	8.86**	−3.12 ns	5.74 ns
	(20.53)	(19.13)	(18.56)	(9.84)	(14.35)	(16.14)

Abbreviations: AAT, Aachen Aphasia Test; ACT, Action Communication Test.
Mean (standard deviation) scores and their changes over time (*P < .05, **P < .01, ns, non-significant). AAT: t-values; ACT, CAL: raw values.

$P = .035$], but not on the naming subscale [$t(33) = 2.01$, $P = .11$].

The CAL revealed a significant main effect of *time* after Greenhouse–Geisser correction [Mauchly's test: $\chi^2(2) = 9.44$, $P = .009$; $F(2,64) = 7.11$, $\varepsilon = .79$, $P = .004$]. Relative to baseline, significantly higher ratings of communicative activity were observed after treatment [$t(32) = 5.17$, $P < .001$], but no longer at LTFU assessment [$t(32) = 2.04$, $P = .15$] (Table 2).

Neurological, Training-Related, and Socio-Demographic Parameters

Analyses focused on performance change from baseline to LTFU (ΔLTFU–PRE). The initial full repeated-measures LMM included the following factors (*time**factor): *age, education, sex, aphasia severity, time since stroke, total hours of ILAT, weekly interjacent SLT, total interjacent SLT,* and *duration of LTFU* (cf. Table 1). All final models included a random factor for *participant* and a fixed factor *aphasia severity*. Relevant interaction terms for final models revealed by the model reduction procedure were *time*age* and *time-*time since stroke* for AAT and *time*age* and *time*duration of LTFU* for ACT results. Both models indicated that only the factor *age* significantly influenced performance stability on the AAT [$F(2,38) = 12.42$, $P < .001$] and ACT [$F(2,34) = 13.61$, $P < .001$]; stability tended to be relatively better for younger patients. A final model for CAL outcomes included the interactions *time*age, time*weekly interjacent SLT, time*total hours of ILAT, time*education,* and *time*sex.* Stability of CAL performance was also influenced by *age* [$F(2,33) = 3.80$, $P = .033$], but furthermore indicated an influence of the *weekly interjacent therapy* [$F(2,33) = 6.08$, $P = .006$], with better stability of CAL ratings in PWCA receiving more weekly interjacent therapy.
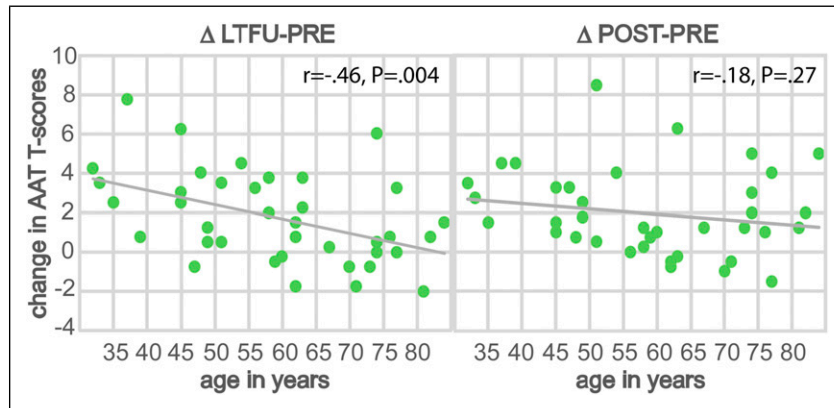
Long-Term Stability and Age

To further examine the relationship between *age* and long-term stability indicated by LMMs, post-hoc analyses focused on AAT performance in specific age groups of the sample. First, younger (<59.5 years, *range* = 32–59, *mean* = 46.9,

$SD = 8.5$, n = 19) and older (>59.5 years, *range* = 60–84, *mean* = 71.2, $SD = 7.6$, n = 19) PWCA were defined by a median-split to ensure equally sized subgroups. An ANOVA including the within-subject factor *time* (PRE and LTFU) and the between-subject factor *age group* (young and old) revealed an interaction of *time*age-group* [$F(1,36) = 8.80$, $P = .005$, $\eta^2_G = .01$]. Post-hoc testing applying Bonferroni correction indicated significant performance gains in the younger ($P < .001$) but not in the older PWCA ($P = .82$). The LMM fitted for AAT results confirmed that age modulation indicated by an interaction effect *time*age* was not present in the younger half of our sample (<59.5 years [$F(2,19) = 3.16$, $P = .07$]), but in subsamples including older PWCA (<70 years [$F(2,26) = 9.14$, $P = .001$]; <80 years [$F(2,35) = 11.21$, $P < .001$]). In all of these subgroups, paired-sample *t*-tests confirmed a significant long-term stability of language performance [ΔLTFU–PRE: <59.5 years: $t(18) = 5.50$, $P < .001$; <70 years: $t(25) = 5.22$, $P < .001$; <80 years: $t(34) = 4.83$, $P < .001$].

To evaluate whether age effects were specific to long-term results or to therapy effects in general, the two-way ANOVA described above was run on the performance change directly after ILAT (POST vs PRE). In this case, results revealed no interaction effect of *time*age-group* [$F(1,36) = 1.86$, $P = .18$]; post-hoc testing indicated significant performance gains in both the younger ($P < .001$) and older ($P = .035$) PWCA. Still, immediate therapy effects appeared as gradually weaker in the older PWCA, as suggested by the LMM where modulatory effects of age were also present in the pre–post comparison [$F(2,38) = 7.18$, $P = .002$]. On the other hand, the correlation of the POST–PRE difference with time was insignificant, whereas that of the stability effect was reliable (Figure 4). Thus, AAT scores significantly improved across ILAT therapy in all age groups, but this gain was differentially maintained, with positive results for relatively younger patients only (cf. Table 3).

Long-Term Stability and Non-Intensive Interjacent Therapy

The frequency of additional interjacent SLT during the LTFU period ranged between 0–4 hours per week, that is, from no to

Figure 4. Long-term stability of language improvements depend on age. Changes of average *t*-scores on the clinical language test (AAT) at LTFU (left) and immediately after ILAT therapy (right) compared to baseline are plotted as a function of age. The patients' age ranged from 32–84 years. A significant correlation with age was present for the stability effect at LTFU. AAT, Aachen Aphasia Test; LTFU, long-term follow-up; ILAT, intensive language–action therapy.

near-intensive treatment; please recall that "intensive" treatment requires 5 or more hours per week.[5] LMM analyses on CAL scores suggested that relatively higher frequency of interjacent SLT led to better maintenance of improvements. Additional analyses using the LMM fitted for CAL data revealed a significant interaction *time***weekly interjacent SLT* also for patients treated with less than 4 hours of weekly interjacent SLT (n = 27) [$F(2,27)$ = 3.77, P = .036], but the effect had vanished in a subsample of PWCA receiving no more than 2.5 hours of weekly extra SLT (n = 19) [$F(2,19)$ = 2.17, P = .14]. Although near-intensive extra SLT might have contributed relatively more to effect stability, no overall long-term effect was documented for CAL results. Importantly, in both subgroups described above, long-term stability of performance was confirmed on the primary outcome measure AAT [<4 hours/week interjacent SLT: $t(30)$ = 3.65, P = .001; ≤2.5 hours/week interjacent SLT: $t(20)$ = 2.80, P = .01]. Moreover, the change in AAT total scores from baseline to LTFU (ΔLTFU–PRE) was comparable in the 3 patients receiving no interjacent SLT (0 hours/week) and in the remaining 35 patients receiving 1–4 hours of interjacent SLT per week (in both groups, 1.75 *t*-score points on average).

Discussion

Here, we report surprisingly long-lasting effects of intensive SLT in chronic aphasia. Improvements of clinical language proficiency after ILAT, an established intensive aphasia treatment, were found to persist for up to 2.5 years. In 38 chronic PWCA, treatment gains obtained over a short period of ILAT, delivered 6–12.5 hours per week for 2–4 weeks, were maintained 8–30 months after the intervention. Stability of language gains was documented by a standardized aphasia test battery, the AAT. Patients' age showed the most robust modulatory effect on long-term stability, with older individuals (>60 years) showing gradually less maintenance of

improvements. The secondary outcome measures, ACT and CAL, which aim to capture communicative effectiveness, confirmed performance improvement across the therapy interval, but did not unambiguously reveal effect stability.

The main finding of this study is that beneficial effects of a short-term intensive aphasia therapy method were maintained over a much longer time period than previously documented. In the few previous studies focusing on the stability of aphasia therapy outcomes, follow-up periods of up to 6 months were most common.[2,7,9,14] As in current clinical practice, most patients might, at best, have access to one intensive therapy period per year, it is of great importance to investigate follow-up intervals of more than 1 year using appropriate group sizes. Based on a properly powered sample (n = 38), our data show that, even after a long period of time, language improvements can persist. It is possible that, similar to ILAT, other massed-practice regimes may lead to lasting beneficial effects, but future research is necessary before any generalizations can be made.

Further advances of our study are the use of both standardized aphasia tests and measures of communicative ability, and the separation of language materials used for training and assessment to exclude "trivial" training effects. Previous work indicated that, in some cases, an effect of aphasia therapy was observed for trained items only. Therefore, it was of great relevance to document that both immediate therapy effects and their long-term maintenance were general and occurred across trained and untrained linguistic materials.

Before discussing results in more detail, we should mention possible limitations of the present work. Speech–language therapists cannot be blinded to the therapy method they deliver, and, although measures to this end were taken, it was not possible to guarantee that testers were always unaware of the milestone of testing. To minimize the possibility of a related bias, we engaged in meticulous control of testing

procedures and results by blinded reviewers (see Methods). Patient selection was guided by MDs' prior assessments, which in 2 cases mismatched with aphasia diagnoses suggested by the Token Test (*t*-values > 63; see Table 1). To exclude a related bias, we repeated analyses without the 2 borderline aphasic patients, which confirmed all main results (significant ΔPOST–PRE and ΔLTFU–PRE differences on the AAT, *P* < 01; n = 36). The range of time lags between intensive SLT and the long-term follow-up examination varied between 8 and 30 months. To reduce the risk of a bias due to great variance or due to relatively short lags, we also repeated the main analyses omitting the single patient with LTFU testing below 12 months (n = 37), which again confirmed the results of the entire sample. Intensive therapy given to the patient sample varied across several dimensions (e.g., duration and intensity); analyses suggesting that this variability was not crucial for the current results are discussed in the Online Supplemental Material.

The pre-existing evidence is mixed regarding the age of patients as a prognostic factor for recovery[29] and the relationship between age and clinical outcomes in aphasia.[4,30,31] Our present results suggest an influence of age on stability effects (Table 3). Explorative evaluations aiming to identify long-term responders showed an effect for younger (<59.5 years), but not for older individuals (>59.5 years) in a split-half analysis. However, when comparing performance before to immediately after ILAT, results indicated significant language improvements in both age groups, even though therapy effects were weaker in older PWCA. In line with this observation, a modulatory effect of age was already present in the pre–post comparison. Therefore, it appears that the age effect was more prominent in long-term outcomes than in immediate therapy effects (ΔPOST–PRE). The average numerical long-term changes on the AAT support this tentative conclusion, as they show a consistent trend toward decreasing stability in the older patients (>60 years). In summary, these results indicate a strong age effect on long-term maintenance of ILAT effects, with some modulatory influence already present during the therapy interval (Figure 4). A possible explanation for these findings are declining neural and cognitive resources with increasing age,[30] which might particularly affect long-term maintenance of newly acquired (language) skills. However, only small samples were available for separate statistical analyses of narrow age groups, so any strong conclusions appear undue. Likewise, given ambiguous pre-existing results of age effects on therapy outcomes,[4,30,31] our present findings would need to be confirmed by future studies to motivate strong conclusions. After all, even our oldest cohort (>70 years) included some PWCA with numerically maintained therapy benefits.

Although most PWCA received some additional non-intensive therapy during the LTFU period, data analysis failed to provide evidence that this "interjacent" therapy was necessary for the documented long-term stability effects. Some modulatory effect of the intensity of extra SLT was

Table 3. Changes in AAT *t*-scores in different age groups.

Age Group	n	ΔPOST–PRE		ΔLTFU–PRE	
		mean	*SD*	*mean*	*SD*
30–40 years	5	3.4	1.3	3.8	2.6
40–50 years	7	2.0	1.0	2.4	2.3
50–60 years	7	2.2	3.1	2.4	1.8
60–70 years	7	0.9	2.5	0.9	1.8
70–80 years	9	1.5	2.3	0.8	2.4
80–90 years	3	2.8	2.0	0.1	1.8

Abbreviations: AAT, Aachen Aphasia Test.

found on communicative activity measured on the CAL, but this non-standardized screening procedure failed to document effect stability. Still, one may argue that patients with near-intensive extra therapy (of 4 hours/week) showed relatively better long-term outcomes. Crucially, however, such near-intensive interjacent therapy was not decisive for the maintenance of treatment effects, as long-term stability of ILAT-induced improvements measured on the primary outcome measure AAT were present in PWCA receiving interjacent SLT of 0–2.5 hours per week. Furthermore, patients receiving no interjacent SLT at all showed numerically comparable stability effects as patients receiving 1–4 hours per week. However, it cannot be excluded that, possibly dependent of the therapy method applied, interjacent therapy delivered at sub-intensive levels might benefit long-term stability. This being said, we should however recall that data on the amount and intensity of interjacent therapy were estimates provided by patients and their accompanying persons. Therefore, they need to be interpreted with care and the issues re-addressed in future work.

Results indicate that improvements in communication achieved during therapy were inconsistently maintained at the LTFU examination, as measured on the non-standardized instruments CAL and ACT. A closer look at the ACT subscales revealed stability effects for requesting but not naming performance. No long-term stability of communicative activity was documented on the CAL questionnaire. Although ILAT may improve motivation and confidence to use language in social communicative interactions,[15] this beneficial effect may disappear over time. Lack of continuous therapeutic intervention and guidance may result in reduced everyday communication and may lead to insufficient positive reinforcement for successful communicative actions. Interestingly, a previous study showed a 6-month stability of improved communication as measured with the Amsterdam–Nijmegen Everyday Language Test (ANELT), a standardized test of communication using role playing.[2] This result, together with our present findings, may suggest that the gradual loss in functional communicative performance may still be insignificant after 6 months but may become significant after 1–2 years. However, we also note that a standardized and

robust test such as the ANELT is better suited for documenting long-term stability than the possibly more variable results of the communicative screenings applied here (ACT and CAL). However, we note that it has recently been argued that all existing tools assessing communicative ability lack sufficient validity, so these possibly diverging results should be interpreted with caution.[32] Therefore, a putative decay of communicative improvements achieved during intensive SLT requires further investigation.

According to previous RCTs, classic aphasia therapy sparsely applied across a long time period might not be efficient in improving language performance in aphasia.[1,5,6,33] Hence, it is necessary to explore new strategies of delivery. In future clinical practice, it may become possible to replace long-term sparse therapy by intervals of short-term intensive treatment. Two necessary milestones have now been achieved in view of this vision: previous studies have already shown the high effectiveness of short-term high-intensity interventions.[2-4,7,15-17,19] Our present study now provides evidence that long-term stability of intensive therapy benefits can be achieved over years. Future research needs to address the questions (1) of how long the interval between subsequent intensive SLT intervals should be to obtain optimal results, (2) whether this lag should be adjusted to age, (3) to what degree interjacent therapy is helpful and with which method and at which intensity this should be given, and (4) whether multiple applications of ILAT and similarly effective methods may lead to accumulating benefits. At the current stage, it appears that a regime including one single 2-week intensive therapy period per year may be an efficient and less resource-consuming alternative to regular sparse therapy on a weekly basis.

Conclusions

This study demonstrates that improvements of language performance in chronic aphasia after ILAT can last for up to 2.5 years. Treatment effects were mapped using test materials different from the items practiced during therapy, thus revealing effect generalization. Long-term stability of treatment effects was confirmed when accounting for standard neurological, therapeutic, and socio-demographic factors. These findings suggest that younger PWCA (<60 years) might maintain therapy-induced language gains more consistently than older patients. Overall, ILAT and possibly other short-term intensive SLT methods may provide an efficient alternative to non-intensive aphasia treatment.

Authors' Contributors

Significant contributions to this study were as follows: Concept and design of the study (BM, FP, and LD), ILAT treatment (BS and VB), statistical analysis (FD, FP, and LD), structural MRI and lesion overlay maps (FD), manuscript drafting and artwork (FP and LD), and revisions (BM, FP, FD, and LD).

ORCID iD

Friedemann Pulvermüller https://orcid.org/0000-0003-3210-7112

Supplemental Material

Supplemental material for this article is available online.

References

1. Brady MC, Kelly H, Godwin J, Enderby P, Campbell P. Speech and language therapy for aphasia following stroke. *Cochrane Database Syst Rev.* 2016; 6: CD000425. doi:10.1002/14651858. CD000425.pub4
2. Breitenstein C, Grewe T, Flöel A, et al. Intensive speech and language therapy in patients with chronic aphasia after stroke: a randomised, open-label, blinded-endpoint, controlled trial in a health-care setting. *Lancet.* 2017; 6736(17): 13-16. doi:10. 1016/S0140-6736(17)30067-3
3. Pulvermüller F, Neininger B, Elbert T, et al. Constraint-induced therapy of chronic aphasia after stroke. *Stroke.* 2001; 32: 1621-1626. doi:10.1161/01.STR.32.7.1621
4. Allen L, Mehta S, Andrew McClure J, Teasell R. Therapeutic interventions for aphasia initiated more than six months post stroke: a review of the evidence. *Top Stroke Rehabil.* 2012; 19(6): 523-535. doi:10.1310/tsr1906-523
5. Bhogal SK, Teasell R, Speechley M. Intensity of aphasia therapy, impact on recovery. *Stroke.* 2003;34(4):987-993. doi: 10.1161/01.STR.0000062343.64383.D0

6. Lincoln NB, Mulley GP, Jones AC, Mcguirk E, Lendrem W, Mitchell JRA. Effectiveness of speech therapy for aphasic stroke patients. *Lancet*. 1984; 323: 1197-1200.

7. Meinzer M, Djundja D, Barthel G, Elbert T, Rockstroh B. Long-term stability of improved language functions in chronic aphasia after constraint-induced aphasia therapy. *Stroke*. 2005; 36(7): 1462-1466. doi:10.1161/01.STR.0000169941.29831.2a

8. Menke R, Meinzer M, Kugel H, et al. Imaging short- and long-term training success in chronic aphasia. *BMC Neurosci*. 2009; 10(10): 118. doi:10.1186/1471-2202-10-118

9. Nardo D, Holland R, Leff AP, Price CJ, Crinion JT. Less is more: neural mechanisms underlying anomia treatment in chronic aphasic patients. *Brain*. 2017; 140(11): 3039-3054. doi: 10.1093/brain/awx234

10. Wisenburn B, Mahoney K. A meta-analysis of word-finding treatments for aphasia. *Aphasiology*. 2009; 23(11): 1338-1352. doi:10.1080/02687030902732745

11. Kurland J, Liu A, Stokes P. Effects of a tablet-based home practice program with telepractice on treatment outcomes in chronic aphasia. *J Speech Lang Hear Res* 2018; 61(5): 1140-1156. doi:10.1044/2018_JSLHR-L-17-0277

12. Grechuta K, Rubio Ballester B, Espín Munne R, et al. Augmented dyadic therapy boosts recovery of language function in patients with nonfluent aphasia. *Stroke*. 2019; 50(5): 1270-1274. doi:10.1161/STROKEAHA.118.023729

13. Difrancesco S, Pulvermüller F, Mohr B. Intensive language-action therapy (ILAT): the methods. *Aphasiology*. 2012; 26(11): 1317-1351.

14. Berthier ML, Green C, Lara JP, et al. Memantine and constraint-induced aphasia therapy in chronic poststroke aphasia. *Ann Neurol*. 2009; 65(5): 577-585. doi:10.1002/ana.21597

15. Stahl B, Mohr B, Dreyer FR, Lucchese G, Pulvermüller F. Using language for social interaction: communication mechanisms promote recovery from chronic non-fluent aphasia. *Cortex*. 2016; 85: 90-99. doi:10.1016/j.cortex.2016.09.021

16. Stahl B, Mohr B, Büscher V, Dreyer FR, Lucchese G, Pulvermüller F. Efficacy of intensive aphasia therapy in patients with chronic stroke: a randomised controlled trial. *J Neurol Neurosurg Psychiatr*. 2017; 89: 586-592. doi:10.1136/jnnp-2017-315962

17. Berthier ML, Pulvermüller F. Neuroscience insights improve neurorehabilitation of poststroke aphasia. *Nat Rev Neurol*. 2011; 7(2): 86-97. doi:10.1038/nrneurol.2010.201

18. Kurland J, Stanek EJ, Stokes P, Li M, Andrianopoulos M. Intensive language action therapy in chronic aphasia: a randomized clinical trial examining guidance by constraint. *Am J Speech Lang Pathol*. 2016; 25: 798–812. doi:10.1044/2016

19. Mohr B, Stahl B, Berthier ML, Pulvermüller F. Intensive communicative therapy reduces symptoms of depression in chronic nonfluent aphasia. *Neurorehabil Neural Repair*. 2017; 31(12): 1053-1062. doi:10.1177/1545968317744275

20. Berthier ML, García-Casares N, Walsh SF, et al. Recovery from post-stroke aphasia: lessons from brain imaging and implications for rehabilitation and biological treatments. *Discov Med*. 2011; 12(65): 275-289. http://www.ncbi.nlm.nih.gov/pubmed/22031666. Accessed November 20, 2017.

21. Huber W, Poeck K, Willmes K. The Aachen Aphasia test. *Adv Neurol*. 1984; 42: 291–303.

22. Oldfield RC. The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*. 1971;9:97–113.

23. Kessels RPC, van Zandvoort MJE, Postma A, Kappelle LJ, de Haan EHF. The Corsi block-tapping task: standardization and normative data. *Appl Neuropsychol*. 2000; 7(4): 252-258. doi: 10.1207/S15324826AN0704_8

24. Pulvermüller F, Berthier ML. Aphasia therapy on a neuroscience basis. *Aphasiology*. 2008; 22(6): 563-599. doi:10.1080/02687030701612213

25. Pulvermüller F. *Aphasische Kommunikation: Grundfragen Ihrer Analyse Und Therapie*. Tübingen: Gunter Narr Verlag; 1990.

26. Kertesz A. *Western Aphasia Battery: Test Manual*. San Antonio: Psychological Corp.; 1982.

27. Stahl B, Mohr B, Dreyer FR, Lucchese G, Pulvermüller F. Communicative-pragmatic assessment is sensitive and time-effective in measuring the outcome of aphasia therapy. *Front Hum Neurosci*. 2017; 11(223): 1-9. doi:10.3389/fnhum.2017.00223

28. Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models usinglme4. *J Stat Software*. 2015;67(1):1-48. doi:10.18637/jss.v067.i01

29. Kertesz A, McCabe P. Recovery patterns and prognosis in aphasia. *Brain*. 1977; 100(1): 1-18. doi:10.1093/brain/100.1.1

30. Ellis C, Urban S. Age and aphasia: a review of presence, type, recovery and clinical outcomes. *Top Stroke Rehabil*. 2016; 23(6): 430-439. doi:10.1080/10749357.2016.1150412

31. Plowman E, Hentz B, Ellis C. Post-stroke aphasia prognosis: a review of patient-related and stroke-related factors. *J Eval Clin Pract*. 2012; 18(3): 689-694. doi:10.1111/j.1365-2753.2011.01650.x

32. Wallace SJ, Worrall L, Rose T, et al. A core outcome set for aphasia treatment research: the ROMA consensus statement. *Int J Stroke*. 2018;14(2):180-185. doi:10.1016/j.irfa.2015.11.004

33. Code C, Petheram B. Delivering for aphasia. *Int J Speech Lang Pathol*. 2011; 13(1): 3-10. doi:10.3109/17549507.2010.520090