



OPEN ACCESS

RESEARCH PAPER

Efficacy of intensive aphasia therapy in patients with chronic stroke: a randomised controlled trial

Benjamin Stahl,^{1,2,3,4} Bettina Mohr,⁵ Verena Büscher,⁶ Felix R Dreyer,⁶ Guglielmo Lucchese,^{2,6} Friedemann Pulvermüller^{6,7}

► Additional material is published online only. To view please visit the journal online (<http://dx.doi.org/10.1136/jnnp-2017-315962>).

¹Department of Neurology, Charité Universitätsmedizin Berlin, Campus Mitte, Berlin, Germany

²Department of Neurology, Universitätsmedizin Greifswald, Greifswald, Germany

³Department of Neurophysics, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

⁴Psychologische Hochschule Berlin, Berlin, Germany

⁵Department of Psychiatry, Charité Universitätsmedizin Berlin, Campus Benjamin Franklin, Berlin, Germany

⁶Department of Philosophy and Humanities, Brain Language Laboratory, Freie Universität Berlin, Berlin, Germany

⁷Berlin School of Mind and Brain, Humboldt-Universität zu Berlin, Berlin, Germany

Correspondence to

Dr Benjamin Stahl, Charité Universitätsmedizin Berlin, Department of Neurology, Charitéplatz 1, 10117 Berlin Germany; benjamin.stahl@charite.de

Received 1 March 2017

Revised 13 November 2017

Accepted 14 November 2017

ABSTRACT

Objective Recent evidence has fuelled the debate on the role of massed practice in the rehabilitation of chronic post-stroke aphasia. Here, we further determined the optimal daily dosage and total duration of intensive speech-language therapy.

Methods Individuals with chronic aphasia more than 1 year post-stroke received Intensive Language-Action Therapy in a randomised, parallel-group, blinded-assessment, controlled trial. Participants were randomly assigned to one of two outpatient groups who engaged in either highly-intensive practice (Group I: 4 hours daily) or moderately-intensive practice (Group II: 2 hours daily). Both groups went through an initial waiting period and two successive training intervals. Each phase lasted 2 weeks. Co-primary endpoints were defined after each training interval.

Results Thirty patients—15 per group—completed the study. A primary outcome measure (Aachen Aphasia Test) revealed no gains in language performance after the waiting period, but indicated significant progress after each training interval (gradual 2-week *t*-score change [CI]: 1.7 [±0.4]; 0.6 [±0.5]), independent of the intensity level applied (4-week change in Group I: 2.4 [±1.2]; in Group II: 2.2 [±0.8]). A secondary outcome measure (Action Communication Test) confirmed these findings in the waiting period and in the first training interval. In the second training interval, however, only patients with moderately-intensive practice continued to make progress (Time-by-Group interaction: $P=0.009$, $\eta^2=0.13$).

Conclusions Our results suggest no added value from more than 2 hours of daily speech-language therapy within 4 weeks. Instead, these results demonstrate that even a small 2-week increase in treatment duration contributes substantially to recovery from chronic post-stroke aphasia.

INTRODUCTION

A long-standing controversy has surrounded the appropriate quantity of speech-language therapy (SLT) in the rehabilitation of chronic post-stroke aphasia. Although clinical research highlights the importance of massed practice in SLT,¹ the effective delivery of intensive regimens is not yet fully understood.² From a conceptual perspective, the outcome of intensive SLT may depend on two discrete features: (i) the amount of weekly practice, and (ii) the total duration of the training period. Literature reviews indeed suggest that a weekly dosage ranging from 5 to 10 hours, referred to

as ‘moderately-intensive’, is sufficient to ensure significantly improved language performance on standardised aphasia test batteries.^{3–5} However, studies so far do not offer insight into whether massed practice in a weekly dosage over and above 10 hours, referred to as ‘highly-intensive’, leads to further gains in SLT. Moreover, weekly practice and overall treatment duration are negatively correlated across studies, making it difficult to determine the influence of the training period on symptom recovery. The current work seeks to address both of these issues.

To date, surprisingly few studies have focused on the amount of weekly practice and duration of the training period in intensive SLT. A randomised controlled trial (RCT) indicated a superiority of Constraint-Induced Aphasia Therapy administered in a highly-intensive fashion (weekly practice: approximately 16 hours; duration: 2 weeks) over traditional, moderately-intensive SLT (weekly practice: 6–8 hours; duration: 4–5 weeks).⁶ The design of this RCT balanced the total number of hours provided in each type of training, whereas it did not match the treatment protocols and clinical settings, such as the one-to-one or patient-group context, along with the selection and variety of therapists. Using well-matched treatment protocols and clinical settings, a subsequent non-RCT study revealed greater progress in language performance with moderately-intensive SLT (weekly practice: 6 hours; duration: 8 weeks) compared with highly-intensive SLT (weekly practice: 16 hours; duration: 3 weeks).⁷ Still, this study did not randomly assign patients to the treatment groups, and therefore prevents definitive conclusions with regard to the ideal amount of weekly practice and duration of the training period.

Here, we present the first RCT evidence on the efficacy of Intensive Language-Action Therapy (ILAT, an expanded version of Constrained-Induced Aphasia Therapy requiring request and planning communication) applied with different degrees of massed practice (12 hours vs 6 hours per week). Both intensity levels reached the estimated minimum dosage of 5–10 hours weekly to assess the potential benefit of further daily practice in individuals with chronic post-stroke aphasia.^{3–5} Apart from the intensity levels, the treatments were based on identical protocols, materials and procedures to overcome methodological problems of previous work. To explore the possible impact of treatment duration on symptom recovery, all patients went



CrossMark

To cite: Stahl B, Mohr B, Büscher V, et al. *J Neurol Neurosurg Psychiatry* Published Online First: [please include Day Month Year]. doi:10.1136/jnnp-2017-315962

through an initial waiting period and two successive training intervals (each phase lasting 2 weeks). Although specific predictions would be premature in light of currently available RCT and non-RCT results, these data do not rule out an added value (i) when increasing the amount of weekly practice over and above 5–10 hours, and (ii) when extending the duration of the training period up to 1 month.⁸

METHODS

Study design and setting

We conducted a randomised, parallel-group, blinded-assessment, controlled trial in an outpatient centre at the Freie Universität Berlin, Germany, from 2015 to 2016. The trial was registered prospectively (German Clinical Trials Register; identifier: DRKS00007829).

Participants

Recruitment was administered in collaboration with several local rehabilitation centres and support groups for individuals with aphasia. After routine referral to the study team, we contacted the potential participants and invited them to a screening session to check their eligibility. The inclusion criteria were as follows: diagnosis of aphasia, as confirmed by the Aachen Aphasia Test (AAT)⁹; chronic stage of symptoms at least 1 year post-onset of stroke to prevent non-treatment effects related to spontaneous recovery; German as first native language; and right-handedness according to the Edinburgh Handedness Inventory.¹⁰ The exclusion criteria were as follows: aphasia due to traumatic brain injury or neurodegenerative disease; severe non-verbal cognitive deficits, as confirmed by the Corsi Block-Tapping Task¹¹; severe uncorrected vision or hearing disorders; other untreated medical conditions; and intensive SLT in the 2 years prior to study enrolment. Thirty individuals with chronic post-stroke aphasia were recruited, screened and agreed to participate in the present RCT (for details, see figure 1). This sample size was calculated in a previous power analysis ($\alpha=0.05$; $1-\beta=0.95$; number of groups: 2; number of repeated measures: 4; Cohen's $f=0.3$ for our primary outcome, the AAT, derived from Stahl *et al.* 2016¹², resulting $n=26$; assumed dropout rate: 10%; final $n=30$).¹³ On average, patients were aged 60.1 years (SD: 15.3 years) and 65.2 months post-onset of stroke (SD: 64.3 months).

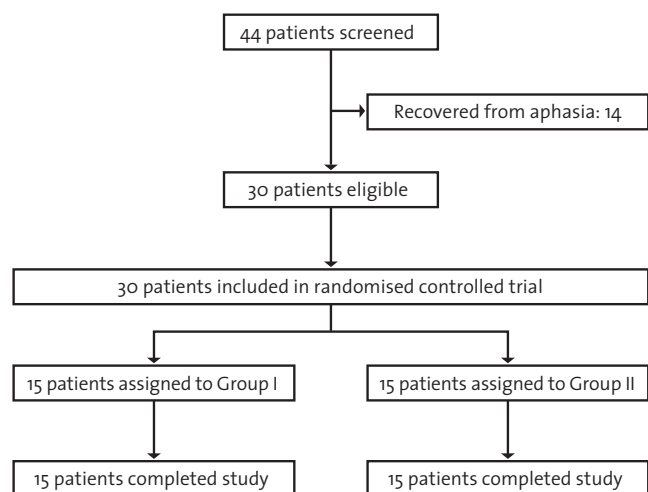


Figure 1 Consolidated Standards of Reporting Trials flow diagram.

Randomisation and blinding

A Python script generated a list of 30 random numbers (0 or 1) resulting in treatment groups of the same size (each $n=15$). The script did not consider any additional variables. Instead, we included key baseline characteristics in our statistical analyses, as specified below. Patients were randomly assigned to one of the two treatment groups receiving ILAT with different degrees of massed practice (Group I: 4 hours of daily training; Group II: 2 hours of daily training). The group allocation was executed by an individual who alone had access to the list of random numbers and who did not participate in any stage of recruiting, screening, therapy or testing. A clinical linguist and neuroscientist performed all diagnostic sessions. The person was blinded to the group assignment and did not have patient contact aside from the testing. Data were unmasked only for final evaluation purposes by the study team who did not attend any of the diagnostic sessions.

Intervention: treatment protocols, materials and procedures

ILAT required patients to engage in everyday request and planning communication with related social interaction.¹⁴ Groups of three patients and a therapist were seated around a table and provided with picture cards showing different objects (e.g., bottle) or action scenes (e.g., drinking). Barriers on the table prevented players from seeing each others' cards. Each card had a duplicate that was owned by one of the other players. The goal was to obtain this duplicate from a fellow player by requesting the depicted object (e.g., 'Give me the [...]') or by proposing an action based on the visualised scene (e.g., 'Let's [...] together'). If the duplicate was available, the players compared the depicted objects or action scenes. In the case of a match, the addressee handed over the corresponding card to the person who initiated the request or action-planning sequence. If the duplicate was not available, the addressee rejected the request or proposed action. In the event of misunderstandings, the players asked clarifying questions. The complexity of the communicative interaction was tailored to the patients' individual language skills by varying the difficulty level of the target words and sentence structures.

The training materials consisted of 336 items presented in separate cards sets, each including 12 picture pairs. For tailoring these sets to individual language skills, the following difficulty levels were available: nouns with high ($n=48$ different pictures), medium ($n=48$) and low ($n=48$) normalised lemma frequency; nouns with phonological similarities (e.g., 'ball' and 'wall'; $n=96$); nouns from only one semantic category ($n=48$); action-related verbs with high ($n=24$) and low ($n=24$) normalised lemma frequency; and action-related verbs applied to one target object (e.g., 'to peel, cut, grate or eat an apple'; $n=48$). Card sets of one difficulty level were matched for mean normalised lemma frequency. All 28 card sets were split into two parts with equal numbers of items per difficulty level and assigned to the two training intervals in counterbalanced order across treatment groups. In all training sessions, the therapist (i) encouraged players by giving positive feedback, (ii) acted as a model by using individually appropriate sentence structures (e.g., polite requests for patients with prevailing agrammatism: 'Would you consider passing me the [...], please?'), and (iii) embedded semantic cues in turn-taking sequences, whenever helpful (e.g., proposals for persons with word-finding deficits: 'May I offer you that tool to cut things?'—'Knife, yes!'). The therapist did not offer other types of cues, whether phonemic (e.g., initial word sounds: 'It starts with /n/...'—'Knife') or graphemic ones (e.g., reading or writing), nor did the patients

Table 1 Intensive regimens

	Group I	Group II
Intervention type	ILAT	ILAT
Daily practice	4 hours	2 hours
Weekly practice	12 hours	6 hours
Therapy frequency	3 weekly sessions	3 weekly sessions
Duration of each trial phase	6 consecutive working days	6 consecutive working days
Total amount of practice	48 hours	24 hours
Total treatment duration	4 weeks	4 weeks

Thirty patients with chronic post-stroke aphasia were randomly assigned to one of two treatment groups: patients receiving ILAT with 4 hours (Group I) or with 2 hours of daily practice (Group II).

ILAT: Intensive Language-Action Therapy.

repeat verbal utterances on instruction (e.g., ‘It is a knife. What is it?’—‘Knife’). Self-cueing strategies were allowed, along with gestures to accompany—but not replace—spoken language.

Treatment was delivered by four therapists who received special training and continuous supervision before and during the trial. Notably, the selection and number of therapists did not differ between the treatment groups. Cohorts of three patients who were relatively heterogeneous with regard to symptom severity underwent ILAT with the degree of massed practice determined by the randomisation procedure described above (4 hours vs 2 hours of daily training). Therapy frequency was consistent across treatment groups (always 3 weekly sessions). Both treatment groups went through an initial waiting period and two successive training intervals. Each phase lasted 2 weeks (six consecutive working days, always separated by a weekend). Depending on the intensity level, the two training intervals

involved overall 48 hours (Group I) or 24 hours of practice (Group II) within 4 weeks (see [table 1](#)). Patients completed all training sessions and did not attend any other form of SLT throughout the entire trial.

Baseline data

Each patient met the diagnostic criteria of aphasia according to the AAT.⁹ Since patients with aphasia often suffer from concomitant deficits in motor planning, it is worth noting that Group I and Group II were similarly affected by apraxia of speech, as confirmed by two independent clinical linguists with high inter-rater agreement (Cohen’s $\kappa = 1.0$). Focusing on non-verbal short-term memory, our patient sample scored, on average, within the normal range on the Corsi Block-Tapping Task.¹¹ Structural T₁-weighted MRI was performed for all individuals using a 3T Magnetom Trio scanner (Siemens Medical Solutions, Erlangen, Germany). All patients had suffered a single cerebrovascular accident with subsequent lesions in parts of the left frontal, parietal and temporal lobes, as well as in adjacent subcortical areas. Two clinical neuroscientists manually delineated and superimposed the precise locations of lesioned voxels in each patient using the software MRIcron (for lesion overlay maps, see [figure 2](#); for individual case histories and baseline test scores, see online supplementary file).¹⁵

Testing and outcomes

All diagnostic sessions were conducted by a clinical linguist and neuroscientist who was blinded to the group allocation, as indicated above. In a repeated-measures design, testing took place 2 weeks before (T₀), 1 day before (T₁) and 1 day after the first training interval (T₂; co-primary endpoint with reference to T₁),

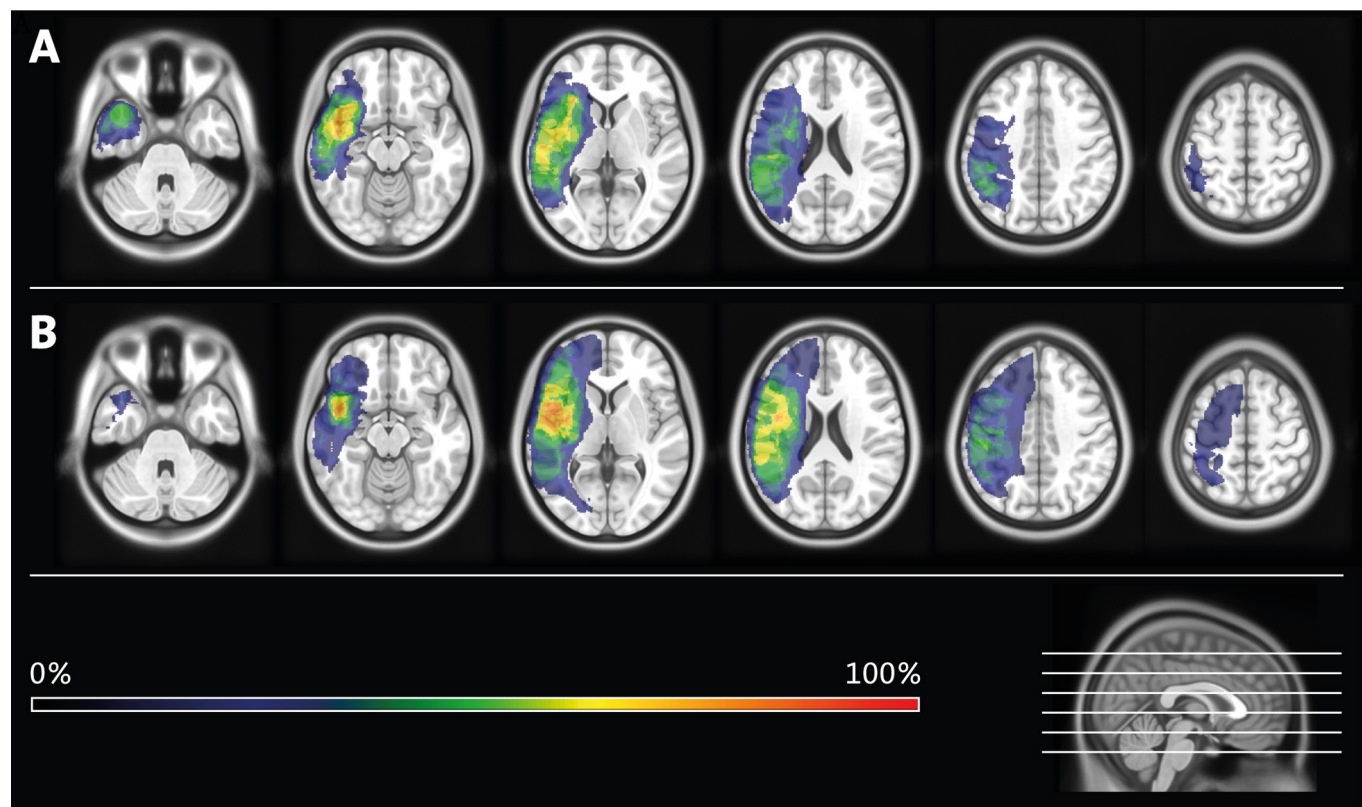


Figure 2 Lesion overlay maps. Patients received Intensive Language-Action Therapy with 4 hours (Group I; see panel A) or with 2 hours of daily practice (Group II; see panel B). Different colours indicate the degree of lesion overlap in each treatment group.

as well as 1 day after the second training interval (T_3 ; co-primary endpoint with reference to T_2).

As a *primary* outcome measure, we administered an impairment-centred aphasia test battery known for its good construct validity and test-retest reliability, the AAT.⁹ Language performance was measured on four subscales of the battery: Token Test, Repetition, Naming and Comprehension. We excluded the AAT subscales Spontaneous Speech (due to its partly insufficient construct validity) and Writing (given the focus on spoken language in our treatment). AAT results were designated as normally distributed *t*-scores, averaged across subscales.

As a *secondary* outcome measure, we used a newly created and published diagnostic instrument, the Action Communication Test (ACT).¹⁶ Motivated by the lack of both linguistic and functional aphasia test batteries with documented psychometric properties, this method reflects impairment-centred and communicative-pragmatic aspects of language processing. In step 1 of the testing, sets of five real generic objects are presented on a table. The patient is asked to name each of these objects, one by one. In step 2 of the testing, the patient verbally requests sets of five objects presented on the table, again one by one. Whenever utterances are correct, the experimenter hands over the requested object to the patient, who in turn places it in a bag. Materials of the ACT consisted of 50 common objects that were allocated to two parallel test versions, each including five sets of five items. Test versions were selected in counterbalanced order across patients. The scoring system was as follows: two points for correctly produced target words; one point for correctly produced target words on the second attempt or incorrect, but semantically or phonologically related utterances; and no point for any further utterances or omissions. Based on these ratings, the average total number of points was expressed as *t*-scores.

Statistical analyses

For each outcome, a repeated-measures analysis of variance (ANOVA) was conducted, including within-subject factor Time (T_0 ; T_1 ; T_2 ; T_3) and between-subject factor Group (Group I; Group II), covaried for pre-treatment performance (T_0) on the AAT or ACT, and for history of previous intensive SLT (weekly dosage ≥ 5 –10 hours) more than 2 years prior to trial onset (yes; no). Two-tailed *P* values and alpha levels of 0.05 were applied for all statistical tests; for multiple comparisons, we used the Bonferroni-Holm correction.

RESULTS

According to independent-sample *t*-tests, the randomisation procedure did not lead to significant differences between Group I and Group II with regard to age, education level, months after onset of disease, non-verbal short-term memory and individual lesion size. Crucially, independent-sample *t*-tests also demonstrated that Group I and Group II did not differ significantly with regard to their performances on the AAT [$t(28)=1.64$, $P=0.62$, not significant (NS)] or on the ACT [$t(28)=1.69$, $P=0.62$, NS] at baseline (T_0). The treatment groups were comparable in terms of gender, clinical diagnoses and history of intensive SLT more than 2 years prior to study enrolment (for group averages and SDs, see online supplementary file).

Focusing on the AAT scores, the repeated-measures ANOVA revealed a significant main effect of the factor Time [$F(3, 78)=4.10$, $P=0.009$, $\eta^2=0.11$]. The ANOVA interaction of Time and Group failed statistical significance [$F(3, 78)=0.80$, NS]. Subsequent post-hoc paired-sample *t*-tests indicated no significant changes in language performance after the initial waiting

period [absolute increase in both groups between T_0 and T_1 (CI): 0.3 (0.5); $t(29)=1.02$, NS], but showed significant progress in each of the two training intervals [increase between T_1 and T_2 : 1.7 (0.4); $t(29)=7.73$, $P<0.001$; Cohen's $d_z=1.4$; between T_2 and T_3 : 0.6 (0.5); $t(29)=2.31$, $P=0.03$; Cohen's $d_z=0.4$] and across the entire therapy phase [increase between T_1 and T_3 : 2.3 (0.7); $t(29)=6.25$, $P<0.001$; Cohen's $d_z=1.1$; see figure 3A and table 2].

Based on the ACT scores, the ANOVA yielded a significant interaction of the factors Time and Group [$F(3, 78)=4.17$, $P=0.009$, $\eta^2=0.13$]. According to post-hoc paired-sample *t*-tests, changes in language performance were absent in the initial waiting period [absolute increase in both groups between T_0 and T_1 (CI): 0.2 (0.4); $t(29)=0.95$, NS] and observed only in the first training interval [increase between T_1 and T_2 : 1.8 (0.7); $t(29)=5.47$, $P<0.001$; Cohen's $d_z=1.0$] as well as across the entire therapy phase [increase between T_1 and T_3 : 1.9 (0.8); $t(29)=4.50$, $P<0.001$; Cohen's $d_z=0.8$]. In the final training interval, only patients with moderately-intensive practice continued to make progress [increase in Group II between T_2 and T_3 : 0.6 (0.5); $t(14)=2.32$, $P=0.04$; Cohen's $d_z=0.6$], while patients with highly-intensive practice did not [decrease in Group I between T_2 and T_3 : -0.4 (0.5); $t(14)=-1.64$, NS; see figure 3B and table 2].

DISCUSSION

The present RCT aimed to determine the ideal amount of daily practice and total duration of the training period in intensive SLT. Thirty individuals with chronic post-stroke aphasia received ILAT in two groups with different degrees of massed practice (4 hours vs 2 hours per day). All patients went through an initial waiting period and two successive training intervals. Each phase lasted 2 weeks. Co-primary endpoints were defined after each training interval. We carefully controlled for the experimental setting, including the patient-group context, the variety of items practised throughout the training sessions, as well as the selection and number of therapists. A standardised aphasia test battery (AAT) revealed no changes in language performance after the waiting period, but indicated significant and clinically relevant progress after each of the two training intervals (ANOVA main effect of Time: $P=0.009$; increase in both groups between T_1 and T_2 [CI]: 1.7 [± 0.4]; between T_2 and T_3 : 0.6 [± 0.5]; medium-to-large effects: $0.4 < \text{Cohen's } d_z \leq 1.4$). Crucially, any such progress did not depend on the intensity level applied (no significant ANOVA interaction of Time and Group; increase between T_1 and T_3 in Group I: 2.4 [± 1.2]; in Group II: 2.2 [± 0.8]). This finding is consistent with the observation that the AAT data showed similar patterns of individual changes in language performance over time, regardless of symptom severity. Both treatment groups completed all training sessions, confirming the good compliance of chronic patients in intensive SLT.⁵ The current results suggest no benefit from more than 2 hours of daily practice within 1 month, whereas a 2-week extension of treatment duration adds to the efficacy of intensive SLT.

The findings reported here may seem in conflict with established Hebbian principles, according to which the repetitive and conjoint firing of neurons is likely to strengthen the synaptic connectivity between them.¹⁷ Although this neurobiological model appears to imply that the functional reorganisation of language increases with the amount of daily practice, it does not postulate unlimited learning capacities in a relatively short period of time.¹⁸ Instead, Hebbian principles do not rule out the possibility of a ceiling effect within a single day, if the

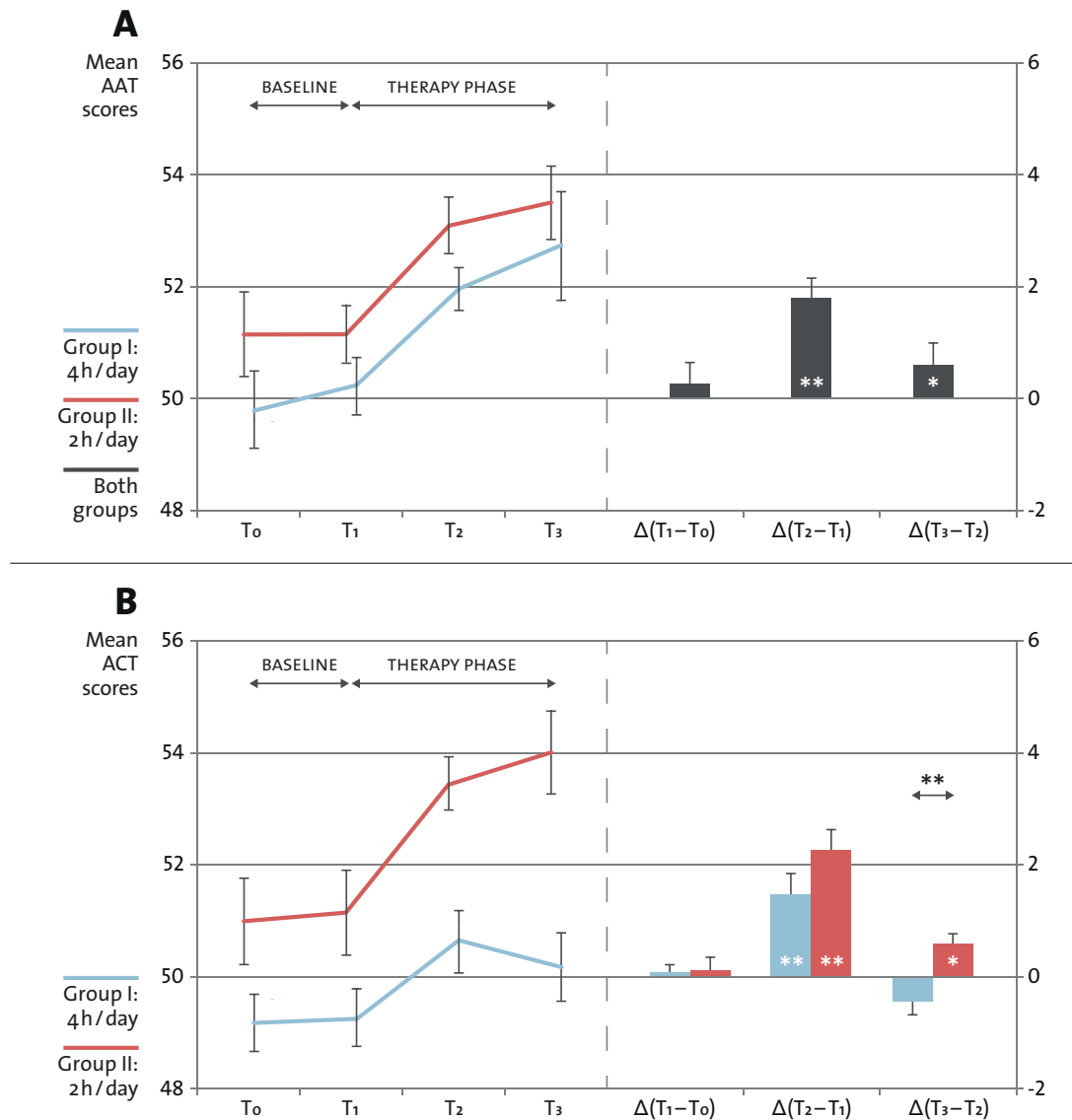


Figure 3 Aphasia test results. Changes in language performance on the Aachen Aphasia Test (AAT; see panel A) and on the Action Communication Test (ACT; see panel B). Thirty individuals with chronic post-stroke aphasia were randomly assigned to one of the two Groups: patients receiving Intensive Language-Action Therapy with 4 hours (Group I) or with 2 hours of daily practice (Group II). All patients went through an initial waiting period ('baseline') and two successive training intervals ('therapy phase'). Each trial phase lasted 2 weeks. Testing took place at four points in Time: 2 weeks before treatment onset (T₀), at treatment onset (T₁), after the first training interval (T₂) and after the second training interval (T₃). Focusing on changes in language performance separately for each trial phase [$\Delta(T_1-T_0)$; $\Delta(T_2-T_1)$; $\Delta(T_3-T_2)$], statistics refer to significant paired-sample *t*-tests (asterisks embedded in bar graphs) and to a significant Time-by-Group interaction, as revealed by repeated-measures analyses of variance (asterisks displayed above bar graphs; **P*<0.05, ***P*<0.01). Error bars represent CIs corrected for between-subject variance.²² Independent-sample *t*-tests confirmed that Group I and Group II did not differ significantly with regard to their performances on the AAT (*P*=0.62) or on the ACT (*P*=0.62) at baseline (T₀).

treatment intensity exceeds a certain threshold. Support for this claim comes from learning psychology, predicting a decline of attention as a consequence of habituation,¹⁹ and from clinical neuroscience, discussing a ceiling effect after intensive SLT and simultaneous dopaminergic medication.²⁰ We nonetheless wish to emphasise that, in the present RCT, we compared two intensive forms of SLT with 4 hours (Group I) or 2 hours of daily practice (Group II). In contrast, traditional SLT in most industrial countries rarely amounts to more than 3 hours of weekly practice, based on our experience. Such a low dosage fails to reach the estimated minimum of 5–10 hours weekly to ensure progress in SLT.^{3–5} In line with this view, a multiple-case study tended to deliver better results for Constraint-Induced Aphasia Therapy if administered in a weekly dosage of 15 hours compared with

only 3 hours, in spite of a higher treatment duration in the latter group.²¹ Therefore, a potential ceiling effect in our RCT should not undermine the importance of offering intensive SLT to individuals with chronic post-stroke aphasia.

The treatment protocols in the current RCT differed in their intensity level, but not in the communicative-pragmatic nature of the training itself. One may thus argue that our results do not generalise to other forms of SLT. However, we do not see any reasons why the observed benefit from a longer training period should be limited to communicative-pragmatic methods in SLT. Rather, it is worth considering that a second RCT with precisely matched degrees of massed practice and treatment duration has indicated a superiority of ILAT over confrontation naming in persons with chronic non-fluent aphasia.¹² Taken together,

Table 2 Aphasia test results

	T ₀	T ₁	T ₂	T ₃	Δ(T ₁ -T ₀)	Δ(T ₂ -T ₁)	Δ(T ₃ -T ₂)	Δ(T ₃ -T ₁)
Mean AAT scores								
Group I (CI)	49.7 (3.7)	50.3 (4.1)	51.9 (4.3)	52.7 (4.7)	0.6 (0.8)	1.6*** (0.6)	0.8 (0.8)	2.4** (1.2)
Group II (CI)	51.2 (4.5)	51.2 (4.3)	53.0 (4.7)	53.4 (4.8)	0.0 (0.6)	1.8*** (0.6)	0.4 (0.6)	2.2*** (0.8)
Both groups (CI)	50.5 (2.9)	50.7 (3.0)	52.5 (3.1)	53.1 (3.3)	0.3 (0.5)	1.7*** (0.4)	0.6* (0.5)	2.3*** (0.7)
Mean ACT scores								
Group I (CI)	49.1 (4.7)	49.2 (4.8)	50.6 (5.0)	50.2 (4.9)	0.1 (0.3)	1.4** (0.9)	-0.4 (0.5)	1.0 (1.0)
Group II (CI)	50.9 (5.5)	51.1 (5.8)	53.4 (5.9)	54.0 (5.6)	0.2 (0.7)	2.3*** (0.9)	0.6* (0.5)	2.9*** (1.2)
Both groups (CI)	50.0 (3.6)	50.2 (3.7)	52.0 (3.8)	52.1 (3.7)	0.2 (0.4)	1.8*** (0.7)	0.1 (0.4)	1.9*** (0.8)

Mean *t*-scores obtained on the Aachen Aphasia Test (AAT)⁹ and on the Action Communication Test (ACT).¹⁶ Thirty individuals with chronic post-stroke aphasia were randomly assigned to one of two treatment groups: patients receiving Intensive Language-Action Therapy with 4 hours (Group I) or with 2 hours of daily practice (Group II). Both treatment groups went through an initial waiting period and two successive training intervals. Each phase lasted 2 weeks. Patients were tested at four points in time: 2 weeks before treatment onset (T₀), at treatment onset (T₁), after the first training interval (T₂) and after the second training interval (T₃). Asterisks refer to significant paired-sample *t*-tests (**P*<0.05, ***P*<0.01, ****P*<0.001) assessing changes in language performance separately for each time interval [Δ(T₁-T₀); Δ(T₂-T₁); Δ(T₃-T₂)] and across the entire therapy phase [Δ(T₃-T₁)].

data from these two RCTs suggest a moderately-intensive use of communicative-pragmatic methods applied over an extended period of time. Additional evidence will be required to determine whether or not a further reduction in the daily amount of practice has an effect on the outcome of SLT. Likewise, it remains to be investigated whether a slight decrease or further increase in treatment duration leads to similar results.

The present RCT included an impairment-centred aphasia test battery (AAT),⁹ along with a dialogue-sensitive diagnostic instrument (ACT).¹⁶ This secondary outcome measure was motivated by the communicative-pragmatic character of the training and its potential relevance to everyday discourse. Both outcome measures indeed revealed congruent changes on the ACT when focusing on the waiting period and the first training interval. In the second training interval, however, the ACT yielded significant progress only in patients with 6 hours of weekly practice (ANOVA interaction of the factors Time and Group: *P*=0.009; $\eta^2=0.13$; *t*-score change between T₂ and T₃ in Group I [CI]: -0.4 [±0.5]; in Group II: 0.6 [±0.5]). This finding is unlikely to arise from discrepant baseline performances between treatment groups, as patients with highly-intensive practice seemed to have marginally—but non-significantly—more room for improvement over time. The finding also converges with previous non-RCT evidence, pointing to optimal gains in SLT with 6 hours of weekly practice provided over an extended period of time.⁷ As one possible reason for the advantage of moderately-intensive practice on the ACT, we propose that keeping the daily quantity of SLT below a critical threshold may diminish post-treatment fatigue, and hence facilitate immediate learning transfer to everyday situations after each training session. Although more research is needed to substantiate the particular influence of moderately-intensive practice on everyday discourse, such considerations should not overshadow the fact that both of our outcome measures consistently confirmed the success of prolonged SLT with at least 6 hours of weekly training.

This is the first RCT to directly compare the efficacy of intensive SLT with different degrees of massed practice and otherwise identical experimental conditions over the course of 4 weeks. The current results suggest no added value of treatment intensity over and above 2 hours of daily practice within 4 weeks. Instead,

these results demonstrate that even a small 2-week increase in treatment duration contributes to recovery from chronic post-stroke aphasia. In light of previous concerns about the feasibility of highly-intensive SLT, we here show that a lower-than-expected dosage of 2 hours per day is sufficient and therefore easier to achieve within the constraints of clinical practice.

Acknowledgements We wish to thank our patients for engaging in the daily therapy sessions, along with our team Valerie Keller, Charlotte Lion, Lena Lorenz and Saskia Millrose. We also wish to thank Laura Besch, Sebastian Kaim, Dr Cora Kim and Laura Schiemann for helping us create our training materials.

Contributors Significant contributions include study concept and design (BS, BM, FP), treatment protocols and materials (BS, BM, FP), trial coordination and therapy sessions (BS, VB), structural MRI and lesion overlay maps (FD, GL), statistical analyses (BS), manuscript drafting and artwork (BS), and revisions (BS, BM, VB, FD, GL, FP).

Funding The current trial was supported by the Deutsche Forschungsgemeinschaft (Pu 97/15-1 to FP) and the Deutsche Akademische Austauschdienst (fellowship to GL).

Competing interests None declared.

Ethics approval The trial was approved by the ethics review board at the Charité University Hospital in Berlin, Germany, with written informed consent obtained from all patients.

Provenance and peer review Not commissioned; externally peer reviewed.

Open Access This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>

© Article author(s) (or their employer(s) unless otherwise stated in the text of the article) 2017. All rights reserved. No commercial use is permitted unless otherwise expressly granted.

REFERENCES

- Breitenstein C, Gewe T, Flöel A, et al. Intensive speech and language therapy improves everyday communication in chronic post-stroke aphasia: a prospective randomized controlled healthcare trial. *The Lancet* 2017;389:1528–38.
- Cherney LR, Patterson JP, Raymer A, et al. Evidence-based systematic review: effects of intensity of treatment and Constraint-Induced Language Therapy for individuals with stroke-induced aphasia. *J Speech Lang Hear Res* 2008;51:1282–99.
- Bhogal SK, Teasell R, Speechley M. Intensity of aphasia therapy, impact on recovery. *Stroke* 2003;34:987–93.

- 4 Allen L, Mehta S, McClure JA, *et al.* Therapeutic interventions for aphasia initiated more than six months post stroke: a review of the evidence. *Top Stroke Rehabil* 2012;19:523–35.
- 5 Brady MC, Kelly H, Godwin J, *et al.* Speech and language therapy for aphasia following stroke. *Cochrane Database Syst Rev* 2016;6:CD000425.
- 6 Pulvermüller F, Neininger B, Elbert T, *et al.* Constraint-induced therapy of chronic aphasia after stroke. *Stroke* 2001;32:1621–6.
- 7 Dignam J, Copland D, McKinnon E, *et al.* Intensive versus distributed aphasia therapy: a nonrandomized, parallel-group, dosage-controlled study. *Stroke* 2015;46:2206–11.
- 8 Robey RR. A meta-analysis of clinical outcomes in the treatment of aphasia. *J Speech Lang Hear Res* 1998;41:172–87.
- 9 Huber W, Poeck K, Willmes K. The Aachen Aphasia Test. *Adv Neurol* 1984;42:291–303.
- 10 Oldfield RC. The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 1971;9:97–113.
- 11 Kessels RP, van Zandvoort MJ, Postma A, *et al.* The Corsi Block-Tapping Task: standardization and normative data. *Appl Neuropsychol* 2000;7:252–8.
- 12 Stahl B, Mohr B, Dreyer FR, *et al.* Using language for social interaction: communication mechanisms promote recovery from chronic non-fluent aphasia. *Cortex* 2016;85:90–9.
- 13 Faul F, Erdfelder E, Buchner A, *et al.* Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. *Behav Res Methods* 2009;41:1149–60.
- 14 Difrancesco S, Pulvermüller F, Mohr B. Intensive Language-Action Therapy (ILAT): the methods. *Aphasiology* 2012;26:1317–51.
- 15 Rorden C, Brett M. Stereotaxic display of brain lesions. *Behav Neurol* 2000;12:191–200.
- 16 Stahl B, Mohr B, Dreyer FR, *et al.* Communicative-pragmatic assessment is sensitive and time-effective in measuring the outcome of aphasia therapy. *Front Hum Neurosci* 2017;11:223.
- 17 Hebb DO. *The organization of behavior*. New York, NY: Wiley, 1949.
- 18 Berthier ML, Pulvermüller F. Neuroscience insights improve neurorehabilitation of poststroke aphasia. *Nat Rev Neurol* 2011;7:86–97.
- 19 Mazur JE. *Learning and behavior*. New York, NY: Routledge, 2016.
- 20 Breitenstein C, Korsukewitz C, Baumgärtner A, *et al.* L-dopa does not add to the success of high-intensity language training in aphasia. *Restor Neurol Neurosci* 2015;33:115–20.
- 21 Mozeiko J, Coelho CA, Myers EB. The role of intensity in Constraint-Induced Language Therapy for people with chronic aphasia. *Aphasiology* 2016;30:339–63.
- 22 Loftus GR, Masson ME. Using confidence intervals in within-subject designs. *Psychon Bull Rev* 1994;1:476–90.

Online Supplement

Efficacy of intensive aphasia therapy in patients with chronic stroke: A randomised controlled trial

**Benjamin Stahl,^{1,2,3,4} Bettina Mohr,⁵ Verena Büscher,⁶ Felix R. Dreyer,⁶
Guglielmo Lucchese^{2,6} and Friedemann Pulvermüller^{6,7}**

¹ Department of Neurology, Charité Universitätsmedizin Berlin, Campus Mitte, Berlin, Germany

² Department of Neurology, Universitätsmedizin Greifswald, Greifswald, Germany

³ Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

⁴ Psychologische Hochschule Berlin, Berlin, Germany

⁵ Department of Psychiatry, Charité Universitätsmedizin Berlin, Campus Benjamin Franklin, Berlin, Germany

⁶ Brain Language Laboratory, Department of Philosophy and Humanities, WE4, Freie Universität Berlin, Berlin, Germany

⁷ Berlin School of Mind and Brain, Humboldt-Universität zu Berlin, Berlin, Germany

Correspondence:

Benjamin Stahl, PhD
Charité Universitätsmedizin Berlin
Charitéplatz 1
10117 Berlin, Germany
Phone: +49 30 450 739 709
Email: *benjamin.stahl@charite.de*

Table I. Patient histories

Patient	Gender	Age (in years)	Education level (in years)	Months after onset of stroke	Left-hemisphere origin
01	Male	74	17	29	MCA ischemia
02	Male	70	13	30	MCA ischemia
03	Male	76	9	76	MCA ischemia
04	Female	58	18	146	MCA ischemia
05	Male	61	19	19	MCA ischemia
06	Female	33	13	12	MCA ischemia
07	Female	82	13	197	MCA ischemia
08	Female	44	12	17	MCA ischemia
09	Female	36	12	27	MCA ischemia
10	Male	64	17	41	MCA ischemia
11	Female	56	21	12	MCA ischemia
12	Male	48	13	12	MCA ischemia
13	Male	67	16	75	BG hemorrhage
14	Male	57	21	74	MCA ischemia
15	Female	51	13	178	MCA ischemia
Group I					
Mean (SD)		58.5 (14.3)	15.1 (3.6)	63.0 (62.3)	
16	Male	67	19	27	MCA ischemia
17	Female	51	18	23	MCA ischemia
18	Female	77	9	23	MCA ischemia
19	Male	74	19	24	MCA ischemia
20	Male	84	11	53	MCA ischemia
21	Female	60	12	253	MCA ischemia
22	Female	41	12	95	MCA ischemia
23	Female	39	11	45	MCA ischemia
24	Male	34	14	54	MCA ischemia
25	Male	65	13	51	MCA ischemia
26	Male	75	16	209	MCA ischemia
27	Male	45	13	23	MCA ischemia
28	Female	58	17	12	MCA ischemia
29	Male	76	10	75	MCA ischemia
30	Male	81	13	44	MCA ischemia
Group II					
Mean (SD)		61.8 (16.5)	13.8 (3.3)	67.4 (70.5)	

Patients are listed according to training intensity: Group I (4 hours per day), and Group II (2 hours per day).

MCA: Middle cerebral artery; BG: Basal ganglia; SD: Standard deviation

Table II. Baseline test scores

Patient	AAT Token Test	AAT Repetition	AAT Naming	AAT Comprehension	CBTT	Aphasia type
01	43	50	42	40	4	Wernicke
02	52	45	49	52	7	Broca
03	63	45	53	52	4	Broca
04	66	59	61	59	5	Broca
05	41	50	38	46	4	Global
06	49	62	53	63	6	Broca
07	47	49	48	46	4	Broca
08	42	55	47	38	4	Broca
09	33	42	42	39	7	Broca
10	42	45	39	37	5	Global
11	49	54	49	37	7	Broca
12	48	51	47	39	6	Broca
13	65	63	62	66	4	Broca
14	54	51	59	64	7	Broca
15	46	50	50	55	6	Broca
Group I						
Mean (SD)	49.3 (9.4)	51.4 (6.3)	49.3 (7.4)	48.9 (10.6)	5.3 (1.3)	
16	46	58	49	55	5	Broca
17	42	43	46	39	5	Amnesic
18	49	41	48	47	6	Broca
19	54	62	61	63	5	Broca
20	58	50	60	50	4	Broca
21	55	59	63	64	5	Broca
22	47	48	51	39	6	Broca
23	59	56	59	78	7	Broca
24	57	59	68	78	7	Broca
25	52	57	58	51	7	Broca
26	46	44	41	36	6	Broca
27	33	48	43	38	6	Global
28	43	44	40	51	6	Global
29	42	42	34	37	3	Global
30	69	54	56	54	6	Broca
Group II						
Mean (SD)	50.1 (8.9)	51.0 (7.2)	51.8 (9.8)	52.0 (13.8)	5.6 (1.1)	

Patients are listed according to training intensity: Group I (4 hours per day), and Group II (2 hours per day).

AAT: Aachen Aphasia Test¹; CBTT: Corsi Block-Tapping Task²; SD: Standard deviation

References

- 1 Huber W, Poeck K, Willmes K. The Aachen Aphasia Test. *Adv Neurol* 1984;42:291–303.
- 2 Kessels RP, van Zandvoort MJ, Postma A, et al. The Corsi BlockTapping Task: standardization and normative data. *Appl Neuropsychol* 2000;7:252–8.