Knoph, Simonsen & Lind (2017). Cross-linguistic transfer effects of verb-production therapy in two cases of multilingual aphasia. Aphasiology *31*(12), 1482–1509.

Cross-linguistic transfer effects of verb-production therapy in two cases of multilingual aphasia*

Monica I. Norvik Knoph[1,2], Hanne Gram Simonsen[1], Marianne Lind [1,2]

[1] MultiLing Center for Multilingualism in Society across the Lifespan (CoE), Department of Linguistics and Scandinavian studies, University of Oslo, Norway
[2] Statped, Department of Speech and Language Disorders, Oslo, Norway

* We thank PN and DT for taking part in this study, and Ingvild Røste for conducting the treatments. We are also grateful to Pernille Hansen for help with the statistics.

This project was funded by the Norwegian ExtraFoundation for Health and Rehabilitation through EXTRA funds (2011/2/0279) and partly supported by the Research Council of Norway through its Centres of Excellence funding scheme, project number 223265.

Address for correspondence:

Monica Norvik Knoph

MultiLing Center for Multilingualism in Society across the Lifespan (CoE)
Department of Linguistics and Scandinavian Studies
University of Oslo
P.O.Box 1102 Blindern
0317 Oslo
Norway
E-mail: monica.knoph@iln.uio.no

Abstract

Background: Verb retrieval is challenging for monolingual and multilingual speakers with aphasia. Previous research on bilingual aphasia shows equivocal results of cross-linguistic transfer and inhibition.

Aims: This study explores the impact of verb-production treatment in the treated and untreated languages of two bilingual speakers with aphasia. The main goals were to explore treatment effects, possible cross-linguistic transfer effects, and to investigate possible inhibition of the untreated languages.

Methods & Procedures: The participants were one trilingual speaker (Portuguese-Ronga-Norwegian) with nonfluent aphasia and one bilingual speaker (English-Norwegian) with fluent aphasia. They received two types of treatment: communication based therapy and Semantic Feature Analysis. Treatment was conducted in Norwegian, a late-acquired language for both speakers. Treatment effects were measured in action naming tasks and narrative tasks in the treated language as well as the untreated languages.

Outcomes & Results: Overall, the participants responded positively to the verb production treatments. This was demonstrated at the lexical level and in discourse production, especially in the treated, but also in the untreated languages. No inhibition of the untreated languages was found.

Conclusions: The data provide evidence for positive effects of verb-retrieval treatment provided in sentence contexts in a late-learned weaker language of multilingual speakers with aphasia. The treatments did not lead to an unwanted inhibition of the untreated language, which is an important finding for clinicians as well as for researchers. The results provide evidence for a shared conceptual network of the languages in bilingual speakers, supporting current models of bilingual language processing.

Keywords: bilingual; aphasia; cross-linguistic transfer; inhibition; verb treatment

Anomia is one of the core symptoms of aphasia and especially challenging is verb retrieval, both for monolingual (e.g. Berndt, Haendiges, & Wozniak, 1997; Mätzig, Druks, Masterson, & Vigliocco, 2009) and multilingual speakers with aphasia (Ansaldo, Ghazi Saidi, & Ruiz, 2010; Faroqi-Shah & Waked, 2010; Kambanaros & van Steenbrugge, 2006). The present study explores the impact of verb-production treatment in the treated and untreated languages of one trilingual speaker (Portuguese-Ronga-Norwegian) with nonfluent aphasia and one bilingual speaker (English-Norwegian) with fluent aphasia. Treatment was conducted in Norwegian, a late-acquired language for both speakers, and treatment effects were measured in restricted (formal language testing) and less restricted (narrative[2] production) tasks.

Cross-linguistic transfer

The overall goal in aphasia therapy of multilingual speakers with aphasia is improved communication abilities in both or all languages needed for participation in meaningful life activities (Kohnert, 2009). Consequently, cross-linguistic transfer is often an aim, since treatment of all the languages of the client is rarely achievable (Roger & Code, 2011). Several studies have reported transfer to the untreated language(s) when treatment was provided in a weaker language (e.g. Edmonds & Kiran, 2004, 2006; Knoph, 2013; Knoph, Lind, & Simonsen, 2015; Kurland & Falcon, 2011; Marangolo, Rizzi, Peran, Piras, & Sabatini, 2009). Moreover, it has been suggested that semantic treatment facilitates cross-linguistic transfer more than phonological treatment does (Croft, Marshall, Pring, & Hardwick, 2011; Edmonds & Kiran, 2004; Kiran & Iakupova, 2011; Kurland & Falcon, 2011). Cross-linguistic transfer is attainable on the assumption of a shared conceptual system of all the languages of a bilingual[1] speaker, as proposed by current models of multilingual language processing (e.g. de Bot, 2004; Kroll & Stewart, 1994; Pavlenko, 2009), and the idea is that the less proficient L2 relies upon the stronger L1.

There is an overall agreement regarding parallel activation of the languages of a multilingual speaker. Studies have shown that, at least at the lexical level all, languages are active when a person is speaking in one of the languages (e.g. Brysbaert & Duyck, 2010; de Bot, 1992; Green, 1998; Kroll, Bobb, & Wodniecka, 2006; Kroll, Dussias, Bice, & Perrotti, 2015; Kroll, Van Hell, Tokowicz, & Green, 2010; Pavlenko, 2009). Several models of multilingual language processing agree that multilinguals have a shared conceptual system for both or all of their languages, with varying connections to separate lexical stores for the individual languages (e.g. Kroll & Stewart, 1994; Pavlenko, 2009). Also, it is well established that the languages of a multilingual speaker mutually influence one another, not only at the lexical level (e.g. Baus, Costa, & Carreiras, 2013; Linck, Kroll, & Sunderman, 2009), but also at the phonological (e.g. Chang, 2013) and syntactic levels (e.g. Dussias & Nuria, 2007; see also Kroll et al., 2015 for a discussion).

While some researchers suggest that structural differences between languages may play a role regarding cross-linguistic transfer (e.g. Goral, Levy, & Kastl, 2010) Ansaldo and Ghazi Saidi (2014) found no evidence of this in their review. On the contrary, a range of studies have shown that treatment effects can transfer across languages, regardless of which family the languages in question belong to, at least within the Indo-European family of languages (Croft et al., 2011; Goral et al., 2010; Kiran & Iakupova, 2011; Kohnert, 2004; Miertsch, Meisel, & Isel, 2009).

Inhibition of untreated languages

A possible unwanted effect of language treatment in one of the languages of a multilingual speaker with aphasia is inhibition, or decline, of the untreated languages. Whilst evidence of cross-linguistic transfer exists, so too does evidence of inhibition of a stronger language when treatment is provided in a postmorbidly weaker language (e.g. Goral, 2012; Goral,

Naghibolhosseini, & Conner, 2013). Goral (2012) reports on a decrease in sentence grammaticality in picture descriptions tasks in the untreated stronger languages of three participants. Also Goral et al. (2013) found decline in the production of complete sentences, more noun–verb agreement errors and more omissions of auxiliary verbs in English, following treatment in the weaker Persian of their participant. In both studies the decrease on the different measures post treatment is interpreted as an inhibition of the untreated languages.

This inhibition is explained by an asymmetric switching cost. That is, when an individual is speaking in one language, the other language needs to be inhibited. Inhibiting a stronger language is more demanding than inhibiting a weaker one, due to a presumably stronger activation of the stronger language in the first place (Costa & Santesteban, 2004; Green, 1998; Meuter & Allport, 1999). However, this question needs further examination, as reviews of therapy studies for multilingual speakers with aphasia have found no clear evidence that therapy in one language harms the untreated languages (Faroqi-Shah, Frymark, Mullen, & Wang, 2010; Kohnert, 2009).

Relevant verb-treatment studies in multilingual aphasia

There are mixed findings in the literature regarding generalisation – or transfer – from a treated language to an untreated language, with evidence of generalisation (e.g. Kiran, Sandberg, Gray, Ascenso, & Kester, 2013; Knoph et al., 2015) as well as evidence of a lack of generalisation (e.g. Croft et al., 2011; Filiputti, Tavano, Vorano, De Luca, & Fabbro, 2002; Miertsch et al., 2009; Miller Amberber, 2011). Many of the published studies focus on noun production. Although there are findings of impaired verb retrieval in multilingual speakers with aphasia (e.g. Hernández, Costa, Sebastián-Gallés, Juncadella, & Reñé, 2007; Kambanaros, 2008, 2010; Kambanaros & van Steenbrugge, 2006; Miozzo, Costa, Hernández, & Rapp, 2010; Mätzig et al., 2009; Weekes & Raman, 2008), treatment studies involving verbs are still relatively scarce

in this field, compared to studies involving nouns (Kambanaros, 2016; Kiran & Bassetto, 2008; Nickels, 2002). In this paper, we will thus focus on verb retrieval, given the smaller amount of studies of verb retrieval and the verb's crucial role in communication. An extensive review on verb treatment-studies for multilingual speakers with aphasia is beyond the scope of this paper.

Goral and colleagues treated a trilingual speaker (Hebrew, English, and French) with nonfluent aphasia in his L2 (English), which was also the language of the environment, in two different studies (Altman, Goral, & Levy, 2012; Goral et al., 2010). The participant had high levels of proficiency in all three languages. Both of the studies targeted language production using complete sentences, not focusing specifically on verb production. In both studies, treatment was provided with low intensity, with a total of 18 hours of therapy. Various tasks were used, e.g. picture description, sentence elicitation, exchange of information, and the participant was encouraged to produce complete sentences. The treatment was divided into two phases, the first focusing on morphosyntactic forms and the second on word-finding strategies. Altman et al. (2012) found statistically significant positive changes to varying degrees in narrative structure (e.g. *Background* and *Setting*, e.g. the introduction of participants), sentence grammaticality (e.g. amount of grammatically simple and complex sentences), and lexical variables (e.g. amount of types and tokens of verbs and nouns) in all the languages, including the L1 (Hebrew). The changes were especially evident following the treatment focusing on morphosyntactic forms. The authors suggest that the participant's improved grammatical production induced more variation in the types of sentences and the content he included in his stories and that this lead to a generalisation of the skills from sentence production, which was targeted in therapy, to narrative production, which was not directly trained during therapy. Also other aspects that were not directly trained, such as grammatical complexity and narrative scheme structure improved following therapy.

In contradiction to the above mentioned findings, little transfer was found to the L1 (Hebrew) in the study by Goral et al. (2010), although improvements in both the treated language (English) and the untreated L3 (French) were found following treatments targeting morphosyntactic constructions and language production rate. The lack of generalisation to the L1 was explained by ceiling performance in this language, as well as a possible differential representation and processing of the treated L2 (English) and the untreated L1 due to structural differences between the two languages. Moreover, the authors suggest that the L1, being the first-acquired language, may have a different mental representation from the other languages. In addition, they assume that, since the L1 was the more-proficient language, this may have affected the occurrence of cross-linguistic transfer to this language. The different results of these two studies, who involve the same participant, point to different focus of analysis. Whereas Altman and colleagues (2012) found transfer to the untreated L1 with the measures used in their study, Goral et al. (2010) failed to find cross-linguistic transfer to the untreated L1.

A similar pattern of a lack of transfer to the untreated L1 was found by Miertsch et al. (2009) and by Knoph et al. (2015). Miertsch et al. (2009) provided therapy in the L3 (French) – not the language of the environment, but a language the participant had a strong emotional relationship to – of a trilingual speaker (German, English, and French) with Wernicke's aphasia. French was neither the strongest nor the weakest language of the participant; he had medium proficiency in this language. The treatment was provided with high intensity, with two sessions every day for 3.5 weeks, and the overall aim was to target the participant's lexical-semantic deficits. Specifically, the treatment consisted of exercises with prepositions, semantic-conceptual relationships between words, and word finding of verbs and nouns in a discourse context. Following 3.5 weeks of treatment, the participant showed significant gains in the treated L3, as well as cross-linguistic transfer to the L2 (English), but not to the L1 (German).

The findings were explained with the participant's close-to-ceiling performance in the L1. In addition, the treatment was provided over a short time span (23 sessions over 3.5 weeks), and the authors suggest that this may not have been long enough to lead to significant improvements in a language with an already stable linguistic performance.

Knoph et al. (2015) investigated transfer effects following treatment in the language of the environment, the weaker Norwegian (L4), of a quadrilingual speaker (Japanese, English, German, and Norwegian) with nonfluent aphasia. The treatment focused on verb production in sentence contexts using Semantic Feature Analysis (SFA), and was provided an intensive manner, with 29 sessions over 2.5 weeks. Following treatment, the naming of trained verbs increased significantly, but there was no improvement of untrained verbs, semantics or syntax in the treated language (Norwegian). However, the participant showed improvements in narrative production in this language, with a significant increase in the types and tokens of verbs and in the proportion of complete and complex sentences. Cross-linguistic transfer at the lexical level and in semantics and syntax was found to the L3 (German) and partially to the L2 (English). Improvements in discourse production were also evident in these two languages. No changes in the L1 (Japanese) were found. Also in this study, the authors argued that the lack of transfer to Japanese could be a result of an already high (close-to-ceiling) proficiency in this language. Additionally, Japanese is the structurally most dissimilar language to Norwegian of all the untreated languages.

An alternative paradigm is to provide treatment in the L1 of the participants. Goral, Rosas, Conner, Maul, and Obler (2012) and Ansaldo et al. (2010) provided treatment in the L1 of their participants. Goral et al. (2012) provided noun and verb retrieval treatment to examine cross-linguistic transfer in a multilingual speaker (Catalan, Spanish, German, French, and English) with nonfluent aphasia. Treatment was administered not only in his strongest language, that is, Spanish (L1), but also in his weakest language, English (L4), which was also the

language of the environment. The treatments were provided in a relatively high-frequent manner, with 22–25 hours over a period of 5 weeks, for each of the languages. The aim was to enhance verb and noun production, and exercises such as a modified version of SFA, sentence production tasks, and a repetition task were carried out. Overall, the treatment in English (L4) resulted in statistically significant improvements in object and action naming in this language, and to cross-linguistic transfer to two of the untreated languages (German and French), where he produced significantly more words in the fluency task and longer narratives. Treatment provided in Spanish (L1), on the other hand, led no significant changes in Spanish and a limited degree of generalisation to the untreated languages. Significant changes in the Verbal Fluency task for French and a significantly greater number of clauses per narrative across all languages were found; however, only in the German narrative, the amount of grammatical clauses increased significantly. The authors suggest that treating the language of the environment, which was also his weakest language (English), may have contributed to the findings. Also Ansaldo et al. (2010) provided treatment in the language of the environment, however, in this case, it was the participant's L1, which also was the strongest language. The participant was a highly proficient speaker of two languages (Spanish and English) with aphasia, for whom verb and noun production treatment was provided in Spanish (L1). Two different treatment protocols were used interchangeably, Switch Back Through Translation and a modified form of SFA, to increase activation of target semantic features. The treatment was provided during a low intensive schedule, with two one-hour sessions a week for three months. The results showed significant improvement on naming of trained nouns and verbs in Spanish. Regarding cross-linguistic transfer to English, no significant transfer was found.

Summing up, the findings from all of these studies indicate that treating a later-acquired language may be beneficial in enhancing cross-linguistic transfer, rather than providing treatment in L1 (Altman et al., 2012; Goral et al., 2010; Goral et al., 2012; Knoph et al., 2015;

Miertsch et al., 2009) , in addition to providing treatment in the language of the environment (Altman et al., 2012; e.g. Ansaldo et al., 2010; Goral et al., 2012; Knoph et al., 2015).

Research questions

The objective of the present study was to investigate the effect of treatment targeting verb production in sentence contexts in two multilingual speakers with aphasia. An important aim was to optimise the possibility of treatment related cross-linguistic transfer, rather than comparing the effect of specific treatment methods. Hence, two treatments with previously reported findings of improved verb retrieval, improved narrative production and also across-language gains were chosen: *communication-based treatment* (CBT) (e.g. Kempler & Goral, 2011) and *Semantic Feature Analysis* (SFA) (e.g. Wambaugh & Ferguson, 2007). Given previous findings of positive effects on discourse production following verb treatment both in monolingual and in multilingual aphasia (e.g. Bastiaanse, Hurkmans, & Links, 2006; Goral et al., 2012; Wambaugh, Mauszycki, & Wright, 2014), one of the aims of the study was improvement of connected speech. Due to the presumably shared conceptual networks between the languages of multilingual individuals, transfer to the same linguistic levels in the untreated languages was anticipated.

The study aimed to answer the following research questions:

1. Related to action naming in sentences:
    a. Does CBT targeting verbs in sentence context lead to improvement of verbs in the treated language?
    b. Does CBT targeting verbs in sentence context lead to transfer, i.e. improvement of verbs in the untreated language?
    c. Does SFA targeting verbs in sentence context lead to improvement of trained verbs in the treated language?

d. Does SFA targeting verbs in sentence context lead to improvement of untrained verbs in the treated language?

e. Does verb-retrieval treatment (i.e. a combination of CBT and SFA) lead to transfer, i.e. improvement of verbs in the untreated language?

2. Related to narrative production:

a. Does verb-retrieval treatment (i.e. a combination of CBT and SFA) lead to improvement of narrative production in the treated language?

b. Does verb-retrieval treatment (i.e. a combination of CBT and SFA) lead to transfer, i.e. improvement of narrative production in the untreated language?

3. Is there evidence of improved verb retrieval in the treated language causing inhibition of the untreated, dominant language?

Method

A multiple-baseline across participants design with five phases was applied: 1) three baselines over two weeks, 2) pre-treatment measures, 3) verb-retrieval treatment in the L2, in addition to weekly generalisation measures and control measures, 4) post-treatment measures, and 5) maintenance measures (for a timeline of assessments and treatments, see Table 1).

Table 1. Timeline of assessments and treatment

Timeframe	**Activity**	**Content**
Week 1	3 baselines for each language	Action-naming test and narrative production
Weeks 2–3	No intervention	
Weeks 4–5	Pre-test, both languages	BAT[1], action-naming test and narrative production, social validation[2]

Weeks 6–9	Treatment + weekly control tests	Ca. 3 sessions of 2-4 hours per week, a total of 22–25 hours + PALPA 8 and 20 untrained verbs
Weeks 10–11	Post-test 1 (no intervention)	BAT[1], action-naming test and narrative production[3]
Weeks 12–15	Treatment + weekly control tests	Ca. 3 sessions of 2-4 hours per week, a total of 22–25 hours + PALPA 8 and 20 untrained verbs
Weeks 16	Post-test 2	BAT[1], action-naming test and narrative production[3]
3-4 months post treatment	Follow-up	BAT[1], action-naming test and narrative production

[1] *Bilingual Aphasia Test (Paradis & Libben, 1987). Will not be reported on.*
[2] *Social validation: Norwegian versions of the CETI (Lomas et al., 2006) & the SALK-39 (Berg, Haaland-Johansen, & Hilari, 2010). Will not be reported on.*
[3] *The post-CBT narrative will not be reported on.*

Participants

Two multilingual speakers with aphasia participated in the study, both had Norwegian as a late-acquired language after the age of 20. PN is a 50-year-old, right-handed female who grew up as a simultaneous bilingual with Portuguese and Ronga[3] in Mozambique. She moved to Norway as an adult and learned Norwegian both by immersion and through language classes. She suffered a left hemisphere stroke ten months prior to the intervention, which caused a moderate nonfluent aphasia of the Broca type[4]. She reported that she used all three of her languages on a daily basis both prior to and after the stroke. The premorbid proficiency level was reported as high for all the languages[5]. Because of fatigue, the assessment tools for this participant were kept to a minimum. Based on the results of a modified *Bilingual Aphasia Test* (BAT) (Paradis & Libben, 1987), Norwegian was the weakest language post-stroke (Appendix 1).

DT is a 75-year-old bilingual, right-handed female. She grew up in Scotland and spoke only English in her childhood. DT learned Latin and German at school, but she rarely used these languages in her adult life. She moved to Norway as an adult, and acquired Norwegian

by immersion. Pre-stroke, she used both English and Norwegian with her husband, and exclusively English with her children and grandchildren. She sustained a left hemisphere stroke causing a moderate to severe fluent aphasia of the Wernicke type, about 18 months prior to the study. This participant also experienced fatigue; hence, the assessment tools were kept to a minimum. English was her strongest language both prior to and following the stroke, according to the BAT (Paradis & Libben, 1987).

Procedures

Each participant received a total of 40–50 hours of treatment over 4–6 weeks. The two different treatment protocols (CBT and SFA) were administered sequentially in an intensive schedule with 20–25 hours each, amounting to approximately 10 hours a week. The order of the treatments was the same for the participants. Each treatment block lasted for 2–3 weeks with a break of two weeks with no treatment in between. The period between post-treatment and follow-up was 13–17 weeks and during this period, no treatment was provided.

Both treatment protocols – CBT (e.g. Kempler & Goral, 2011) and SFA (Wambaugh & Ferguson, 2007) – focused on production of verbs in sentence contexts, and were provided in Norwegian, a late-acquired language for both of the participants. A trained SLT, with extensive experience of aphasia therapy, conducted the therapy with each of the participants. Consent was obtained from both participants prior to the study, and the Norwegian Social Science Data Services (NSD) approved the ethical standards of the project.

The CBT involved no pre-selected verbs and any relevant verb to describe the pictures was approved of. The treatment included duplicate picture materials, and the client and the SLT had a cardholder each, so that they could not look at each other's pictures. They took turns in describing and guessing the pictures (e.g. 'Do you have *a girl who is sitting*?' 'Yes, I have *a girl who is sitting*.'). The aim of the procedure was to produce complete sentences to describe

different pictures, in order for the interlocutor to identify the pictures. The method has certain similarities to *Promoting Aphasics' Communication Effectiveness* (PACE) (Davis & Wilcox, 1985), including several of the four PACE principles. The SLT and the client participate equally as sender and receiver of the information, as the dynamic exchange of new information between the SLT and participant is important. The method furthermore includes natural feedback provision. Contrary to PACE, the participants were encouraged (although not constrained) to use spoken language as the means of communication. The method thus corresponds to *intensive language-action therapy*[6] (ILAT) (Difranscesco, Pulvermüller, & Mohr, 2012). The CBT includes also other principles of ILAT, namely high intensity treatment with massed practice and action-embedded language use relevant for daily life. Different versions of ILAT and PACE have shown to be fruitful for improving language in speakers with aphasia (e.g. Davis, 2005; Goral & Kempler, 2009; Kirmess & Lind, 2011; Kirmess & Maher, 2010). The SLT consistently modelled correct sentence structure (e.g. when describing the selected card the SLT would produce simple, complete sentences such as *The man is fishing*), and if needed the participants were reminded of the sentence structure being used during treatment. When sentences were produced incorrectly, that is, if the client for instance used a verb not appropriate in the context (e.g. *eating* instead of *drinking*), the SLT would prompt or suggest a response. The prompts could be either semantic or phonological. Grammatical errors, such wrong or missing inflections, were not corrected. The participants were encouraged to produce simple, but complete sentences containing verbs. In this method, any verb in a sentence that exchanges relevant information is accepted (Kempler & Goral, 2011).

The other treatment method, SFA, is an approach that aims at enhancing lexical retrieval by improving the access to the semantic network of the speaker (e.g. Boyle & Coelho, 1995; Wambaugh & Ferguson, 2007). The method was initially developed for (single word) noun production, where it has shown promising results in both monolingual and multilingual

speakers with aphasia (e.g. Boyle, 2004, 2010; Boyle & Coelho, 1995; Coelho, McHugh, & Boyle, 2000; Edmonds & Kiran, 2006; Kiran & Edmonds, 2004; Lowell, Beeson, & Holland, 1995). SFA has also been adapted to treatment of verb retrieval in studies of monolingual speakers (e.g. Wambaugh & Ferguson, 2007; Wambaugh et al., 2014), and and in at least one study with a multilingual speaker with aphasia (Knoph et al., 2015).

In the present study, the procedure was adapted from Wambaugh and colleagues (2007; 2014) (cf. also Knoph et al., 2015). The semantic features used were mostly related to argument structure and semantic roles and included *agent/experiencer*, *theme/patient*, *purpose, location, instrument,* and *association*. The semantic features were triggered by asking the following questions: 'Who usually does this?' (agent), 'What is it done to?' (theme), 'Why does this happen?' (purpose), 'Where does this action happen?' (location), 'What part of the body or what tool is used to make this happen?' (instrument), and 'What does it make you think of?' (association).

For PN the procedure was carried out in line with other SFA studies (e.g. Knoph et al., 2015; Wambaugh & Ferguson, 2007; Wambaugh et al., 2014) in the following fashion: A picture of the target action was placed in the centre of a diagram (Appendix 2), and the participant was asked to produce a relevant verb referring to the illustration. Whether she managed to do this or not, the SLT guided her through each of the six semantic features, one at a time. This was done by asking each of the questions above, to elicit information about each of the six semantic features. PN's responses were written down on the diagram. When she could not access an appropriate feature, the SLT would suggest a response, or semantically or phonologically prompt a possible feature. PN was then requested to name the target verb. If she could not produce the verb after naming all the features, the SLT would provide the target verb for her to repeat. Finally, she was asked to produce a simple sentence with the target verb. In the cases where she did not succeed in producing a correct sentence, the SLT would suggest a

simple sentence orally, which PN repeated. Positive feedback was provided, and the next picture was presented.

Initially, the method was applied in the same way for DT as well. However, it soon became apparent that she did not fully grasp the rationale behind this treatment method, probably due to her difficulties with auditory comprehension. She could not generate the required semantic features. Her approach was rather to describe the pictures in a relatively detailed manner. Thus, the treatment procedure was modified. Instead of requesting her to produce the six predetermined semantic features, any three relevant features produced by DT for each picture were accepted. Frequently she would produce irrelevant features, and the SLT would suggest relevant features to the pictures. She was also requested to produce the target verb and a simple sentence to describe the picture. Often, she did not succeed in producing a sentence, and the SLT would suggest a simple sentence for her to repeat. She did not always manage to repeat these sentences, which was not surprising, given her fluent aphasia of the Wernicke type (Potagas, Kasselimis, & Evdokimidis, 2013).

For each of the participants, the words for the SFA treatment were selected based on their individual performances on the Norwegian baselines. The verbs tested at baselines and also pre-post treatments were selected from a shortened Norwegian version of the *Newcastle University Aphasia Treatment Resources* (NATR), comprising 50 verbs (the full set consists of 120 action pictures) (Morris, Webster, Whitworth, & Howard, 2012). The verbs in the NATR are all everyday words with high naming agreement in both the English (Morris, Webster, Whitworth, & Howard, 2009) and the Norwegian version (Morris et al., 2012). The verbs in the action-naming test are not matched for relevant psycholinguistic variables that can affect the performance of speakers with aphasia, like frequency, imageability, and number of syllables (Whitworth, Webster, & Howard, 2014). However, a post-hoc analysis of the verbs by means of two-tailed two-sample Wilcoxon tests showed that there were no significant differences

between each of the participant's trained and untrained verbs of any of the variables, except one: PN's trained verbs were significantly less frequent than the untrained verbs ($W = 102$, $p = .014$).

Verbs that were produced correctly at the baselines were not included in the treatment set. PN was able to produce 26 verbs spontaneously twice or thrice in the baselines. The remaining 24 verbs were selected for training. A supplementary set of 22 verbs randomly selected from the Norwegian version of the NATR (Morris et al., 2012) were included for treatment, to yield a larger practice set of 46 verbs. However, due to PN's fatigue, it was decided to minimize the burden of testing; hence, only the 24 targeted items were tested before and after each treatment block. A further 20 verbs were selected from the Norwegian version of the NATR (Morris et al., 2012) to serve as untrained controls. These were not tested at baseline, but they were assessed weekly. DT did not produce any verbs at all three baselines, and only eight verbs were named twice. For her, due to her more severe language impairments, a smaller set of 30 verbs was selected for treatment, while 20 were chosen to serve as untrained controls which were not tested at baseline, but assessed weekly. All the verbs were chosen from the Norwegian version of the NATR (Morris et al., 2012).

Pre-treatment measures and outcome measures

Preceding the intervention, multiple baselines were conducted in all the languages where assessment materials were available for both participants (i.e. all the languages except Ronga). At each baseline all the languages were assessed on the same day. The baselines consisted of an action-naming test, where the aim was to produce a simple sentence for each of the 50 pictures. Pictures depicting actions from the English version of the NATR were used for assessing English and Portuguese, and the Norwegian version of the NATR was used for Norwegian (Morris et al., 2012). The verbs from the English version of the NATR were

translated into Portuguese by a professional interpreter, as there is no Portuguese version of these resources. In addition, a personal narrative on a different topic was collected each time (see below). The baselines were conducted thrice for each participant (cf. Table 1).

Pre- and post-treatment measures included a modified version of the BAT (Paradis & Libben, 1987), the action-naming test, and a personal narrative, and were conducted in all languages of the participants (except Ronga), in addition to social validation tests, conducted in Norwegian (cf. Table 1). For the narrative elicitation, the participants produced personal narratives (as opposed to a story retell or a cartoon description (e.g. Webster, Whitworth, & Morris, 2015)) about a movie they had seen or a book they had read, or a happy moment or a vacation (cf. e.g. Altman et al., 2012; Goral & Kempler, 2009; Kempler & Goral, 2011). They were encouraged to tell a different story at each measurement point, so that they did not practice telling the same story every time. Additionally, two types of control measures, one related and one unrelated were collected weekly during the treatment blocks to ascertain experimental control and to assess the effects of treatment and possible generalisation. These were collected in the language of treatment, and comprised naming of 20 untrained Norwegian verbs and a non-word repetition task in Norwegian (PALPA 8) (Kay, Coltheart, & Lesser, 2009). The test sessions including the narratives were recorded on audio- and videotape.

In this paper we report on the action-naming test and the narratives. We use three measurement points for the narratives: pre-CBT, post-SFA, and at follow-up. The scores from the action-naming test will be reported for four measurement points: pre-CBT, post-CBT, post-SFA, and at follow-up.

Analysis

The *McNemar* test was used to test the statistical significance on the treatment change on the action-naming test at baselines, pre-post each of the treatments, and at follow-up (i.e. pre-CBT–

follow-up). This is a two-tailed non-parametric test, based on chi square statistics, and it is commonly used to measure changes in participants' scores on, for instance, language tests (Field, Miles, & Field, 2012).

For the narratives, we focused on the overall increase in verbal production. Hence, the target category for treatment (verb types and tokens) was analysed in addition to the total number of words produced (including false starts and repetitions), speech rate (words per minute), and the number of utterances. The basic analytical sentence level unit was the *Analysis of Speech Unit* (AS-unit) (Foster, Tonkyn, & Wigglesworth, 2000). The AS-unit is defined as 'a single speaker's utterance consisting of an independent clause, or sub-clausal unit, together with any subordinate clause(s)' (p. 365). In addition, to investigate whether the narratives improved in terms of content, the number of correct information units (CIUs) was counted (cf. e.g. Wambaugh & Ferguson, 2007; Wambaugh et al., 2014). This is a measure of content production in discourse which comprises words that are 'accurate, relevant and informative relative to the eliciting stimuli' (Nicholas & Brookshire, 1993, p. 340). In addition, the percentage of words in the sample that conveyed information appropriate to the content of the narrative was calculated (CIUs/total verbal units). This measure has a good ecological validity, as changes in CIU measures of information content are perceived even by naïve listeners in some studies (Jacobs, 2001; Ross & Wertz, 1999).

Effect size was calculated to compare the pre- and post-intervention scores. The Busk and Serlin's d_1 (1992) is based on within-case variation, and the calculation is done by subtracting the average pre-treatment from the average post-treatment values and dividing them by the standard deviation of the pre-treatment measures. To calculate the standard deviation at least two measurement points pre-intervention are necessary; hence, effect size was calculated for the narrative measures only (where multiple baselines were obtained). Effect sizes greater

than one were taken as significant (e.g. Goral & Kempler, 2009; Goral et al., 2013; Knoph et al., 2015).

Reliability

To ensure the reliability of the testing, the first author briefed the interpreters about conducting the tests and about communicating with people with aphasia. Native speakers or highly skilled speakers with a university degree in the relevant language transcribed the responses on the action-naming test and the narratives orthographically. On the action-naming test, a sentence was scored as correct if it contained the target verb (in any form). Authors MK and ML checked all the transcriptions, apart from the Portuguese ones. For Portuguese, two native speakers of this language transcribed the narratives to ensure the reliability of the transcriptions. Transcription disagreements were discussed and resolved before scoring. In collaboration authors MK and ML segmented a minimum of 20 percent of the transcripts into AS-units to ensure that this was done in the same way for all the transcripts, and the rest was divided between them. For the interrater reliability of the coding of the narratives, authors MK and ML divided the Norwegian and the English transcripts between them and scored them. The interrater agreement varied between 87 percent and 99 percent for the different measures (e.g. AS-units, grammaticality of sentences, nouns and verbs, types and tokens, total verbal units, number and percentage of CIUs). It thus exceeded a coefficient of 0.8, which is considered reliable (Kratochwill et al., 2013; Pring, 2005).

Results

Both participants demonstrated stable baselines on the action-naming test, with no significant changes between the trials (PN: $p = .220$; DT: $p = .220$). The control measures (untrained verbs and nonword repetition) were stable throughout the intervention, and there were no significant

changes between the trials, for either of the participants (untrained verbs: PN: $p = .617$; DT: $p = .450$, nonword repetition: PN: $p = .248$; DT: $p = .343$).

The results are presented in relation to the three main research questions with subquestions presented above.

1) Research questions related to action naming in sentences:

a) Did CBT targeting verbs in sentence context lead to improvement of verbs in the treated language?

Since there are no pre-selected trained verbs in this treatment method, we will report on the total number of verbs assessed with the action-naming test pre-post CBT (cf. Table 2 for all results on the action-naming test). The post-CBT results were collected between the two treatment protocols, thus we cannot report on maintenance of these findings.

We found no significant changes on the action-naming test in the treated language following CBT for either of the participants (PN: $p = .220$; DT: $p = .479$).

Table 2. Action-naming test results pre- and post- both treatments for PN and DT (percentage accuracy and significance measures)

	Pre CBT	**Post CBT**	**Pre SFA**	**Post SFA**	**Follow-up**
PN (Portuguese-Norwegian)					
Portuguese	70 %	82 %	82 %	90 %	88 %
Norwegian, all	52 %	64 %	64 %	86 %	86 %
Norwegian, trained	n/a	n/a	29 %	**79 %*** **	**75 %*** **
Norwegian, untrained	n/a	n/a	96 %	92 %	92 %
DT (English-Norwegian)					
English	13 %	24 %	24 %	18 %	8 %
Norwegian, all	14 %	18 %	18 %	31 %	20 %
Norwegian, trained	n/a	n/a	14 %	**43 % ***	27 %
Norwegian, untrained	n/a	n/a	13 %	19 %	10 %

Significant effects are in bold.
*** $p < .001$; * $p < .05$

b) Did CBT targeting verbs in sentence context lead to transfer, i.e. improvement of verbs in the untreated language?

Following CBT, there were no significant changes on the action-naming test in the untreated languages of the participants (PN: p =.238; DT: p =.332).

c) Did SFA targeting verbs in sentence context lead to improvement of trained verbs in the treated language?

In SFA, there were pre-selected verbs that were trained and other verbs remained untrained. We will therefore distinguish between the trained and the untrained verbs in the following.

For PN, the production of trained verbs increased significantly pre-post SFA treatment (p = .001), and this was maintained at follow-up (p = .001). DT showed an improvement pre-post SFA treatment of the trained verbs (p = .013), but this was not significant at follow-up (p = .134).

d) Did SFA targeting verbs in sentence context lead to improvement of untrained verbs in the treated language?

For PN's untrained verbs, i.e. those she was able to produce during the baseline test and which were therefore not included in therapy, she maintained a high score at subsequent measurement points, with no significant change (p = .480). DT showed a similar pattern, with no significant changes on the untrained verbs in Norwegian (p = 1).

e) Did verb-retrieval treatment (i.e. a combination of CBT and SFA) lead to transfer, i.e. improvement of verbs in the untreated language?

For PN, there was a significant change on the Portuguese action-naming test pre-post both treatments ($p = .034$). This increase was not maintained at follow-up ($p = .052$). For DT, there was no significant change on the English action-naming test ($p = .773$).

2) Research questions related to narrative production:

a) Did verb-retrieval treatment (i.e. a combination of CBT and SFA) lead to improvement of narrative production in the treated language?

<TABLES 3 AND 4>

As the narratives are measures on more general language production, and not specific to trained and untrained words the results will be presented pre-CBT–post-SFA. This also allows us to present effect sizes based on these measures. All the results from the narratives can be found in Table 3 for PN and in Table 4 for DT.

PN showed an increase of verb tokens ($d_1 = 4.36$) and of verb types ($d_1 = 4.78$) in the Norwegian narratives. None of these measures were maintained at follow-up. There was also an increase in the total number of words post-treatment ($d_1 = 3.06$) in the narrative in Norwegian as well as in speech rate ($d_1 = 3.94$). Moreover, there was an increase in the total number of CIUs post-treatment ($d_1 = 3.62$). None of the other measures changed significantly, and no follow-up effects were found.

DT showed an increase in verb types only post-treatment ($d_1 = 1.00$). At the discourse level DT produced more words per minute ($d_1 = 4.26$), which was maintained at follow-up ($d_1 = 3.56$). The number of CIUs increased post-treatment ($d_1 = 1.79$), but this was not maintained. None of the other measures changed significantly in any direction.

b) Did verb-retrieval treatment (i.e. a combination of CBT and SFA) lead to transfer, i.e. improvement of narrative production in the untreated language?

In the Portuguese narratives, PN's amount of verbs increased in tokens as well as types post-treatment (verb tokens: d_1 = 12.53, verb types: d_1 = 11.36). This was maintained at follow-up (verb tokens: d_1 = 6.16, verb types: d_1 = 6.05). At the discourse level, PN produced more words in Portuguese post-treatment (d_1 = 11.89), and an increase in words per minute was found (d_1 = 4.84). There was an increase in the number of utterances (d_1 = 4.35) and in the number of CIUs (d_1 = 8.81); however, the percentage of CIUs decreased with 15 percent. All these changes remained at their high levels at follow-up, and for words that were CIUs, the decrease diminished (-7%).

DT also improved both in verb tokens (d_1 = 1.37) and verb types in English post-treatment (d_1 = 1.13). This was not maintained at follow-up. The discourse of DT also improved, with an increased number of utterances (d_1 = 1.25), which was maintained at follow-up (d_1 = 1.01). The total number of words and the amount of words per minute did not change post-treatment, but an increase was evident at follow-up (total words: d_1 = 1.20; words/minute: d_1 = 1.81). Finally, a considerable increase of CIUs was found post-treatment (d_1 = 4.26), and this was maintained at follow-up (d_1 =1.23). A small, insignificant decrease of the percentage of CIUs was apparent post-treatment, this continued to decrease to a significant level at follow-up (-16 %).

3) Was there evidence of improved verb retrieval in the treated language causing inhibition of the untreated, dominant language?

Decreases in the language measures in the untreated languages post treatment are interpret as inhibition.

As pointed out above, there were only two instances of negative findings, or evidence of decline, in the untreated languages: the percentage of words that were CIUs decreased in the untreated language for both participants (PN's Portuguese narrative at post-test: -15 %, DT's English narrative at follow- up: -16 %). This has to be interpreted in relation to the excessive increase of the total number of words.

Discussion

This study investigated treatment effects on an action naming task and on narrative production following verb production therapy in two multilingual speakers with speakers with different types of aphasia. The findings are discussed in relation to the three main topics of the study: therapy effects in the treated language, cross-linguistic transfer, and inhibition.

Therapy effects in the treated language

Following the SFA treatment, the participants' trained verbs improved significantly as measured by the action-naming test, and for PN this increase was maintained several months post treatment. As pointed out previously, the post-hoc analysis of the properties of the trained and untrained verbs showed that PN's trained verbs were less frequent than the untrained ones. Many speakers with aphasia have easier access to high-frequency words than low-frequency words (Kittredge, Dell, Verkuilen, & Schwartz, 2004; Whitworth et al., 2014). The finding of PN's improvement of the trained, less frequent words are therefore an additional support to the effectiveness of the SFA treatment.

Not surprisingly, neither of the participants achieved higher scores on the action-naming test following CBT, as they did not directly practice the verbs used in the outcome measure during this treatment.

The lack of generalisation to untrained verbs as measured by the action-naming test was disappointing. The findings are, however, consistent with other studies of verb retrieval with the SFA specifically (Knoph et al., 2015; Wambaugh & Ferguson, 2007; Wambaugh et al., 2014). Similar results are also found in the majority of verb-treatment studies with monolingual speakers (see reviews by Conroy, Sage, & Lambon Ralph, 2006; and Webster & Whitworth, 2012). In a recent meta-analysis, de Aguiar, Bastiaanse, and Miceli (2016) found that generalisation to untrained verbs occurred for less than 15 % of the participants (166 individual treatment outcomes, obtained from 30 articles). It is worth noting that the studies that report generalisation to untrained verbs have used therapy methods that are comparable to the treatments employed in the present study. As in SFA, they include a discussion of the semantic features of the verb (Carragher et al., 2013; Rose & Sussmilch, 2008), as in the CBT, they focus on informative exchanges (Maul, Conner, Kempler, Radvanski, & Goral, 2014). The lack of generalisation to untrained verbs for DT may be attributed to her challenges in understanding the rationale behind the SFA method and how to perform the task. In the case of PN, on the other hand, the lack of generalisation to untrained verbs may be linked to her grammatical impairments (cf. that she has aphasia of the Broca type, which is associated with grammatical impairment (Potagas et al., 2013)). de Aguiar et al. (2016) found that patients with grammatical impairment who did not receive direct morphological training were less likely to improve in the production of untreated verbs than those who received such training. The lack of generalisation to untrained items underlines the importance of selecting words for treatment that are relevant to the speaker with aphasia.

In narrative production, PN showed an extensive improvement in the treated language, both in the production and variety of verbs, and on the different discourse variables. In line with results of prior studies of verb production in sentence contexts in monolingual aphasia (e.g. Bastiaanse et al., 2006; Fink, Schwartz, Saffran, & Myers, 1992; Schneider & Thompson, 2003), PN produced longer narratives post-treatment, and she did so more efficiently, with a larger number of words per minute. PN's sentences were also more informative post-treatment, indicating that not only did she produce more words, but also that the words she used were correct in the context to a larger extent than pre-treatment. This is also in keeping with findings from several verb-treatment studies in multilingual aphasia (e.g. Altman et al., 2012; Goral et al., 2010; Knoph et al., 2015) where statistically significant positive changes have been reported in the participants' narratives in the treated language.

DT also showed improvements in the narratives, although to a lesser extent than PN. DT's verb production became more diverse, she increased her speech rate, and her sentences were more informative post-treatment. As noted, there were great differences in the extent to which the participants were able to implement the SFA treatment. DT's inability to identify and produce the required semantic features in the SFA may reflect her language deficits and aphasia type. Speakers with fluent aphasia of the Wernicke type have been described to have predominantly lexical-semantic impairments, in addition to auditory comprehension deficits (Bastiaanse & Edwards, 2004). Regardless of DT's challenges with the original SFA procedure, improvement of verb production especially in the treated language was found, thus showing a clear treatment effect. Moreover, the treatment may have strengthened the semantic network, with an increase in the amount of verbs and in the number of CIUs as a consequence.

Comparable results as the ones we found for PN and DT in the present study have been found in previous studies of verb-production treatment with SFA (e.g. Knoph et al., 2015; Wambaugh & Ferguson, 2007) and CBT (Altman et al., 2012), indicating that treatment of

verbs in sentence contexts may generalise to discourse production. In addition, the implementation of semantic feature generation as a strategy for the person with aphasia may have been beneficial for improving the discourse (Wambaugh et al., 2014).

Transfer to untreated languages

Cross-linguistic transfer is an overarching goal in treatment of multilingual speakers with aphasia. Clinically, this is important since there is often a lack of SLTs speaking the language combination of their clients. Such transfer effects are also theoretically interesting. If therapy gains in one language generalise to an untreated language, it implies that the two languages share structures and representations (Goral et al., 2010). Structural similarities between languages have been suggested to play a role for cross-linguistic transfer, in that it may be difficult to find transfer between structurally different languages (e.g. Goral et al., 2010), although the review of Ansaldo and Ghazi Saidi (2014) showed cross-linguistic transfer between a range of languages within the Indo-European family, as does the language of treatment, Norwegian. The languages of PN can be argued to be structurally more dissimilar than the languages of DT, since Portuguese is a Romance language, while English belongs to the Germanic languages, together with Norwegian (Lewis, Simons, & Fennig, 2013).[7] Even though structural similarities between the languages may be important for cross-linguistic transfer, it is not the only influencing factor. For instance, aphasia type and severity may also play a role.

Regarding verb production, transfer to Portuguese was found for PN. Transfer was evident in the action-naming test, and also in the Portuguese narratives the verbs improved significantly following treatment, both in amount and variety. The maintenance effect for Portuguese was considerable for these lexical variables, indicating that PN's word-finding abilities improved also in the untreated language. The findings are in keeping with Goral et al.

(2012), where cross-linguistic transfer effects in action naming were found following SFA and sentence generation treatment.

At the discourse level, PN showed significant improvements, and all these improvements had lasting effects in Portuguese. The finding of significant improvements on nearly all discourse measures in the untreated language indicates that PN's communication skills in Portuguese improved substantially. Portuguese was PN's strongest pre therapy. It was also the language she used at home, with her family and friends (together with Ronga). So even though she received 40–50 hours of language therapy in Norwegian, she was getting a lot more practice in Portuguese at home. Improvements in discourse have been found for some participants in SFA-studies in monolinguals (e.g. Wambaugh & Ferguson, 2007; Wambaugh et al., 2014) and in one multilingual speaker with aphasia (Knoph et al., 2015). Also CBT has shown to improve discourse in studies with monolingual speakers (Kempler & Goral, 2011) and multilingual speakers (Altman et al., 2012). This furthermore supports the suggestions of treating verbs in sentence contexts (Bastiaanse et al., 2006; Links, Hurkmans, & Bastiaanse, 2010), indicating that this may be fruitful for cross-linguistic transfer.

DT also showed treatment-related transfer to verbs in the untreated language, English, although these improvements were not maintained. Improvements in the connected speech in English for DT were evident. Her narratives in English showed an increased amount of utterances as well as improvement of the content of her sentences, measured with the amount of CIUs (both with lasting results). This implies that the treatment of verbs in sentence contexts had a positive effect on the discourse in the untreated language, as well. The results in this study are therefore congruent with the findings of Altman et al. (2012) and support the proposal that treatments targeting sentence generation, without necessarily practicing specific items, can lead to positive carryover outcome to connected language production.

In several studies, cross-linguistic transfer has in several studies shown to be difficult to accomplish in multilingual speakers with aphasia (cf. reviews by Ansaldo & Ghazi Saidi, 2014; Faroqi-Shah et al., 2010; Kohnert, 2009), thus the findings from the present study are promising and should be investigated further in later studies (cf. Clinical implications and conclusions).

Current models of lexical access assume that the search for words is semantically driven (Hall, 2011), therefore, semantic treatment is often a preferred approach in treating word-retrieval difficulties in speakers with aphasia (Peach & Reuter, 2010). Several models of multilingual language processing argue that unbalanced bilinguals (i.e. speakers with a higher proficiency in one language than in another) access the conceptual store through the L1 lexicon (cf. Kroll & Stewart, 1994; Pavlenko, 2009). This way, treatment gains, especially on the lexical-semantic level, in the L2 may transfer to the active, but untreated L1 lexicon through the conceptual store. Additional support for this is the finding of improvement in the number of CIUs in the untreated languages for both participants. The semantic nature of the treatment in as much as the participants were able to apply semantic feature generation as a strategy in combination with more phonological techniques – like the phonological prompting provided by the SLT and the participants' repetition of the correct verb – may have contributed to the findings. It is also conceivable that by stimulating the semantic network, and by activating a concept in the target language, semantically related words in the other languages were in turn activated (Costa & Caramazza, 1999; de Bot, 1992; Edmonds & Kiran, 2006; Green, 1998). The sharing of semantic features between languages, poses for a possible bridge between languages, through the conceptual system.

Inhibition of the untreated language

Inhibition of the untreated language(s) of a multilingual speaker when provided with therapy in only one language is a highly undesirable effect of treatment. Thus, this study specifically

investigated such negative outcome. Inhibition has been reported, especially when treatment has been administered in the postmorbidly weakest language, as in the present study (Goral, 2012; Goral et al., 2013). A decrease in the assessment results in the untreated language is interpreted as inhibition. No inhibitory effects were detected for any of the participants in the present study. On one single measure there was a decrease in the narratives of both participants, namely on the percentage of words that were CIUs. Following treatment, the participants produced significantly longer narratives. It seems, however that the increase in the total number of words produced came at the expense of how informative the discourse was (measured in the percentage of words that were CIUs). This interplay was not lasting for PN though.

The results of this study are inconsistent with the results of Goral and colleagues (2012; 2013) and suggest no inhibition of the untreated languages of the two participants, following treatment in a weaker language (cf. also Knoph et al., 2015 for similar results in a quadrilingual speaker with aphasia). Given the low cut-off for significance in the current project, the results are considered valid. The findings are thus in line with the conclusions of Kohnert (2009) and Faroqi-Shah et al. (2010), in that no harm was done to any of the untreated languages when treatment was provided in a later-acquired language.

Limitations

There are caveats to most clinical studies, including the present one. The two treatment protocols were provided in the same order for both participants: CBT followed by SFA. Hence, we cannot exclude that the improvement shown in the action-naming test following SFA results from the cumulative effect of the two treatments. However, even if this is possible, we do not think it is likely, given that the participants showed stable baselines prior to the study and no gains on the action-naming test following the CBT alone. Moreover, they were in a chronic phase of aphasia, where spontaneous recovery is not to be expected.

To ascertain client compliance, we decided to keep the amount of measurement points to a minimum, for instance by skipping multiple assessments between the treatment blocks. This was crucial, especially since the participants had to be assessed in both their languages, and they both experienced fatigue. This hampered the possibility of comparing the effect size of the two treatment methods. However, we were able to use the McNemar test, which provided measures of significance for the action-naming test, although maintenance measures were not obtainable for each of the methods.

A further possible weakness was the selection of trained verbs for the SFA treatment. As pointed out earlier, verbs were selected for treatment based on the performance of each of the participants on the baseline tests in Norwegian. While it seems reasonable to exclude items that the participant is able to produce correctly from the treatment set, because of the variability on language tasks by people with aphasia (e.g. Howard, Patterson, Franklin, Morton, & Orchard-Lisle, 1984), this might result in 'regression to the mean' (Howard, Best, & Nickels, 2015). By selecting items that the speaker fails to produce on one occasion, a better performance in the next (post-treatment) probe may wrongly be interpreted as a treatment effect. However, the verbs selected for treatment were not selected based on one measurement point only, but on three baseline measurement points. In addition, the baseline scores for the selected verbs did not change significantly, demonstrating that they were stable in their low performance. In addition, control measures of both related and unrelated materials were included, and these too showed no significant changes. This indicates that the treatment effects found are valid, and not a regression to the mean. Additionally, regardless of the clinical desire for generalisation to untrained items, the lack of generalisation to the untrained verbs may provide additional support to the design of the study, in that the improvements were treatment specific (Nickels, Best, & Howard, 2015).

Various measures (e.g. of lexical variation) depend on sample size (Harris Wright, Silverman, & Newhoff, 2003; Malvern, Richards, Chipere, & Durán, 2004), and different sample sizes are used in different studies. For instance, Saffran, Berndt, and Schwartz (1989) use 150 narrative words, whereas Edwards and Bastiaanse (1998) use 300 words. Such long samples are, however, often hard to obtain from speakers with aphasia; both between and within speakers the variation may be great. Thus, since we are presenting two case studies, we have chosen to follow the practice of prior clinical case studies (e.g. Altman et al., 2012; Goral et al., 2012; Kempler & Goral, 2011) in not standardising the narrative sample sizes.

Clinical implications and conclusions

The results of the present study suggest that verb-retrieval treatment provided in sentence contexts in a late-learned weaker language may be propitious for multilingual speakers with fluent and nonfluent aphasia. This is demonstrated at the lexical level and in discourse production, in the treated language as well as in the untreated languages. Finally, the treatment did not lead to an unwanted inhibition of the untreated language. The findings are important for clinicians, as the growing number of multilingual speakers with aphasia pose a clinical challenge. For some multilingual speakers with aphasia bilingual treatment may be preferable, in order to utilise both languages; however, this is often not feasible. Multilingual clinicians who speak (one or several of) the languages of their multilingual clients, are not available in many communities (Roger & Code, 2011). Rather than refusing to offer services to multilingual clients on the grounds that the SLT does not know all of the client's languages, the findings of the present study should encourage SLTs to provide treatment in one of the languages, even if this is a late-learned weaker language of the client.

The results are consistent with the findings of Goral et al. (2012), where selective cross-linguistic transfer effects have been found following SFA and sentence generation treatment. It

is furthermore plausible that treatments like the CBT, that target sentence generation without the practice of pre-selected items, can lead to a positive carryover to connected language production, as proposed by Altman et al. (2012). Both the participants showed generalisation effects on some of the discourse variables, and especially participant PN showed great improvements. Hence, findings largely support the idea of shared networks of the languages in multilingual individuals, indicated by the improvement for the participants.

References

Altman, C., Goral, M., & Levy, E. S. (2012). Integrated narrative analysis in multilingual aphasia: The relationship among narrative structure, grammaticality, and fluency. *Aphasiology, 26*(8), 1029-1052. doi: 10.1080/02687038.2012.686103

Ansaldo, A. I., & Ghazi Saidi, L. (2014). Aphasia Therapy in the Age of Globalization: Cross-Linguistic Therapy Effects in Bilingual Aphasia. *Behavioural Neurology, 2014*, 1-10. doi: 10.1155/2014/603085

Ansaldo, A. I., Ghazi Saidi, L., & Ruiz, A. (2010). Model-driven intervention in bilingual aphasia: Evidence from a case of pathological language mixing. *Aphasiology, 24*(2), 309-324. doi: 10.1080/02687030902958423

Bastiaanse, R., & Edwards, S. (2004). Word order and finiteness in Dutch and English Broca's and Wernicke's aphasia. *Brain and Language, 89*(1), 91-107. doi: 10.1016/s0093-934x(03)00306-7

Bastiaanse, R., Hurkmans, J., & Links, P. (2006). The training of verb production in Broca's aphasia: A multiple-baseline across-behaviours study. *Aphasiology, 20*(2/3/4), 298-311. doi: 10.1080/02687030500474922

Baus, C., Costa, A., & Carreiras, M. (2013). On the effects of second language immersion on first language production. *Acta Psychologica, 142*(3), 402–409. doi: 10.1016/j.actpsy.2013.01.010

Berndt, R. S., Haendiges, A. N., & Wozniak, M. A. (1997). Verb retrieval and sentence processing: Dissociation of an established symptom association. *Cortex, 33*(1), 99-114. doi: 10.1016/S0010-9452(97)80007-X

Berthier, M. L., Green, C., Lara, J. P., Higueras, C., Barbancho, M. A., Dávila, G., & Pulvermuller, F. (2009). Memantine and constraint-induced aphasia therapy in chronic poststroke aphasia. *Annals of Neurology, 65*(5), 577-585. doi: 10.1002/ana.21597

Boyle, M. (2004). Semantic Feature Analysis Treatment for Anomia in Two Fluent Aphasia Syndromes. *American Journal of Speech-Language Pathology, 13*(3), 236-249. doi: 10.1044/1058-0360(2004/025)

Boyle, M. (2010). Semantic Feature Analysis Treatment for Aphasic Word Retrieval Impairments: What's in a Name? *Top Stroke Rehabilitation, 17*(6), 411-422. doi: 10.1310/tsr1706-411

Boyle, M., & Coelho, C. A. (1995). Application of Semantic Feature Analysis as a Treatment for Aphasic Dysnomia. *American Journal of Speech-Language Pathology, 4*(4), 94-98. doi: 10.1044/1058-0360.0404.94

Brysbaert, M., & Duyck, W. (2010). Is it time to leave behind the Revised Hierarchical Model of bilingual language processing after fifteen years of service? *Bilingualism: Language and Cognition, 13*(3), 359-371. doi: 10.1017/S1366728909990344

Busk, P. L., & Serlin, R. (1992). Meta-analysis for single-case research. In T. R. Kratochwill & J. R. Levin (Eds.), *Single-case research design and analysis: New directions for psychology and education* (pp. 187-212). Hillsdale, NJ: Lawrence Erlbaum Associates.

Carragher, M., Sage, K., & Conroy, P. (2013). The effects of verb retrieval therapy for people with non-fluent aphasia: Evidence from assessment tasks and conversation. *Neuropsychological Rehabilitation, 23*(6), 846-887. doi: 10.1080/09602011.2013.832335

Chang, C. B. (2013). A novelty effect in phonetic drift of the native language. *Journal of Phonetics, 41*(6), 520-533. doi: http://dx.doi.org/10.1016/j.wocn.2013.09.006

Coelho, C. A., McHugh, R. E., & Boyle, M. (2000). Semantic feature analysis as a treatment for aphasic dysnomia: A replication. *Aphasiology, 14*(2), 133-142. doi: 10.1080/026870300401513

Conroy, P., Sage, K., & Lambon Ralph, M. A. (2006). Towards theory - driven therapies for aphasic verb impairments: A review of current theory and practice. *Aphasiology, 20*(12), 1159-1185. doi: 10.1080/02687030600792009

Costa, A., & Caramazza, A. (1999). Is lexical selection in bilingual speech production language specific? Further evidence from Spanish-English and English-Spanish bilinguals. *Bilingualism: Language and Cognition, 2*(3), 231-244. doi: 10.1017/S1366728999000334

Costa, A., & Santesteban, M. (2004). Lexical access in bilingual speech production: Evidence from language switching in highly proficient bilinguals and L2 learners. *Journal of Memory and Language, 50*(4), 491-511. doi: 10.1016/j.jml.2004.02.002

Croft, S., Marshall, J., Pring, T., & Hardwick, M. (2011). Therapy for naming difficulties in bilingual aphasia: which language benefits? *International Journal of Language & Communication Disorders, 46*(1), 48-62. doi: 10.3109/13682822.2010.484845

Davis, G. A. (2005). PACE revisited. *Aphasiology, 19*(1), 21-38. doi: 10.1080/02687030444000598

Davis, G. A., & Wilcox, M. J. (1985). *Adult aphasia rehabilitation: Applied pragmatics*. San Diego, CA: College-Hill Press.

de Aguiar, V., Bastiaanse, R., & Miceli, G. (2016). Improving Production of Treated and Untreated Verbs in Aphasia: A Meta-Analysis. *Frontiers in Human Neuroscience, 10*, 468. doi: 10.3389/fnhum.2016.00468

de Bot, K. (1992). A Bilingual Production Model: Levelt's 'Speaking' Model Adapted. *Applied Linguistics, 13*(1), 1-24. doi: 10.1093/applin/13.1.1

de Bot, K. (2004). The Multilingual Lexicon: Modelling Selection and Control. *International Journal of Multilingualism, 1*(1), 17-32. doi: 10.1080/14790710408668176

Difranscesco, S., Pulvermüller, F., & Mohr, B. (2012). Intensive language-action therapy (ILAT): The methods. *Aphasiology, 26*(11), 1317-1351. doi: 10.1080/02687038.2012.705815

Dussias, P. E., & Nuria, S. (2007). The effect of exposure on syntactic parsing in Spanish–English bilinguals. *Bilingualism: Language and Cognition, 10*(1), 101-116. doi: 10.1017/S1366728906002847

Edmonds, L. A., & Kiran, S. (2004). Confrontation naming and semantic relatedness judgements in Spanish/English bilinguals. *Aphasiology, 18*(5-7), 567-579. doi: 10.1080/02687030444000057

Edmonds, L. A., & Kiran, S. (2006). Effect of Semantic Naming Treatment on Crosslinguistic Generalization in Bilingual Aphasia. *Journal of Speech, Language, and Hearing Research, 49*(4), 729-748. doi: 10.1044/1092-4388(2006/053)

Edwards, S., & Bastiaanse, R. (1998). Diversity in the lexical and syntactic abilities of fluent aphasic speakers. *Aphasiology, 12*(2), 99-117. doi: 10.1080/02687039808250466

Faroqi-Shah, Y., Frymark, T., Mullen, R., & Wang, B. (2010). Effect of treatment for bilingual individuals with aphasia: A systematic review of the evidence. *Journal of Neurolinguistics, 23*(4), 319-341. doi: 10.1016/j.jneuroling.2010.01.002

Faroqi-Shah, Y., & Waked, A. N. (2010). Grammatical category dissociation in multilingual aphasia. *Cognitive Neuropsychology, 27*(2), 181-203. doi: 10.1080/02643294.2010.509340

Field, A., Miles, J., & Field, Z. (2012). *Discovering Statistics Using R*. London, UK: SAGE Publications Ltd.

Filiputti, D., Tavano, A., Vorano, L., De Luca, G., & Fabbro, F. (2002). Nonparellel recovery of languages in a quadrilingual aphasic patient. *International Journal of Bilingualism, 6*(4), 395-410. doi: 10.1177/13670069020060040201

Fink, R. B., Schwartz, M. F., Saffran, E. M., & Myers, J. L. (1992). Facilitation of Verb Retrieval Skills in Aphasia: A Comparison of Two Approaches. *Clinical Aphasiology, 263-275.*

Foster, P., Tonkyn, A., & Wigglesworth, G. (2000). Measuring Spoken Language: A Unit for All Reasons. *Applied Linguistics, 3*(21), 354-375. doi: 10.1093/applin/21.3.354

Goral, M. (2012). Cross-Language Treatment Effects in Multilingual Aphasia. In M. R. Gitterman, M. Goral & L. K. Obler (Eds.), *Aspects of Multilingual Aphasia* (pp. 106-121). Bristol, UK: Multilingual Matters.

Goral, M., & Kempler, D. (2009). Training verb production in communicative context: Evidence from a person with chronic non-fluent aphasia. *Aphasiology, 23*(12), 1383-1397. doi: 10.1080/02687030802235203

Goral, M., Levy, E. S., & Kastl, R. (2010). Cross - language treatment generalisation: A case of trilingual aphasia. *Aphasiology, 24*(2), 170-187. doi: 10.1080/02687030902958308

Goral, M., Naghibolhosseini, M., & Conner, P. S. (2013). Asymmetric inhibitory treatment effects in multilingual aphasia. *Cognitive Neuropsychology, 30*(7-8), 564-577. doi: 10.1080/02643294.2013.878692

Goral, M., Rosas, J., Conner, P. S., Maul, K. K., & Obler, L. K. (2012). Effects of language proficiency and language of the environment on aphasia therapy in a multilingual. *Journal of Neurolinguistics, 25*(6), 538-551. doi: 10.1016/j.jneuroling.2011.06.001

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition, 1*(2), 67-81. doi: 10.1017/S1366728998000133

Hall, M. L. (2011). Bilingual Picture-Word Studies Constrain Theories of Lexical Selection. *Frontiers in Psychology, 2*, 1-19. doi: 10.3389/fpsyg.2011.00381

Harris Wright, H., Silverman, S., & Newhoff, M. (2003). Measures of lexical diversity in aphasia. *Aphasiology, 17*(5), 443-452. doi: 10.1080/02687030344000166

Hernández, M., Costa, A., Sebastián-Gallés, N., Juncadella, M., & Reñé, R. (2007). The organisation of nouns and verbs in bilingual speakers: A case of bilingual grammatical category-specific deficit. *Journal of Neurolinguistics, 20*(4), 285-305. doi: 10.1016/j.jneuroling.2006.10.002

Howard, D., Best, W., & Nickels, L. (2015). Optimising the design of intervention studies: critiques and ways forward. *Aphasiology, 29*(5), 526-562. doi: 10.1080/02687038.2014.985884

Howard, D., Patterson, K., Franklin, S., Morton, J., & Orchard-Lisle, V. (1984). Variability and consistency in picture naming by aphasic patients. In F. C. Rose (Ed.), *Advances in Neurology, 42: Progress in Aphasiology* (1984/01/01 ed., Vol. 42, pp. 263-276). New York, NY: Raven Press.

Jacobs, B. J. (2001). Social Validity of Changes in Informativeness and Efficiency of Aphasic Discourse Following Linguistic Specific Treatment (LST). *Brain and Language, 78*(1), 115-127. doi: 10.1006/brln.2001.2452

Kambanaros, M. (2008). The trouble with nouns and verbs in Greek fluent aphasia. *Journal of Communication Disorders, 41*(1), 1-19. doi: 10.1016/j.jcomdis.2007.02.001

Kambanaros, M. (2010). Action and object naming versus verb and noun retrieval in connected speech: Comparisons in late bilingual Greek-English anomic speakers. *Aphasiology, 24*(2), 210-230. doi: 10.1080/02687030902958332

Kambanaros, M. (2016). Verb and noun word retrieval in bilingual aphasia: a case study of language- and modality-specific levels of breakdown. *International Journal of Bilingual Education and Bilingualism, 19*(2), 169-184. doi: 10.1080/13670050.2015.1037717

Kambanaros, M., & van Steenbrugge, W. (2006). Noun and verb processing in Greek-English bilingual individuals with anomic aphasia and the effect of instrumentality and verb-noun name relation. *Brain and Language, 97*(2), 162-177. doi: 10.1016/j.bandl.2005.10.001

Kay, J., Coltheart, M., & Lesser, R. (2009). *Psykolingvistisk kartlegging av språkprosessering hos afasirammede (PALPA)* (Bredtvet kompetansesenter, Logopedtjenesten – Helse Bergen, Statped Vest & Øverby kompetansesenter, Trans.). Oslo, Norway: Novus forlag.

Kempler, D., & Goral, M. (2011). A comparison of drill- and communication-based treatment for aphasia. *Aphasiology, 25*(11), 1327-1346. doi: 10.1080/02687038.2011.599364

Kiran, S., & Bassetto, G. (2008). Evaluating the Effectiveness of Semantic-Based Treatment for Naming Deficits in Aphasia: What Works? *Seminars in speech and language, 29*(1), 71-82. doi: 10.1055/s-2008-1061626

Kiran, S., & Edmonds, L. A. (2004). Effect of semantic naming treatment on crosslinguistic generalization in bilingual aphasia. *Brain and Language, 91*(1), 75-77. doi: 10.1016/j.bandl.2004.06.041

Kiran, S., & Iakupova, R. (2011). Understanding the relationship between language proficiency, language impairment and rehabilitation: Evidence from a case study. *Clinical Linguistics & Phonetics, 25*(6-7), 565-583. doi: 10.3109/02699206.2011.566664

Kiran, S., Sandberg, C., Gray, T., Ascenso, E., & Kester, E. (2013). Rehabilitation in Bilingual Aphasia: Evidence for Within- and Between-Language Generalization. *American Journal of Speech-Language Pathology, 22*, 298-309.

Kirmess, M., & Lind, M. (2011). Spoken language production as outcome measurement following constraint induced language therapy. *Aphasiology, 25*(10), 1207-1238. doi: 10.1080/02687038.2011.589986

Kirmess, M., & Maher, L. M. (2010). Constraint induced language therapy in early aphasia rehabilitation. *Aphasiology, 24*(6-8), 725-736. doi: 10.1080/02687030903437682

Kittredge, A. K., Dell, G. S., Verkuilen, J., & Schwartz, M. F. (2004). Where is the effect of frequency in word production? Insights from aphasic picture-naming errors. *Cognitive Neuropsychology, 25*(4), 463-492. doi: 10.1080/02643290701674851

Knoph, M. I. K. (2013). Language intervention in Arabic-English bilingual aphasia: A case study. *Aphasiology, 27*(12), 1440-1458. doi: 10.1080/02687038.2013.832139

Knoph, M. I. N., Lind, M., & Simonsen, H. G. (2015). Semantic Feature Analysis targeting verbs in a quadrilingual speaker with aphasia. *Aphasiology, 29*(12), 1473-1496. doi: 10.1080/02687038.2015.1049583

Kohnert, K. (2004). Cognitive and cognate-based treatments for bilingual aphasia: A case study. *Brain and Language, 91*(3), 294-302. doi: 10.1016/j.bandl.2004.04.001

Kohnert, K. (2009). Cross-language generalization following treatment in bilingual speakers with aphasia: a review. *Seminars in Speech and Language, 30*(3), 174-186. doi: 10.1055/s-0029-1225954

Kratochwill, T. R., Hitchcock, J. H., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shadish, W. R. (2013). Single-Case Intervention Research Design Standards. *Remedial and Special Education, 34*(1), 26-38. doi: 10.1177/0741932512452794

Kroll, J. F., Bobb, S. C., & Wodniecka, Z. (2006). Language selectivity is the exception, not the rule: arguments against a fixed locus of language selection in bilingual speech. *Bilingualism: Language and Cognition, 9*(2), 119-135. doi: 10.1017/S1366728906002483

Kroll, J. F., Dussias, P. E., Bice, K., & Perrotti, L. (2015). Bilingualism, Mind, and Brain. *Annual Review of Linguistics, 1*, 377-394. doi: 10.1146/annurev-linguist-030514-124937

Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language, 33*(2), 149-174. doi: 10.1006/jmla.1994.1008

Kroll, J. F., Van Hell, J. G., Tokowicz, N., & Green, D. W. (2010). The Revised Hierarchical Model: A critical review and assessment. *Bilingualism: Language and Cognition, 13*(3), 373-381. doi: 10.1017/S136672891000009X

Kurland, J., & Falcon, M. (2011). Effects of cognate status and language of therapy during intensive semantic naming treatment in a case of severe nonfluent bilingual aphasia. *Clinical Linguistics & Phonetics, 25*(6-7), 584-600. doi: 10.3109/02699206.2011.565398

Lewis, M. P., Simons, G. F., & Fennig, C. D. (2013). Ethnologue: Languages of the World, Seventeenth edition. Retrieved March 05, 2015

Linck, J. A., Kroll, J. F., & Sunderman, G. (2009). Losing access to the native language while immersed in a second language: evidence for the role of inhibition in second-language learning. *Psychological Science, 20*(12), 1507-1515. doi: 10.1111/j.1467-9280.2009.02480.x

Links, P., Hurkmans, J., & Bastiaanse, R. (2010). Training verb and sentence production in agrammatic Broca's aphasia. *Aphasiology, 24*(11), 1303-1325. doi: 10.1080/02687030903437666

Lowell, S., Beeson, P. M., & Holland, A. L. (1995). The Efficacy of a Semantic Cueing Procedure on Naming Performance of Adults With Aphasia. *American Journal of Speech-Language Pathology, 4*(4), 109-114. doi: 10.1044/1058-0360.0404.109

Malvern, D. D., Richards, B. J., Chipere, N., & Durán, P. (2004). *Lexical diversity and language development. Quantification and assessment*. New York, NY: Palgrave Macmillan.

Marangolo, P., Rizzi, C., Peran, P., Piras, F., & Sabatini, U. (2009). Parallel Recovery in a Bilingual Aphasic: A Neurolinguistic and fMRI Study. *Neuropsychology, 23*(3), 405-409. doi: 10.1037/a0014824

Maul, K. K., Conner, P. S., Kempler, D., Radvanski, C., & Goral, M. (2014). Using Informative Verbal Exchanges to Promote Verb Retrieval in Nonfluent Aphasia. *American Journal of Speech-Language Pathology, 23*, 407-420.

Meuter, R., & Allport, A. (1999). Bilingual Language Switching in Naming: Asymmetrical Costs of Language Selection. *Journal of Memory and Language, 40*(1), 25-40. doi: 10.1006/jmla.1998.2602

Miertsch, B., Meisel, J. M., & Isel, F. (2009). Non-treated languages in aphasia therapy of polyglots benefit from improvement in the treated language. *Journal of Neurolinguistics, 22*(2), 135-150. doi: 10.1016/j.jneuroling.2008.07.003

Miller Amberber, A. (2011). Language intervention in French-English bilingual aphasia: Evidence of limited therapy transfer. *Journal of Neurolinguistics, 25*(6), 588-614. doi: 10.1016/j.jneuroling.2011.10.002

Miozzo, M., Costa, A., Hernández, M., & Rapp, B. (2010). Lexical processing in the bilingual brain: Evidence from grammatical/morphological deficits. *Aphasiology, 24*(2), 262-287. doi: 10.1080/02687030902958381

Morris, J., Webster, J., Whitworth, A., & Howard, D. (2009). *Newcastle University Aphasia Therapy Resources. University of Newcastle Upon Tyne*. Newcastle Upon Tyne, UK: University of Newcastle Upon Tyne.

Morris, J., Webster, J., Whitworth, A., & Howard, D. (2012). *Newcastle University undervisningsmateriell for afasirammede: Setningsprosessering* (I. Røste, L. Haaland-Johansen & E. Qvenild, Trans.). Oslo, Norway: Statped sørøst.

Muñoz, M. L., Marquardt, T. P., & Copeland, G. (1999). A Comparison of the Codeswitching Patterns of Aphasic and Neurologically Normal Bilingual Speakers of English and Spanish. *Brain and Language, 66*(2), 249-274. doi: 10.1006/brln.1998.2021

Mätzig, S., Druks, J., Masterson, J., & Vigliocco, G. (2009). Noun and verb differences in picture naming: Past studies and new evidence. *Cortex, 45*(6), 738-758. doi: 10.1016/j.cortex.2008.10.003

Nicholas, L. E., & Brookshire, R. H. (1993). A System for Quantifying the Informativeness and Efficiency of the Connected Speech of Adults With Aphasia. *Journal of Speech and Hearing Research, 36*(2), 338-350. doi: 10.1044/jshr.3602.338

Nickels, L. (2002). Therapy for naming disorders: Revisiting, revising, and reviewing. *Aphasiology, 16*(10-11), 935-979.

Nickels, L., Best, W., & Howard, D. (2015). Optimising the ingredients for evaluation of the effects of intervention. *Aphasiology, 29*(5), 619-643. doi: 10.1080/02687038.2014.1000613

Paradis, M., & Libben, G. (1987). *The Assessment of Bilingual Aphasia*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Pavlenko, A. (2009). Conceptual Representation in the Bilingual Lexicon and Second Language Vocabulary Learning. In A. Pavlenko (Ed.), *The Bilingual Mental Lexicon: Interdisciplinary Approaches* (pp. 125-160). Bristol, UK: Multilingual Matters.

Peach, R. K., & Reuter, K. A. (2010). A discourse-based approach to semantic feature analysis for the treatment of aphasic word retrieval failures. *Aphasiology, 24*(9), 917-990. doi: 10.1080/02687030903058629

Potagas, C., Kasselimis, D. S., & Evdokimidis, I. (2013). Elements of Neurology Essential for Understanding the Aphasias. In I. Papathanasiou, P. Coppens & C. Potagas (Eds.), *Aphasia and Related Neurogenic Communcation Disorders* (pp. 23-47). Burlington, MA: Jones & Barlett Learning.

Pring, T. (2005). *Research Methods in Communication Disorders*. London, UK: Whurr Publishers.

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

Roger, P., & Code, C. (2011). Lost in translation? Issues of content validity in interpreter-mediated aphasia assessments. *Internatinal Journal of Speech and Language Pathology, 13*(1), 61-73. doi: 10.3109/17549507.2011.549241

Rose, M., & Sussmilch, G. (2008). The effects of semantic and gesture treatments on verb retrieval and verb use in aphasia. *Aphasiology, 22*(7-8), 691-706. doi: 10.1080/02687030701800800

Ross, K. B., & Wertz, R. T. (1999). Comparison of impairment and disability measures for assessing severity of, and improvement in, aphasia. *Aphasiology, 13*(2), 113-124. doi: 10.1080/026870399402235

Saffran, E. M., Berndt, R. S., & Schwartz, M. F. (1989). The quantitative analysis of agrammatic production: procedure and data. *Brain and Language, 37*(3), 440-479. doi: 10.1016/0093-934X(89)90030-8

Schneider, S. L., & Thompson, C. K. (2003). Verb production in agrammatic aphasia: The influence of semantic class and argument structure properties on generalisation. *Aphasiology, 17*(3), 213-241. doi: 10.1080/729255456

Wambaugh, J. L., & Ferguson, M. (2007). Application of semantic feature analysis to retrieval of action names in aphasia. *Journal of Rehabilitation Research & Development, 44*(3), 381-394. doi: 10.1682/jrrd.2006.05.0038

Wambaugh, J. L., Mauszycki, S., & Wright, S. (2014). Semantic feature analysis: Application to confrontation naming of actions in aphasia. *Aphasiology, 28*(1), 1-24. doi: 10.1080/02687038.2013.845739

Webster, J., & Whitworth, A. (2012). Review. Treating verbs in aphasia: exploring the impact of therapy at the single word and sentence level. *International Journal of Language & Communication Disorders, 47*(6), 619-636. doi: 10.1111/j.1460-6984.2012.00174.x

Webster, J., Whitworth, A., & Morris, J. (2015). Is it time to stop "fishing"? A review of generalisation following aphasia intervention. *Aphasiology, 29*(11), 1240-1264. doi: 10.1080/02687038.2015.1027169

Weekes, B. S., & Raman, I. (2008). Bilingual deep dysphasia. *Cognitive Neuropsychology, 25*(3), 411-436. doi: 10.1080/02643290802057311

Whitworth, A., Webster, J., & Howard, D. (2014). *A Cognitive Neuropsychological Approach to Assessment and Intervention in Aphasia* (2 ed.). Hove, UK: Psychology Press.

Footnotes

[1] In this paper, the terms bilingual and multilingual will be used somewhat interchangeably. By not separating these two groups, the use of the cumbersome term *bi- and/or multilinguals* is avoided.

[2] The terms narrative, connected speech, and discourse will be used interchangeably in the paper.

[3] Ronga is a Bantu language spoken mainly in Mozambique. Assessment of this language is not included, as neither the Bilingual Aphasia Test nor the action-naming test exist in this language. Narratives were collected in Ronga, but we have not been able to find qualified assistants to transcribe and analyse the data.

[4] For both participants, assessment of aphasia type and severity is based on the Bilingual Aphasia Test (Paradis & Libben, 1987) and clinical judgement, since the BAT does not provide information on aphasia type.

[5] Information on the language use and the proficiency level of each of the languages was obtained with the *Language Use Questionnaire* (Muñoz, Marquardt, & Copeland, 1999) and part A of the BAT (Paradis & Libben, 1987) for both participants.

[6] Also referred to as *constraint-induced language therapy* (CILT) or *constraint-induced aphasia therapy* (CIAT) (e.g. Berthier et al., 2009; Pulvermüller & Berthier, 2008).

[7] Ronga belongs to the Niger–Congo language family, and is structurally very different from Norwegian. Since Ronga was not assessed in this study, we do not know whether the treatment in Norwegian affected this language in any way.

Appendix 1. Short form of the BAT

OVERVIEW OF SUBTESTS ON THE SHORT FORM OF THE BILINGUAL APHASIA TEST

Part B

Language background. Items 4–17

Pointing. Items 23–32

Commands. Items 33–47

Syntactic Comprehension. *Items 66–70, 71–76, 81–88, 111–114, 125–128, 137, 140, 142, 143, 145, 148, 150, 151*

Semantic Categories. Items 125–157

Synonyms. Items 158–162

Antonyms. *Items 163–172*

Verbal Fluency. *Items 263–266 (phonological fluency). In addition was one task of semantic fluency added, due to the focus of the treatment*

Naming. *Items 269–288*

Sentence Construction. *Items 289–313*

Picture Description. Items 344–346

Listening Comprehension. *Items 362–366*

<APPENDIX 2>

Table 3. Narrative results for PN in Portuguese (L1) and Norwegian (L3), with effect sizes

Narrative		Portuguese (L1)					Norwegian (L3)				
		Pre-treatment (mean)	**Post-treatment**	**Effect size**	**Follow-up**	**Effect size (follow-up)**	**Pre-treatment (mean)**	**Post-treatment**	**Effect size**	**Follow-up**	**Effect size (follow-up)**
Lexical variables	**Verb tokens**	9.80 (3.77)	57	**d_1 = 12.53**	33	**d_1 = 6.16**	27 (8.04)	62	**d_1 = 4.35**	25	d_1 = 0.25
	Verb types	9.20 (3.77)	52	**d_1 = 11.36**	32	**d_1 = 6.05**	25 (6.48)	56	**d_1 = 4.78**	24	d_1 = -0.15
Discourse variables	**Total words**	55.40 (22.42)	322	**d_1 = 11.89**	190	**d_1 = 6.00**	170.5 (49.92)	323	**d_1 = 3.06**	167	d_1 = -0.68
	Words/ minute	8.38 (2.51)	20.50	**d_1 = 4,84**	19.66	**d_1 = 1.44**	11.54 (1.73)	18.37	**d_1 = 3.94**	14.90	**d_1 = 1.94**
	Number of utterances	17	57	**d_1 = 4.35**	39	**d_1 = 2.41**	59	79	d_1 = 0.89	64	d_1 = 0.22
	CIUs	41.00 (17.13)	192	**d_1 = 8.81**	130	**d_1 = 5.20**	93.75 (27.99)	195	**d_1 = 3.62**	85	d_1 = -0.31
	CIUs/total verbal units	75%	60%	**-15%**	68%	-7%	55%	60%	5%	51%	-4%

Significant effects are in bold

Table 4. Narrative results for DT in English (L1) and Norwegian (L2), with effect sizes

Narrative		English (L1)					Norwegian (L2)				
		Pre-treatment (mean)	**Post-treatment**	**Effect size**	**Follow-up**	**Effect size (follow-up)**	**Pre-treatment (mean)**	**Post-treatment**	**Effect size**	**Follow-up**	**Effect size (follow-up)**
Lexical variables	**Verb tokens**	44 (13.63)	63	**d_1 = 1.37**	55	d_1 = 0.78	55 (24.64)	77	d_1 = 0.89	62	d_1 = 0.28
	Verb types	39 (14.74)	56	**d_1 = 1.13**	49	d_1 = 0.66	48 (22.91)	71	**d_1 = 1.00**	56	d_1 = 0.35
Discourse variables	**Total words**	187 (58.73)	245	d_1 = 0.99	257	**d_1 = 1.20**	335 (99.63)	391	d_1 = 0.56	360	d_1 = 0.25
	Words/minute	28.55	28.82	d_1 = 0.06	37.61	**d_1 = 1.81**	17.00	30.60	d_1 = **4.26**	28.4	**d_1 = 3.56**
	Number of utterances	27	42	**d_1 = 1.25**	39	**d_1 = 1.01**	81	90	d_1 = 0.53	81	d_1 = -0.02
	CIUs	104	124	**d_1 = 4.26**	110	**d_1 = 1.23**	57	177	d_1 = **1.79**	129	d_1 = 0.84
	CIUs/total verbal units	59 %	51 %	-8 %	43 %	**-16 %**	24 %	45 %	**21 %**	36 %	12 %

Significant effects are in bold

Appendix 2. SFA diagram

HVEM GJØR DETTE VANLIGVIS? *Who does this usually?* _____	HVA BLIR DET GJORT MED? *What/who is it done to/with?* _____	HVOR SKJER DETTE VANLIGVIS? *Where does this usually happen?* _____
	MÅLBILDE *Picture of target word*	
HVORFOR SKJER DET (MENINGEN)? *Why does it happen?* _____	HVILKEN DEL AV KROPPEN/HVILKET VERKTØY? *Which part of the body/What tool is used?* _____	HVA FÅR DET DEG TIL Å TENKE PÅ? *What does it make you think of?* _____

English translation in italics.