ASHA American Speech-Language-Hearing Association

AJSLP

Research Article

Treatment-Induced Recovery Patterns Between Nouns and Verbs in Mandarin–English Bilingual Adults With Aphasia

Ran Li[a] **and Swathi Kiran**[b]

[a] Human Communication, Development, and Information Sciences, Faculty of Education, The University of Hong Kong, Hong Kong
[b] Department of Speech, Language & Hearing Sciences, Boston University, MA

ARTICLE INFO

Article History:
Received July 27, 2022
Revision received November 24, 2022
Accepted April 28, 2023

Editor-in-Chief: Katherine C. Hustad
Editor: Brent Archer

https://doi.org/10.1044/2023_AJSLP-22-00236

ABSTRACT

Purpose: This study aimed to investigate treatment effects of naming therapy targeting nouns and verbs in Mandarin–English bilingual adults with aphasia (BWA).
Method: Twelve Mandarin–English bilingual adults with chronic aphasia completed a 40-hr semantic-based naming treatment for either nouns or verbs. Eight of these participants completed both noun and verb treatment, and the other four completed either noun or verb treatment. Participants were trained in either Mandarin or English for both treatment cycles. Weekly naming probes were measured to capture the direct treatment gain and within- and cross-language generalizations. Performance on the standardized language assessments was analyzed to examine further generalizations beyond the word level and to standardized naming tasks.
Results: Responses in the weekly naming probes showed significant treatment gains in both noun and verb treatment, but the effect was greater in verb treatment. Generalization to semantically related items was captured in noun treatment only. Cross-language generalization was identified in both noun and verb treatment with a larger effect in verb treatment. Additionally, widespread generalizations beyond the word level and to standardized naming tasks were found following both noun and verb treatment, but the effect was larger following noun treatment in discourse and verb naming tasks.
Conclusions: Findings from this study suggested robust treatment effects of semantic-based naming treatment targeting nouns and verbs in Mandarin–English BWA. However, patterns of treatment gains and generalizations differed between these word categories. This study provides strong evidence of bilingual aphasia rehabilitation in Mandarin–English BWA.
Supplemental Material: https://doi.org/10.23641/asha.23818299

Bilingual aphasia refers to the loss of function in one or both languages (Paradis, 2001). Different patterns of language impairment may emerge due to complex interactions among bilingual language history and cross-linguistic differences (Paradis, 2001). In the United States, Chinese is the third mostly spoken language, followed by English and Spanish (Zeigler & Camarota, 2019). The growing bilingual population coincides with an overall increase of older people at risk for neurogenic disorders (i.e., stroke; Hoeffel et al., 2012). However, evidence of language recovery in bilingual adults with aphasia (BWA) is mainly derived from individuals speaking Indo-European languages (e.g., Spanish–English). Given the increase in Chinese-speaking individuals, it is imperative to establish the evidence base for this bilingual population with aphasia.

Language intervention is a critical component of recovery for communication deficits and quality-of-life changes in BWA. Most previous studies examining the effect of naming therapy in BWA have reported significant improvement in the trained items (Gil & Goral, 2004;

Correspondence to Ran Li: rli92@hku.hk. ***Disclosure:*** *No potential conflict of interest was reported by the first author. The second author is a scientific consultant for The Learning Corporation (FKA Constant Therapy), but there is no scientific overlap with the work in this project.*

Kiran et al., 2013; Lerman et al., 2019; Li et al., 2021). However, one conceptual challenge in bilingual aphasia rehabilitation is whether training one language can generalize to the untrained language (i.e., cross-language generalization). Evidence of cross-language generalization is inconsistent across previous studies as some have identified both treatment gains and cross-language generalization (Edmonds & Kiran, 2006; Goral et al., 2012; Kiran & Roberts, 2010; Li et al., 2021; Miertsch et al., 2009), whereas others have only observed direct treatment gains (Galvez & Hinckley, 2003).

Among various treatment approaches targeting lexical impairment in BWA, semantic-based treatment has shown positive treatment gains and generalizations (Edmonds & Kiran, 2006; Kiran & Roberts, 2010; Kiran et al., 2013). Theoretically, since bilingual lexicons share the same semantic system (Costa et al., 2006), training lexical semantics should induce spreading activation from the semantic system not only to semantically related items in the same language (i.e., within-language generalization) but also to items in the untreated language (i.e., cross-language generalization). Kiran and colleagues have previously modified semantic feature analysis (SFA; Boyle, 2004; Boyle & Coelho, 1995) and replicated in bilinguals with a variety of language combinations (Edmonds & Kiran, 2006; Kiran & Roberts, 2010; Kiran et al., 2013). In general, these studies have shown a robust treatment gain and within-language generalization, but patterns of cross-language generalization varied across participants. These results suggest that semantic-based treatment effectively improves lexical retrieval ability in BWA.

Most previous studies in bilingual aphasia rehabilitation have focused on nouns. Nevertheless, lexical retrieval of other word categories (i.e., verbs) may also be affected after brain damage (Bastiaanse & Jonkers, 1998; Vigliocco et al., 2011). There has been evidence suggesting that patterns of lexical impairment differ by a specific grammatical category (i.e., noun or verb) in both monolingual (Kim & Thompson, 2000) and bilingual (Dai et al., 2012; Faroqi-Shah & Waked, 2010; Kambanaros & van Steenbrugge, 2006; Li & Kiran, 2023) individuals with aphasia. However, to what extent similar patterns of treatment-induced language recovery emerge in nouns and verbs remains unclear. Several single-subject studies and case series have implemented semantic-based treatment targeting verb retrieval in bi/multilinguals with aphasia (Ansaldo et al., 2010; Goral et al., 2012; Knoph et al., 2015; Li et al., 2021). Findings from these studies indicate significant improvements in the trained verbs with varying patterns of within- and cross-language generalizations across individuals.

Although similar findings have been reported in either noun or verb treatment studies, none have directly compared treatment effects between these grammatical categories. Knowing whether treatment effect differs by a specific category is essential for developing treatment plans for individuals with lexical retrieval deficits. Three studies to date have targeted both nouns and verbs in bi/multilinguals with aphasia (Ansaldo et al., 2010; Goral et al., 2012; Miertsch et al., 2009). In Miertsch et al. (2009), a German–English–French trilingual with aphasia was trained in the third language (L3; French), targeting lexical semantic processing. Treatment effect was captured via subtests of the Bilingual Aphasia Test (Paradis, 1987) in all three languages. This participant demonstrated significant improvement in word production in both third language (French) and second language (L2; English). In Goral et al. (2012), a Spanish–German–French–English multilingual with aphasia received two phases of naming treatment targeting both verbs and nouns, one in Spanish and one in English. Significant gains in noun and verb naming were reported following both treatment phases. Another study targeted verb and noun retrieval in a Spanish–English bilingual with pathological switching (Ansaldo et al., 2010). Results showed significant improvement in both word categories in the trained language (Spanish) but no generalization to the untrained language (English). Altogether, results from these three studies pointed to a robust treatment gain targeting either verbs or nouns. Nevertheless, none of them have investigated patterns of treatment-induced recovery as a function of grammatical category.

As mentioned above, evidence of bilingual aphasia rehabilitation has been chiefly derived from Indo-European languages (i.e., Spanish–English and French–English). It is important to examine the treatment effect between nouns and verbs in bilinguals speaking typologically different languages (i.e., Mandarin–English). Concretely, Mandarin is a "verb-friendly" language in which verbs do not carry rich morphosyntactic structures (Gentner, 2006; Vigliocco et al., 2011). In addition, pronouns can be dropped in a Mandarin sentence (Huang, 1989), making verbs at the sentence-initial position and more salient. Hence, patterns of treatment-induced language recovery may differ between Mandarin and English, particularly for verbs. To date, one study has examined the effect of a verb therapy in two Mandarin–English BWA (Li et al., 2021). Both participants received a semantic-based naming treatment in their first language (L1; Mandarin). While positive treatment effect and generalizations were reported, therapy was only delivered in one language and evidence was limited to a small sample size.

Improvements beyond the word level can further promote communication skills in individuals with aphasia. Nevertheless, generalization to sentence and discourse production has not been systematically investigated across treatment studies in BWA. Several studies have reported significant improvements in these tasks following verb therapy (Goral et al., 2012; Knoph et al., 2015; Lerman et al., 2019; Li et al., 2021), which corroborates the assumption that verbs can activate nouns being assigned to thematic roles (i.e., agent and patient; McRae et al., 2005). Hence, training the target verb may have the potential for generalization beyond the single-word level.

In summary, the bulk of previous research in bilingual aphasia rehabilitation has been limited to noun treatment, and only a few studies have targeted verbs. Although positive treatment effect has been identified in studies targeting either nouns or verbs, to what extent this effect is similar in both grammatical categories remains unclear. Hence, the same treatment steps need to be implemented so that treatment effects can be examined as a function of grammatical category. Moreover, most evidence of bilingual aphasia rehabilitation has stemmed from Indo-European languages. Given the cross-linguistic differences in verb morphology and verb salience between Mandarin and English, it is important to investigate if similar patterns of treatment-induced language recovery emerge in both languages. Therefore, this study aimed to investigate patterns of treatment gains and generalizations between training nouns and verbs in a larger sample of Mandarin–English BWA who were trained in either Mandarin or English. Results from this study would help us better understand bilingual aphasia treatment targeting different grammatical categories.

Specifically, we investigated patterns of the direct treatment gain, within-language generalization, cross-language generalization, and widespread generalizations to sentence and discourse production and standardized naming tasks. It was hypothesized that Mandarin–English BWA would show significant improvements in naming the trained and semantically related items in treated and untreated languages following both noun and verb treatment (Ansaldo et al., 2010; Goral et al., 2012; Kiran et al., 2013; Lerman et al., 2019; Li et al., 2021). The treatment effect would be larger in verb treatment since verbs may be more impaired than nouns at the semantic representation level (Dai et al., 2012; Faroqi-Shah & Waked, 2010; Kambanaros & van Steenbrugge, 2006). We also hypothesized widespread generalizations to sentence and discourse production and standardized naming tasks after both noun and verb treatment (Goral et al., 2012; Knoph et al., 2015; Lerman et al., 2019; Li et al., 2021). The effects would be larger following verb therapy given the priming effect from verbs to their thematic roles (McRae et al., 2005). Therefore, we tested the above hypotheses by addressing the following aims:

1. Do differences in treatment acquisition, within-language generalization, and cross-language generalization emerge between training nouns and verbs in Mandarin–English BWA?
2. Do differences in generalizations to sentence and discourse production emerge between training nouns and verbs in Mandarin–English BWA?
3. Do differences in standardized naming tasks emerge between training nouns and verbs in Mandarin–English BWA?

Method

Participants

Twelve Mandarin–English bilinguals with chronic aphasia were enrolled in this study (see Table 1; six female participants, mean age = 52.6 ± 18.5 years, mean years of education = 17.7 ± 3.4, mean months postonset = 51.4 ± 48.4, L1: Mandarin). Eleven of them had a single left hemisphere stroke, and one had a traumatic brain injury (P4). These individuals met the following inclusion criteria: (a) fluent in speaking both Mandarin and English before onset (Grosjean, 1982), (b) diagnosed with aphasia based on the Western Aphasia Battery–Revised (WAB-R; Kertesz, 2007) for English and the Aphasia Battery in Chinese (ABC; Gao, 1993) for Mandarin, (c) were between 18 and 85 years old, (d) presented with normal/near-normal or corrected-to-normal hearing and vision, (e) were premorbid right-handed, and (f) had no other neurological condition (i.e., dementia) or learning disorders. Participants were recruited from local and national hospitals, rehabilitation centers, and aphasia support groups. All enrolled participants gave consent according to the Boston University Institutional Review Board protocol.

The bilingual language history was collected via the Language Use Questionnaire (Kastenbaum et al., 2019; see Table 1). Specifically, language usage measured the proportion of time that participants and their conversation partners spent using Mandarin and English during weekdays and weekends. Lifetime exposure captured the average proportion of time that participants heard, spoke, and read each language. Language ability rating (LAR) indicated the average self-rated scores of premorbid abilities to listen, speak, read, and write in each language. All these factors contributed to the overall language proficiency in bilinguals (Peñaloza et al., 2020). Our participants reported higher usage of L2 (English) as they lived in the United States for work or study during the period of this study. Although most of them were proficient in both languages, the overall LAR was relatively higher in L1.

Experimental Design

At the group level, a crossover treatment design was implemented; that is, participants began with either noun

Table 1. Patient demographics, language use history, and treatment assignment.

ID	Sex	Age (years)	Edu (years)	MPO	AoA	Usage %		Exposure %		LAR %		Tx Lang	Tx Assignment
						L1	L2	L1	L2	L1	L2		
P1	F	75.2	18.0	110.8	16.0	4.4	95.6	32.9	67.1	100.0	100.0	C	N ➔ V
P2	M	72.7	20.0	165.2	10.0	33.4	66.6	19.5	80.5	100.0	100.0	C	N ➔ V
P3	M	31.5	25.0	45.0	10.0	43.9	56.1	63.3	36.7	100.0	80.0	E	N ➔ V
P4	F	29.3	17.0	19.2	8.0	0.0	100.0	39.2	60.8	100.0	100.0	C	N ➔ V
P5	F	67.9	13.0	20.8	17.0	22.5	77.5	47.0	53.0	100.0	68.6	E	N ➔ V
P6	F	25.2	20.0	8.2	9.0	56.6	43.4	74.1	25.9	100.0	68.6	E	N ➔ V
P7	M	57.9	20.0	77.3	10.0	19.5	80.5	70.7	29.3	100.0	80.0	C	V ➔ N
P8	M	42.8	16.0	17.4	12.0	50.0	50.0	80.2	19.8	100.0	60.0	C	V ➔ N
P9	F	61.7	19.0	75.5	13.0	36.9	63.1	50.9	49.1	100.0	100.0	E	V
P10	F	53.0	15.0	20.5	12.0	29.2	70.8	43.6	56.4	88.6	82.9	E	V
P11	M	38.7	16.0	50.1	12.0	32.0	68.0	23.1	76.9	90.0	100.0	E	N
P12	M	74.7	13.0	6.2	20.0	50.0	50.0	80.8	19.2	100.0	48.6	E	N
M		52.6	17.7	51.4	12.4	31.5	68.5	52.1	47.9	98.2	82.4		
SD		18.5	3.4	48.4	3.6	17.7	17.7	21.5	21.5	4.2	18.1		

Note. Edu = education; MPO = months postonset; AoA = age of acquisition; Usage = the proportion of time that participants and their conversation partners spent using Mandarin and English during weekdays and weekends; Exposure = the average proportion of time that participants heard, spoke, and read each language; LAR = language ability rating, the average self-rated scores of premorbid abilities to listen, speak, read, and write in each language; Tx Lang = treatment language; Tx = treatment; F = female, M = male; L1 = Mandarin, L2 = English; C = Mandarin Chinese; N ➔ V = noun treatment then verb treatment; E = English; V ➔ N = verb treatment then noun treatment; V = verb treatment; N = noun treatment.

or verb treatment. Within each participant, a multiple-baseline approach (Connell & Thompson, 1986) was implemented for each treatment cycle; that is, each participant served as their own control by completing a baseline phase, a treatment phase, and a posttreatment phase (see Figure 1). Eight of the 12 participants completed a two-cycle treatment (P1–P8), one for verbs and one for nouns (see Table 1). They either completed a noun therapy followed by a verb therapy, or vice versa. Additionally, a 4- to 6-week washout period was added between each cycle,[1] during which participants did not receive any individual speech therapy. The other four participants (P9–P12) completed a one-cycle treatment targeting either nouns or verbs. The same language (either Mandarin or English) was targeted in both treatment cycles, counterbalanced across participants. Hence, a total of five participants were assigned to each treatment condition: (a) English noun, (b) English verb, (c) Mandarin noun, and (d) Mandarin verb.

Standardized Language Assessments and Scoring

A battery of standardized language assessments was administered before and after each treatment cycle, that is, 4 times for P1–P8 (i.e., pretreatment of the first treatment cycle, posttreatment of the first treatment cycle, pretreatment of the second treatment cycle, and posttreatment of the second treatment cycle) and twice for P9–P12 (i.e., pretreatment of the first treatment cycle and posttreatment of the first treatment cycle). Testing sessions were counterbalanced by the target language (i.e., first session: English; second session: Mandarin). All the assessments were conducted remotely via Zoom (https://zoom.us/) during the COVID-19 pandemic (Dekhtyar et al., 2020). Scores were calculated based on guidelines within each test manual and aimed to characterize language impairment in both languages comprehensively.

WAB-R and ABC

The WAB-R (Kertesz, 2007) and the ABC (Gao, 1993) were administered to measure the overall aphasia severity in English and Mandarin, respectively, as characterized by the Aphasia Quotient (AQ). Improvement more than 5.03 points on the AQ was considered clinically significant (Gilmore et al., 2019).

Boston Naming Test

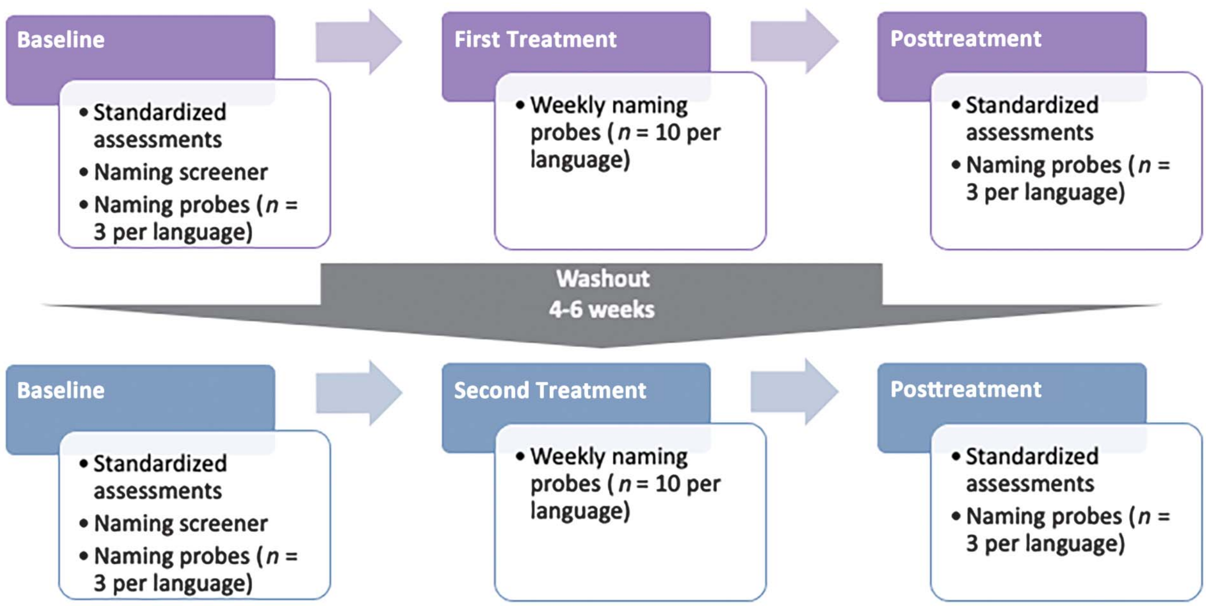
Noun retrieval in English was evaluated with the Boston Naming Test (BNT) long form (Kaplan et al., 2001). A 30-item version (the same items from the long form) was administered to assess noun retrieval in Mandarin (Chen et al., 2014). Improvement of more than 3.30 points was considered clinically significant (Gilmore et al., 2019).

Northwestern Naming Battery and Northwestern Assessment of Verbs and Sentences

Noun and verb naming were further administered using the Northwestern Naming Battery (NNB) and the Verb

[1]The washout period for P3 and P4 was 4 and 6 months, respectively, due to health and personal reasons.

Figure 1. Experimental design. Group-level crosssover treatment design: Participants completed either noun or verb treatment first. Within-subject multiple-baseline approach: Each participant served as their own control by completing baseline, treatment, and posttreatment phases for each treatment cycle. Washout period: no individual speech therapy.



Naming Test from the Northwestern Assessment of Verbs and Sentences (NAVS) for both English (Thompson, 2012; Thompson et al., 2012) and Mandarin (Wang & Thompson, 2016). For the NNB, items that comprised the noun–verb ratio were administered (*n* = 16 each for nouns and verbs).

Verb retrieval in sentences was assessed via the Argument Structure Production Test (ASPT) from the NAVS for English (Thompson, 2012) and Mandarin (Wang & Thompson, 2016). Following previous treatment studies (Edmonds et al., 2009; Li et al., 2021), the target verb was not shown in the ASPT. Responses were coded for the following measures to address Aim 2: (a) complete utterance (CU; 0 = *incomplete*, 1 = *complete*; Edmonds et al., 2009; Lerman et al., 2019), (b) verb production (0 = *inaccurate*, 1 = *accurate*), (c) proportion of argument/noun production (i.e., agent and/or patients), and (d) a 0–5 rating scale capturing the accuracy and completeness of an utterance (see Supplemental Material S1, Section 1.5, for scoring details).

Pyramids and Palm Trees Test

A three-picture version of the Pyramids and Palm Trees Test (Howard & Patterson, 1992) was administered to evaluate semantic processing. This test was administered in the dominant or preferred language, given that bilinguals have a shared semantic system across languages.

Cognitive Linguistic Quick Test

Executive functions were assessed using the Cognitive Linguistic Quick Test (Helm-Estabrooks, 2001), including three nonlinguistic tasks (i.e., symbol trails, mazes, and design generation) and one linguistic task (i.e., generative naming). All the nonlinguistic tasks were administered in the dominant or preferred language, and the linguistic tasks were evaluated in both languages.

Discourse

Connected speech samples were collected in Mandarin and English during separate sessions using a sequential picture ("Umbrella"), a single picture ("Cat Rescue"), and a storytelling task ("The Tortoise and the Hare") from AphasiaBank (https://aphasia.talkbank.org/; see Supplemental Material S1, Section 1.6 and Figure S1). These tasks were chosen as they are relevant to the Chinese population (Kong, 2017) and have been commonly implemented to assess discourse production for both clinical and research purposes (MacWhinney et al., 2011). Responses were video-recorded and transcribed by the first author using the Computerized Language Analysis program (MacWhinney, 2000). All narratives were coded using the Computerized Quantitative Production Analysis command for the total numbers of (a) CUs, (b) utterances, (c) narrative words, (d) nouns, and (e) verbs. The primary outcome measure was the total numbers of words (i.e., narrative words, nouns, and verbs), which have been applied to capture generalizations to discourse production following naming treatment (Dai et al., 2012). The secondary outcome measure was the total numbers of utterances and CUs, given their clinical applicability and high sensitivity to generalization effects

(Edmonds et al., 2009; Lerman et al., 2019; Li et al., 2021).

Stimuli

Either a 298-item noun-naming screener or a 230-item verb-naming screener was administered before the targeted naming therapy (see Supplemental Material S1, Section 1.1). Inaccurate items in both languages were selected for treatment and control stimuli, which also served as naming probes. For each participant, six 15-item sets were selected for nouns (see sample stimuli in Supplemental Material S1, Table A1): (a) English Set 1, trained items (ESN1); (b) Mandarin translation of ESN1 (MSN1); (c) English Set 2, semantically related to Set 1 (ESN2); (d) Mandarin translation of ESN2 (MSN2); (e) English Set 3, unrelated/control items (ESN3); and (f) Mandarin translation of ESN3 (MSN3). Another six 15-item sets were chosen for verbs: (a) English Set 1, trained items (ESV1); (b) Mandarin translation of ESV1 (MSV1); (c) English Set 2, semantically related to Set 1 (ESV2); (d) Mandarin translation of ESV2 (MSV2); (e) English Set 3, unrelated/control items (ESV3); and (f) Mandarin translation of ESV3 (MSV3). Cognates (i.e., more than 50% phoneme overlap) were excluded given their potential cross-linguistic facilitation effect (Costa et al., 2005). Lexical frequency, the number of syllables, familiarity, and imageability were matched within participant[2] (see Supplemental Material S1, Section 1.2). Semantic features for noun and verb stimuli were selected from previous databases (Buchanan et al., 2019; Sandberg et al., 2020). At least 12 features (i.e., six applied and six did not apply) were assigned to each trained item. Feature categories were then assigned to the applied features based on noun and verb SFA treatment protocols (see details in Supplemental Material S1, Section 1.4, and examples in Supplemental Material S1, Tables S2 and S3).

Procedure

Baseline

Three consecutive sessions of probe testing were administered in both languages at the beginning of each treatment cycle. All the probe stimuli (i.e., six 15-item sets, total $n = 90$) were presented in language blocks, counterbalanced across sessions (i.e., Session 1: English first, Mandarin second; Session 2: Mandarin first, English second; Session 3: English first, Mandarin second). Stimuli were pseudorandomized in each probe to ensure that (a) two semantically related items were not presented sequentially and (b) no more than two items from the trained set (i.e., ESN1, MSN1, ESV1, and MSV1) were presented sequentially.

Treatment and Posttreatment Probes

During the treatment phase, the 90-item naming probes in both languages were administered at the beginning of every other treatment session ($n = 10$ per language). Three posttreatment probes for both languages were administered immediately following the treatment phase. Verbal instructions and scoring criteria followed that for naming screeners (see Supplemental Material S1, Section 1.3). Responses were audio- and video-recorded for data analysis.

Treatment

Treatment was delivered via Zoom during the COVID-19 pandemic using PsychoPy3 (Peirce, 2007) on Pavlovia (https://pavlovia.org/). In each treatment cycle, participants received 2-hr sessions twice per week for 10 weeks (total = 40 hr per cycle). One treatment cycle was terminated after all 20 sessions were completed. Treatment was conducted by the first author and trained research assistants who are fluent in both Mandarin and English. For each participant, 15 treatment stimuli (i.e., Set 1 items) were randomized across sessions.

The steps for delivering noun therapy followed previous bilingual aphasia studies (Edmonds & Kiran, 2006; Kiran et al., 2013) and were adapted for verb treatment based on the verb SFA protocol (Wambaugh & Ferguson, 2007). A total of six steps were administered for each target word: (a) spontaneous naming, (b) feature selection and assignment, (c) word association, (d) feature verification, (e) spontaneous naming, and (f) sentence production (see details in Supplemental Material S1, Table S5).

Treatment Reliability

To ensure the reliability of treatment administration, two trained research assistants conducted a fidelity check for 25% of all the videotaped sessions. Point-by-point interrater reliability between Raters 1 and 2 was performed for 20% of their rated sessions using Pearson correlations.

Data Analysis

Descriptive and statistical analyses were conducted to address each research aim. All analyses were performed in RStudio (Version 4.0.3). The significance threshold was set at $p < .05$.[3]

[2]Stimuli were not matched across participants given that not all participants were enrolled at the same time.

[3]Threshold of statistical significance: $p < .05$. Findings with $p < .01$ indicated stronger statistical significance.

Treatment Gain and Within- and Cross-Language Generalizations

Probe responses in all treatment phases (i.e., baseline, treatment, and posttreatment) served as the primary outcome measure. To capture treatment effects in noun treatment, two generalized linear mixed-effects models (GLMMs) were conducted to estimate (a) the direct treatment gain and within-language generalization in the treated language and (b) cross-language generalization in the untreated language. In both models, the item-level accuracy was the dependent variable (0 = *inaccurate*, 1 = *accurate*). Fixed factors included probe session (i.e., Probe 1, Probe 2, etc.), item set (i.e., SN1, SN2, SN3), and a Session × Set interaction term. The Western Aphasia Battery–Aphasia Quotient (WAB-AQ) was included as a covariate. Random intercepts for subject and item and by-subject and by-item random slopes for session were added. Both models were repeated to assess treatment effects in verb treatment. Two additional GLMMs were performed to examine treatment effects between nouns and verbs in the treated and untreated languages. In each model, the item-level accuracy across all stimuli (i.e., nouns and verbs) was the dependent variable. The independent variables included probe session, item set (i.e., SN1, SN2, SN3, SV1, SV2, SV3), treatment category (i.e., verb, noun), and a Session × Set × Category three-way interaction. The treatment order (i.e., first vs. second) and WAB-AQ were entered as covariates. Random intercepts for item and subject and by-item and by-subject random slopes for session were also included. Individual effect sizes were further calculated using the Cohen's *d* statistic (Cohen, 1988; see benchmarks in Supplemental Material S1, Section 2).

Generalizations to Sentence and Discourse Production

To estimate generalization to sentence production, responses from the ASPT in both Mandarin (n = 20 items) and English (n = 22 items) were included.[4] Mixed-effects models were conducted for each dependent variable in the treated and untreated languages separately: CUs (0 or 1), verb production (0 or 1), percent argument/noun production (%), and rating scale (0–5). Fixed factors included time point of assessment (i.e., pre, post), treatment category (i.e., verb, noun), and their interaction term. The WAB-AQ and treatment order were included as covariates. Random intercepts for subject and item were also added.

To estimate generalization to discourse production, Poisson GLMMs were conducted for each dependent variable in the treated and untreated languages separately: total counts of narrative words, nouns, verbs, utterances, and CUs across all three discourse tasks. Fixed factors included time point of assessment (i.e., pre, post), treatment category (i.e., verb, noun), and their interaction term. The WAB-AQ and treatment order were included as covariates, and random intercept for subject was added. Twenty percent of the narratives were randomly assigned to a trained student and transcribed to check for reliability. Interrater reliability was calculated for all the outcome measures using Pearson correlations.

Generalizations to Untrained Verbs and Nouns

To address the last research aim, responses from the standardized naming tasks were included. Noun stimuli were the same items on the Mandarin and English BNTs (n = 30 per language) and items that contributed to the noun–verb ratio on the NNB (n = 16 per language). Verb stimuli were the nonredundant items from the NNB and the NAVS (n = 21 for Mandarin, n = 31 for English). Four GLMMs were performed to estimate the effect of generalization to noun and verb naming in the treated and untreated languages. In each model, the item-level naming accuracy was the dependent measure. Fixed factors included time point (i.e., pre, post), treatment category (i.e., noun, verb), and their two-way interaction. The WAB-AQ and treatment order were included as covariates. Random intercepts for subject and item were also added.

Results

This section includes group-level results. Individual performance is illustrated in Supplemental Material S1, Section 3.[5]

Standardized Language Assessments

As shown in Table 2, the average WAB-AQ in the treated language improved significantly following noun treatment in L1 Mandarin (> 5.03 points; Gilmore et al.,

[4]Twelve items in the English ASPT included the same target verb but optional argument structure (i.e., "the man drives the car" and "the man drives"), which were excluded from the analysis.

[5]Standardized assessments (see Supplemental Material S1, Table S6), treatment probes (see Supplemental Material S1, Figure S2), treatment effect sizes (see Supplemental Material S1, Table S7), sentence production (see Supplemental Material S1, Table S8), discourse production (see Supplemental Material S1, Table S9), and standardized naming tasks (see Supplemental Material S1, Table S10).

2019), and the average BNT score in the treated language increased significantly after verb treatment in L2 English (> 3.30 points; Gilmore et al., 2019). Although other measures did not reveal any statistical significance (paired t tests, $p > .05$), patterns of gains were observed in both treated and untreated languages. These results suggested that both noun and verb treatment facilitated gains in the overall aphasia severity and lexical retrieval ability.

Treatment Gain and Within- and Cross-Language Generalizations

Group-level performance (see Table 3) in noun treatment (see Figures 2a and 2b) exhibited significant gains in the trained item, semantically related untrained items, and translations of the trained items, suggesting a remarkable treatment gain, within-language generalization, and cross-language generalization. Responses in verb treatment (see Figures 2c and 2d) showed significant improvement in the trained but not semantically related items and significant gains in the translations of both trained and untrained items. These results indicated a direct treatment gain and a cross-language generalization in verb treatment, but no within-language generalization. When directly comparing treatment effects between nouns and verbs, the trained items ($p < .05$) and translations of both trained ($p < .01$) and untrained ($p < .05$) items improved to a significantly larger extent in verb treatment, suggesting that verb treatment promoted greater therapy gains and cross-language generalization than noun treatment. Among the 25% of sessions checked for fidelity, the treatment protocol was followed with a reliability of 99.6%. Interrater reliability showed a significant correlation ($r = .2$, $p < .05$).

Generalization to Sentence Production

The results in sentence production (see Table 3) did not reveal any significant improvement in CUs (see Figure 3a) or the target verbs (see Figure 3b) following either noun or verb treatment. However, argument/noun production (see Figure 3c) significantly improved in both treated and untreated languages following noun treatment and in the treated language following verb treatment. The rating scale (see Figure 3d) also significantly gained in both treated and untreated languages following noun and verb treatment. None of these measures revealed any significant differences between training nouns and verbs (i.e., two-way interactions, ps $> .05$). These findings suggested that both noun and verb treatment enhanced argument production and the overall accuracy of sentence production to a similar extent in the treated and untreated languages.

Generalization to Discourse Production

There were significant gains in the total number of narrative words (see Figure 4a) in both treated and untreated languages following noun treatment and in the untreated language following verb treatment (see Table 3). The extent of this generalization in the treated language was significantly greater after noun treatment (i.e., two-way interaction, $p < .05$). Other measures also significantly improved following noun treatment, including noun production (see Figure 4b) in the untreated language and verb production (see Figure 4c) and CUs (see Figure 4d) in both treated and untreated languages. Since generalizations were mostly observed in noun treatment, follow-up analyses were conducted to examine whether performance in verb treatment differed by training in L1 versus L2. The number of narrative words in both treated ($p < .05$) and untreated ($p < .01$) languages and the number of nouns in the treated language ($p < .05$) significantly improved after L1 verb treatment, whereas the number of narrative words in the untreated language ($p < .05$) remarkably improved after L2 verb treatment. Altogether, noun treatment led to better overall lexical retrieval and more CUs, whereas verb treatment mainly generalized to lexical retrieval, particularly when treatment targeted L1. All discourse measures showed high interrater reliability in both L1 and L2 (see Supplemental Material S1, Table S4).

Generalization to Standardized Naming

Results in object noun naming tasks (see Figure 5a) captured significant improvement in the treated language after noun treatment (see Table 3). In action verb naming tasks (see Figure 5b), significant gains were identified in the treated language after both noun and verb treatment and in the untreated language following noun treatment. The magnitude of this generalization in the untreated language was significantly larger after noun treatment (i.e., two-way interaction, $p < .05$). Follow-up analyses were conducted to examine if generalization patterns differed by training in L1 versus L2. Findings showed that L2 noun treatment facilitated verb naming in the treated language ($p < .05$), whereas L1 noun treatment enhanced verb naming in the untreated language ($p < .05$). These results indicated that both noun and verb treatment promoted generalization to standardized naming, but the effect was larger following noun treatment.

Discussion

This study aimed to investigate patterns of treatment-induced language recovery between nouns and verbs via

Table 2. Group-level performance on standardized assessments by each treatment cycle.

Test	TL	Noun Tx								Verb Tx							
		Tx Lang: L1 (*n* = 5)				Tx Lang: L2 (*n* = 5)				Tx Lang: L1 (*n* = 5)				Tx Lang: L2 (*n* = 5)			
		Pre		Post		Pre		Post		Pre		Post		Pre		Post	
		M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
WAB-AQ	T	**55.6**	**19.2**	**61.9***	**18.4**	81.0	2.6	79.7	1.2	**61.1**	**16.8**	**61.7**	**16.8**	84.2	9.8	83.2	9.4
	U	**52.5**	**29.4**	**54.8**	**29.8**	**84.2**	**5.9**	**86.1**	**7.6**	**55.3**	**31.7**	**57.0**	**30.0**	85.6	4.6	85.5	5.6
BNT Raw	T	**7.6**	**8.3**	**8.0**	**7.9**	**21.6**	**4.0**	**24.3**	**5.6**	**8.0**	**8.2**	**8.8**	**9.1**	**26.2**	**7.0**	**29.8***	**6.6**
	U	15.4	19.4	14.8	20.8	**20.0**	**8.1**	**20.2**	**7.8**	**17.2**	**20.0**	**17.6**	**19.8**	19.8	5.1	19.6	5.5
NNB obj %	T	**38.8**	**34.9**	**57.5**	**27.7**	**78.8**	**12.2**	**83.8**	**7.1**	**45.0**	**38.4**	**50.0**	**28.0**	83.8	9.5	82.5	14.3
	U	48.8	46.4	48.8	45.6	**86.3**	**20.9**	**95.0**	**5.2**	**47.5**	**47.3**	**50.0**	**45.7**	**92.5**	**10.3**	**96.3**	**5.6**
NNB act %	T	**26.3**	**35.2**	**32.5**	**32.0**	**58.8**	**18.0**	**68.8**	**11.7**	**36.3**	**29.4**	**41.3**	**30.2**	**70.0**	**19.5**	**76.3**	**21.4**
	U	**25.0**	**36.2**	**35.0**	**36.6**	**68.8**	**17.1**	**80.0**	**12.8**	**33.8**	**37.1**	**38.8**	**38.6**	87.5	15.9	78.8	16.3
VNT %	T	**21.0**	**37.1**	**28.0**	**38.2**	**57.3**	**9.0**	**68.2**	**9.2**	**29.0**	**37.7**	**33.0**	**32.0**	**69.1**	**9.4**	**80.9**	**19.5**
	U	**30.9**	**41.7**	**39.1**	**40.9**	**71.0**	**7.6**	**74.0**	**5.0**	**34.5**	**41.5**	**41.8**	**41.4**	82.0	5.0	71.0	13.5
ASPT ALL %	T	**25.0**	**33.2**	**42.0**	**35.1**	83.1	14.2	80.6	11.6	**37.0**	**29.0**	**45.0**	**30.5**	87.5	18.6	76.3	18.6
	U	**36.3**	**35.4**	**42.5**	**40.9**	**77.0**	**16.4**	**90.0**	**16.7**	**38.1**	**43.8**	**41.3**	**39.4**	**85.0**	**15.7**	**87.0**	**16.4**
CLQT EF-NV		19.2	6.8	17.5	1.7	22.6	9.2	22.2	10.4	22.4	6.3	21.6	6.4	23.6	9.5	21.8	10.2
PPT %		85.7	9.1	85.0	9.9	**94.4**	**5.8**	**95.0**	**2.8**	**82.8**	**15.8**	**89.1**	**10.4**	95.9	2.4	92.5	7.4

Note. Bolded values indicate higher posttreatment than pretreatment scores. Asterisk (*) indicates significant change based on the previous benchmarks. Tx = treatment; Tx Lang = treatment language; L1 = Mandarin, L2 = English; Pre = pretreatment; Post = posttreatment; TL = target language; WAB-AQ = Western Aphasia Battery–Aphasia Quotient (total = 100); BNT Raw = Boston Naming Test raw score (total = 60 for L2, total = 30 for L1); NNB obj = Northwestern Naming Battery Object Naming (total = 16 in L2 and L1); NNB act = Northwestern Naming Battery Action Naming (total = 16 in L2 and L1); VNT = Verb Naming Test (total = 22 in L2, total = 20 in L1); ASPT ALL = Argument Structure Production Test with all arguments (total = 32 in L2, total = 20 in L1); CLQT EF-NV = Cognitive Linguistic Quick Test Executive Function Nonverbal (total = 31); PPT = Pyramids and Palm Trees Test (total = 64).

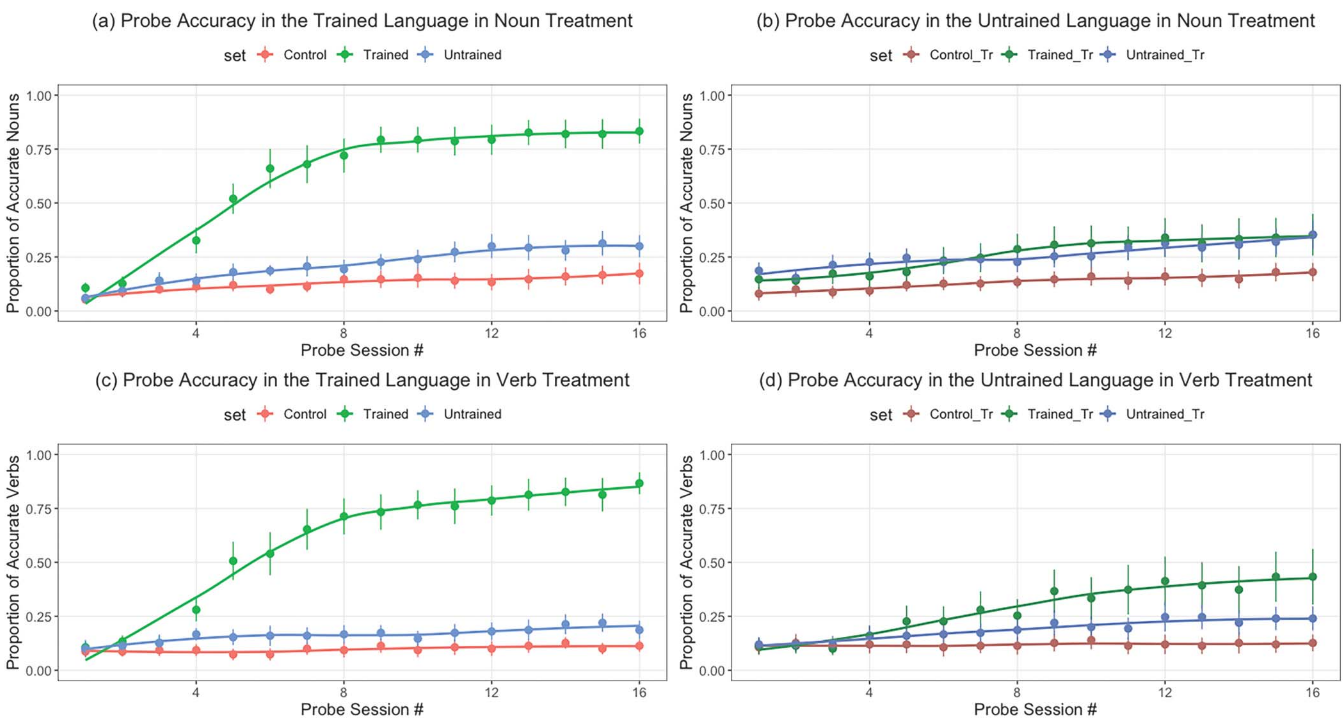
Table 3. Group-level regression results.

RQ 1: Treatment gain and within- and cross-language generalizations										
Predictors	**Noun Tx Model 1: treated**	**Noun Tx Model 2: untreated**	**Verb Tx Model 1: treated**	**Verb Tx Model 2: untreated**	**Noun + Verb Tx Model 1: treated**	**Noun + Verb Tx Model 2: untreated**				
Item Condition (trained) × Session	.45 (.04)***	.13 (.04)***	.50 (.03)***	.31 (.03)***						
Item Condition (untrained) × Session	.07 (.04)**	.03 (.04)	.05 (.03)	.13 (.03)***						
Item Condition (trained) × Session × Category (ref: noun)					.12 (.05)**	.14 (.05)***				
Item Condition (untrained) × Session × Category (ref: noun)					–.02 (.05)	.10 (.05)**				
RQ 2: Generalization to sentence production										
Predictors	**CUs**		**Verb production**		**Argument production**		**Rating scale**			
	Treated	**Untreated**	**Treated**	**Untreated**	**Treated**	**Untreated**	**Treated**	**Untreated**		
Time point (category = noun)	.29 (.28)	–.00 (.39)	.44 (.27)	.41 (.31)	.12 (.02)***	.04 (.02)**	.80 (.19)***	.94 (.23)***		
Time point (category = verb)	.23 (.30)	.61 (.39)	.18 (.29)	.60 (.33)	.07 (.02)***	.00 (.02)	.77 (.20)***	.54 (.23)**		
Category (ref: noun) × Time Point (ref: pre)	–.06 (.41)	.61 (.55)	–.27 (.39)	.19 (.45)	–.05 (.03)	–.04 (.03)	–.03 (.27)	–.40 (.32)		
RQ 2: Generalization to discourse production										
Predictors	**Narrative words**		**Nouns**		**Verbs**		**CUs**		**Utterances**	
	Treated	**Untreated**	**Treated**	**Untreated**	**Treated**	**Untreated**	**Treated**	**Untreated**	**Treated**	**Untreated**
Time point (category = noun)	.20 (.04)***	.23 (.04)***	.12 (.08)	.25 (.08)***	.19 (.09)**	.17 (.09)**	.30 (.12)**	.25 (.12)**	.19 (.10)	.12 (.09)
Time point (category = verb)	.07 (.04)	.17 (.04)***	.11 (.08)	.14 (.08)	.01 (.09)	.14 (.09)	.01 (.13)	.19 (.12)	.11 (.10)	.12 (.09)
Category (ref: noun) × Time Point (ref: pre)	–.13 (.06)**	–.07 (.05)	–.01 (.11)	–.11 (.12)	–.19 (.13)	–.03 (.12)	–.29 (.18)	–.06 (.17)	–.08 (.13)	–.01 (.13)
RQ 3: Generalization to standardized naming										
Predictors	**Noun naming**		**Verb naming**							
	Treated	**Untreated**	**Treated**	**Untreated**						
Time point (category = noun)	.47 (.19)**	.19 (.23)	.68 (.23)***	.79 (.28)***						
Time point (category = verb)	.32 (.20)	.05 (.22)	.63 (.25)**	–.04 (.29)						
Category (ref: noun) × Time Point (ref: pre)	–.15 (.27)	–.14 (.31)	–.05 (.34)	–.83 (.40)**						

Note. Coefficients (standard error) are reported. RQ = research question; Tx = treatment; treated = treated language; untreated = untreated language; ref = reference level; category = treated word category (verb or noun); CUs = complete utterances.

**$p < .05$ (significance threshold). ***$p < .01$.

Figure 2. Group-level performance on naming probes in noun and verb treatment. *X*-axis is Probe Sessions 1–16: 1–3 (baseline phase), 4–13 (treatment phase), 14–16 (posttreatment phase). *Y*-axis is the average response accuracy across participants at each assessed time point (1.00 = 100%). Control_Tr = translations of control items; Trained_Tr = translations of trained items; Untrained_Tr = translations of semantically related untrained items.



the implementation of the same treatment steps for both noun and verb treatment. Specifically, we examined if differences in the following treatment and generalization effects emerged between training nouns and verbs: (a) treatment acquisition and within- and cross-language generalizations, (b) generalizations to sentence and discourse production, and (c) generalizations to standardized naming. Our study captured seven key findings (see Table 4), which will be elaborated in the following paragraphs. This study, to our knowledge, is the first one in bilingual aphasia rehabilitation that included the largest sample size of Mandarin–English BWA. Findings provide evidence of treatment-induced language recovery targeting nouns and verbs and help grow the evidence base for bilingual aphasia rehabilitation in Mandarin–English BWA.

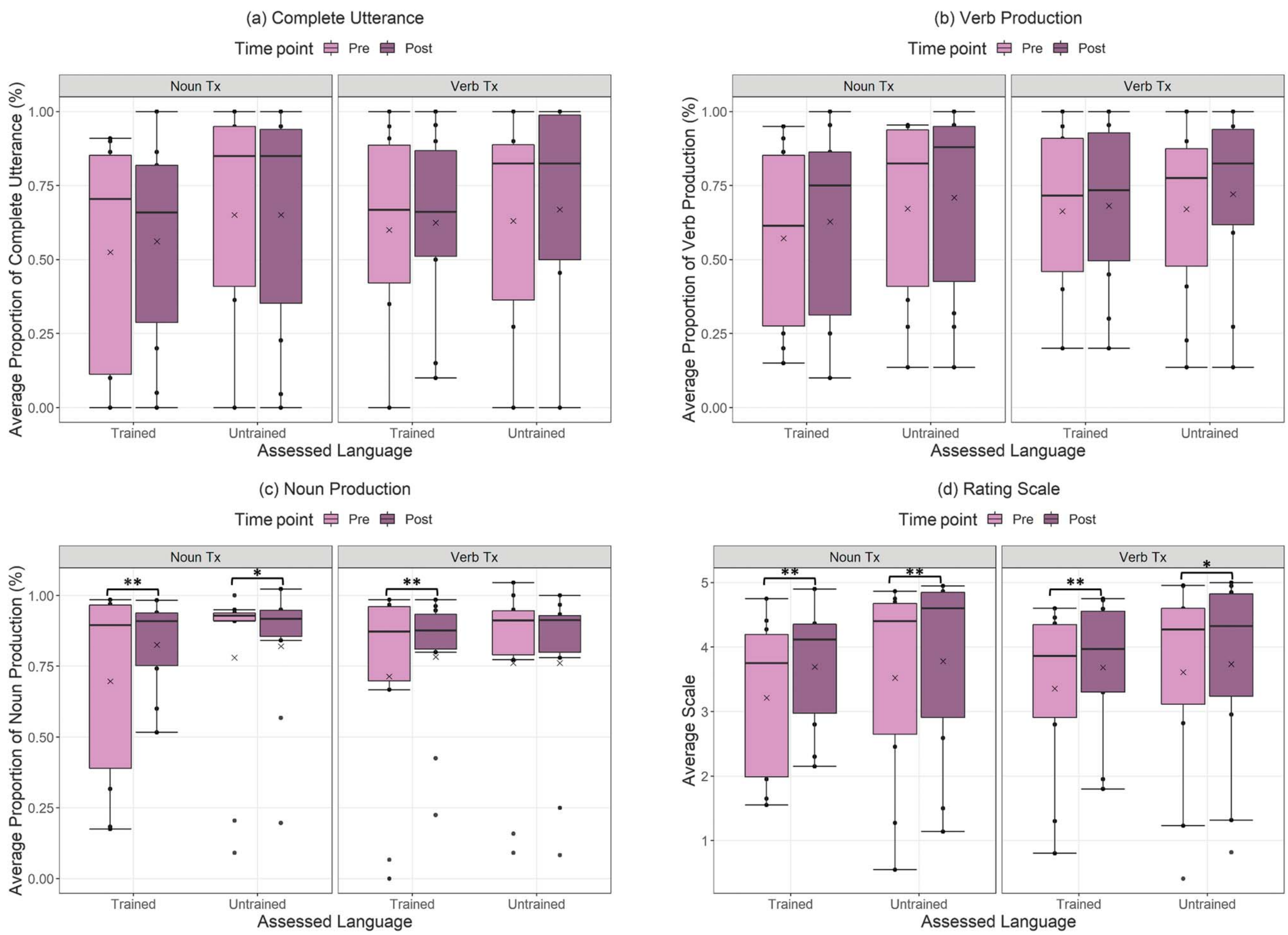
Standardized Language Assessments

Our participants demonstrated significant improvements on the WAB-AQ in the treated language following noun treatment in L1 (i.e., Mandarin) and on the BNT in the treated language following verb treatment in L2 (i.e., English; Finding 1). These findings suggest that practicing lexical retrieval and associated features of an item facilitated the overall aphasia severity and lexical retrieval ability (Edmonds & Kiran, 2006; Kiran & Roberts, 2010; Li et al., 2021).

Treatment Gain and Within- and Cross-Language Generalizations

This study replicated a significant direct treatment effect in both noun and verb treatment in Mandarin–English BWA (Gil & Goral, 2004; Kiran et al., 2013; Lerman et al., 2019; Li et al., 2021), but the magnitude of this effect was significantly larger in verb treatment (Finding 2). Our findings support the hypothesis that repeated activation of labels and their semantic features strengthens the links between semantic representations and associated lexical representations, leading to increased lexical retrieval accuracy (Boyle & Coelho, 1995). Despite the cross-linguistic differences in verb morphology and verb salience between Mandarin and English (Gentner, 2006; Vigliocco et al., 2011), training in either L1 or L2 facilitated the direct treatment gain. This finding is particularly important in bilingual language rehabilitation, as targeting the lexical semantic representation may have the potential for improving lexical retrieval in wider bilingual populations. In addition, a larger treatment effect in verb treatment suggested that retrieving nouns associated with the thematic

Figure 3. Group-level performance on sentence production before and after both noun and verb treatment. Box plots capturing the average performance in the trained and untrained languages based on measures including (a) average proportion of complete utterance (1.00 = 100%), (b) average proportion of accurate verb production (1.00 = 100%), (c) average proportion of noun production (1.00 = 100%), and (d) average rated scale (0–5). Cross (×) indicates the mean value across participants. Asterisk (*) indicates significant improvement from pre- to posttreatment based on pairwise comparisons: **p* < .05. ***p* < .01.



roles of a verb may facilitate verb retrieval (Edmonds et al., 2009; Li et al., 2021; Wambaugh & Ferguson, 2007).

Mandarin–English BWA also demonstrated significant improvements in the semantically related items in noun treatment (Finding 3), which is in line with previous treatment studies (Kiran & Roberts, 2010; Kiran et al., 2013). However, we did not find any within-language generalization in verb treatment. This could be due to conceptual factors that affect the degree of semantic relatedness of verbs. For example, somatotopic information (i.e., body part) is automatically activated in verb processing and primes verbs sharing the same features (e.g., "licking" ➔ "kissing"; Faroqi-Shah et al., 2010). In addition, the naming accuracy tends to be higher for instrumental verbs (i.e., involve an artificial instrument, e.g., "sweep") than non-instrumental verbs in individuals with aphasia (Bastiaanse & Jonkers, 1998). Another explanation is the extent of feature specificity (Faroqi-Shah et al., 2010). Prior studies have shown that training general features (e.g., use tools) of a word (e.g., "dig") was insufficient in facilitating retrieval of another tool verb (e.g., "scoop"). Therefore, generalization to semantically related verbs may require specific features that directly link to the target names.

Items in the untrained language improved significantly in both noun and verb treatment, suggesting the potential for cross-language generalization (Goral et al., 2012; Li et al., 2021). This effect was again more robust when treatment targeted verbs, as evidenced by a larger magnitude of improvement in translations of both trained and untrained items (Finding 4). Such differences between noun and verb treatment may be due to cross-language inhibitory control (Kiran et al., 2013). When both L1 and L2 lexicons are activated, language selection is achieved by inhibiting lexical representations in the nontarget

Figure 4. Group-level performance on discourse tasks before and after both noun and verb treatment. Box plots showing scores averaged across three discourse tasks as measured by the total numbers of (a) narrative words, (b) nouns, (c) verbs, (d) complete utterances (CUs), and (e) utterances. Cross (×) indicates the mean value across participants. Asterisk (*) indicates significant improvement from pre- to post-treatment based on pairwise comparisons: *p < .05. **p < .01.

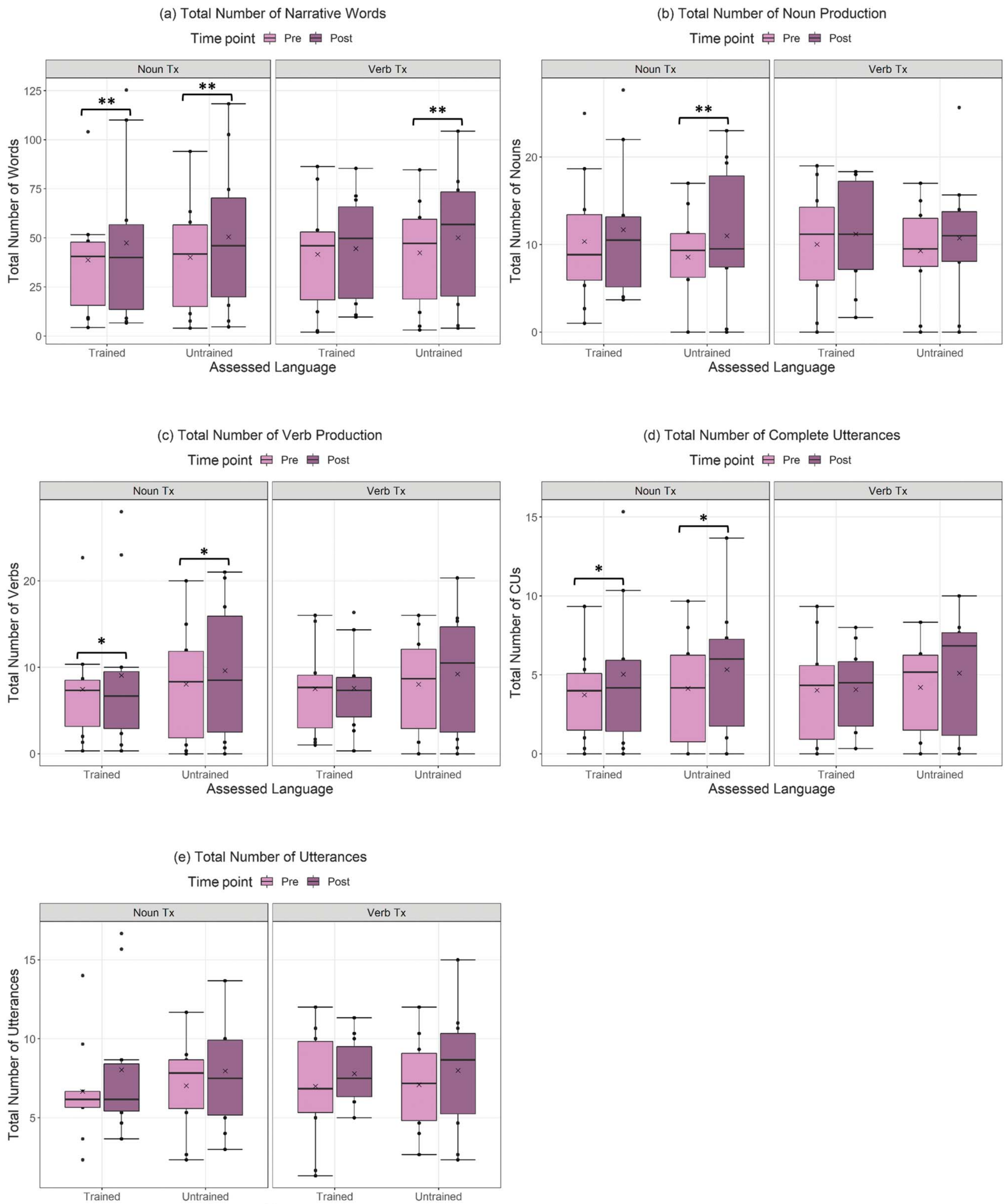
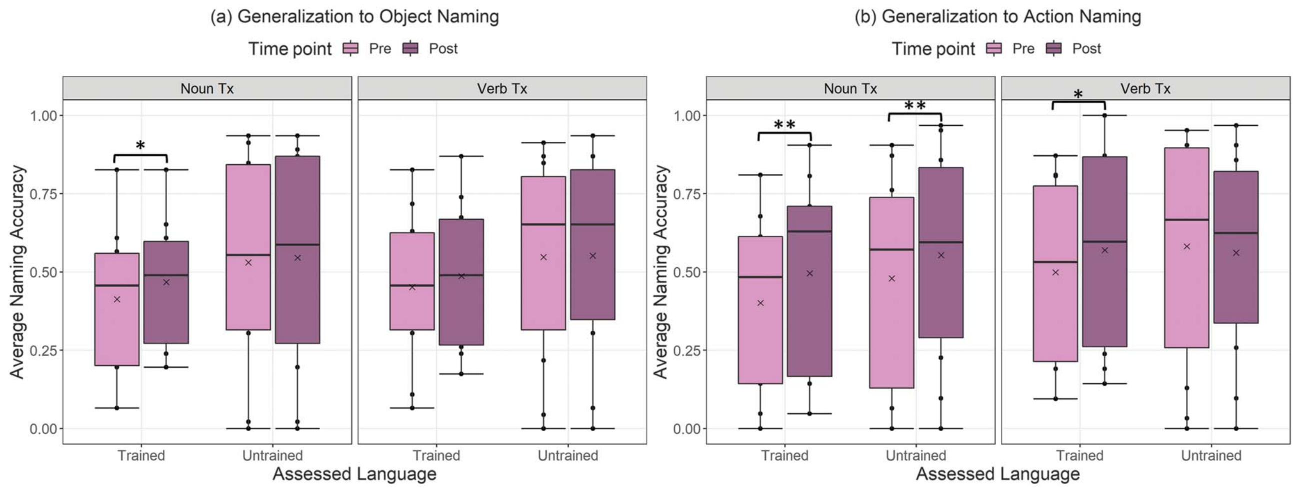


Figure 5. Group-level performance on standardized (a) object naming and (b) action naming tasks. *Y*-axis: naming accuracy averaged across participants. Cross (×) indicates the mean value. Asterisk (*) indicates significant improvement from pre- to posttreatment based on pairwise comparisons: *$p < .05$. **$p < .01$.



language (inhibitory control model; Green, 1998). Since the cross-linguistic semantic overlap is higher for concrete than abstract words (de Groot, 1992), training nouns may yield more cross-linguistic interference, leading to stronger between-languages inhibitory control (Green, 1998). This finding may also be attributed to less complicated verb morphology in Mandarin, which makes lexical access to verbs easier (Li et al., 2019). In fact, three of the four participants who received verb treatment in L2 showed significant effect sizes in the untreated language (i.e., Mandarin; P3, P6, and P9). Altogether, these findings suggested that training semantic features of nouns and verbs promoted generalizations to the untreated language, but the underlying mechanism may depend on individual and language-specific factors.

Generalizations to Sentence and Discourse Production

Both noun and verb treatment promoted generalizations to sentence production to a similar extent, as evidenced by gains in argument production and the overall sentence accuracy (Finding 5). These findings indicated that both types of treatment facilitated lexical retrieval of nouns associated with thematic roles in a sentence (McRae et al., 2005), leading to increased argument production. In addition, since argument structure of a verb is an integral part of the semantic representation (Webster & Whitworth, 2012), repeated feature analysis of a verb's subject and object may have strengthened its predicate argument structure and basic syntactic structure.

Table 4. Summary of key findings.

Outcome measures	**Noun treatment**	**Verb treatment**
1. Standardized language assessments	✓	✓
	WAB-AQ in L1 (Tx lang = L1)	BNT in L2 (Tx lang = L2)
2. Direct treatment gain	✓	✓✓
3. Within-language generalization	✓	–
4. Cross-language generalization	✓	✓✓
5. Generalization to sentences	✓	✓
	Lexical retrieval, overall accuracy	Lexical retrieval, overall accuracy
6. Generalization to discourse	✓✓	✓
	Lexical retrieval, CUs	Lexical retrieval
7. Generalization to standardized naming	✓✓	✓
	Noun naming; verb naming in L2	Verb naming

Note. A checkmark indicates a significant treatment or generalization effect; two checkmarks indicate a larger treatment or generalization effect; a minus sign indicates no significant effect. WAB-AQ = Western Aphasia Battery–Aphasia Quotient; L1 = Mandarin; Tx lang = treatment language; BNT = Boston Naming Test; L2 = English; CUs = complete utterances.

Our participants further demonstrated varying generalization patterns in discourse production, as reflected by significant improvements in the numbers of narrative words, nouns, verbs, and CUs in both treated and untreated languages. However, most of these generalizations were observed following noun treatment (Finding 6). Since generalizations were mainly confined to lexical retrieval, some levels of sentence production ability might be necessary for successful generalization to discourse (Wambaugh et al., 2014). Our follow-up analyses showed that verb treatment targeting L1 (Mandarin) promoted lexical retrieval to a larger extent than L2 (English). This could be due to a higher L1 versus L2 LAR (see Table 1) that may facilitate the generalization effect in BWA (Gil & Goral, 2004). Our results also did not reveal any significant changes in CUs after verb treatment, which have been mainly reported in sentence-based treatment (i.e., Verb Network Strengthening Treatment; Edmonds et al., 2009). Hence, it may be necessary to integrate verbs in a sentence context to promote this generalization. These findings reflect diverse mechanisms of generalization beyond the single-word level that differ by the trained grammatical category.

Generalization to Standardized Naming

Finally, our participants exhibited widespread generalizations to standardized naming of nouns following noun treatment and to naming of verbs following both noun and verb treatment, but the magnitude of the generalization to verb naming was significantly larger after noun than verb treatment (Finding 7). These findings support the hypothesis that strengthening semantic knowledge could lead to better overall lexical retrieval (Lerman et al., 2019; Li et al., 2021). From a semantic perspective, features such as location, time, and function of an object may activate an event (Lancaster & Barsalou, 1997). In addition, nouns are less semantically constrained than verbs and can prime verb processing (McRae et al., 2005). Therefore, training nouns from a variety of semantic categories might have strengthened their semantic connections to verbs.

Our post hoc analyses revealed that training nouns in L2 facilitated generalization to verb naming in the treated language, whereas training nouns in L1 led to generalization to verb naming in the untreated language. In bilingual language processing, inhibitory control in the dominant language (i.e., L1) is stronger when the target language is the less dominant language (Green, 1998). Hence, when our participants received noun therapy in L2, the between-languages inhibitory control might be stronger than the generalized activation between nouns and verbs, leading to increased verb naming in L2. However, when receiving noun therapy in L1, the within-language semantic interference (Belke et al., 2005) might be stronger than between-languages inhibitory control and thus precluded generalization to verb retrieval in the treated language. These findings indicated that generalization to standardized naming may depend on the semantic constraints of nouns and verbs, language-specific properties, and bilingual inhibitory control.

Findings from this study provide clinical guidance for bilingual aphasia rehabilitation in Mandarin–English BWA. Specifically, lexical retrieval of both nouns and verbs should be thoroughly assessed across different linguistic contexts (i.e., naming, sentence/discourse production). Additionally, if the treatment goal is to improve the direct lexical retrieval, semantic-based interventions (i.e., SFA) can benefit both noun and verb retrieval regardless of the target language. However, if the goal is lexical retrieval of the untrained items, treatment stimuli should be carefully selected to maximize this generalization effect. Moreover, if therapy aims to improve lexical retrieval in the untrained language, targeting verbs may facilitate greater generalizations than nouns. Finally, language-specific properties such as verb morphology and verb salience should be considered if therapy is provided in Mandarin Chinese.

Limitations and Future Directions

First, we included a heterogeneous sample that varied in bilingual language history and aphasia profile. It is important for future research to increase the sample size and investigate the effect of these factors on treatment outcomes. Second, our verb treatment did not show any within-language generalization, possibly due to factors affecting the semantic relatedness across stimuli. Future studies should control for variables such as instrumentality and somatotopic information of verbs to maximize this type of generalization. Third, while widespread generalizations beyond the trained items were evident in standardized naming tasks, patterns varied between training nouns and verbs. Hence, future research should examine the effects of semantic constraints of nouns and verbs, cross-linguistic differences between Mandarin and English, and bilingual language control. Finally, although our study demonstrated various significant treatment effects, findings are limited to Mandarin–English BWA. Therefore, future studies should replicate the current findings in other bilingual populations.

Conclusions

This study investigated treatment and generalization effects between nouns and verbs in a semantic-based naming therapy in 12 Mandarin-English BWA. Participants demonstrated significant improvements on the overall aphasia severity (i.e., WAB-AQ) and lexical retrieval ability (i.e., BNT). Responses from weekly naming probes showed direct treatment gains in both noun and verb treatment, but the effect was greater in verb treatment. Generalization to semantically related items was identified in noun treatment but not in verb treatment. Cross-language generalization was observed in both noun and verb treatment, but the effect was significantly larger in verb treatment. Widespread generalizations were captured in sentence and discourse production and standardized naming tasks. However, patterns varied by the target grammatical category. Our findings support the effectiveness of semantic naming treatment targeting both nouns and verbs and help grow the evidence base for aphasia rehabilitation in Mandarin–English BWA.

Data Availability Statement

The data sets generated and/or analyzed during this study are not publicly available due to ethical restrictions but are available from the corresponding author on reasonable request.

Acknowledgments

This work was supported by Boston University Sargent Student Research Grant (Ran Li: 9250001170). The authors were funded by National Institute on Deafness and Other Communication Disorders Grant (Swathi Kiran: 1U01DC014922). The authors would like to thank all the study participants and the members of the Aphasia Research Laboratory for their support throughout the development and execution of this study. They would also like to thank undergraduate research assistants Hongli Zheng, ShiMin Chen, Gengchen Wei, and Jianing Long for their help with treatment administration and data analysis.

References

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

Bastiaanse, R., & Jonkers, R. (1998). Verb retrieval in action naming and spontaneous speech in agrammatic and anomic aphasia. *Aphasiology, 12*(11), 951–969. https://doi.org/10.1080/02687039808249463

Belke, E., Meyer, A. S., & Damian, M. F. (2005). Refractory effects in picture naming as assessed in a semantic blocking paradigm. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology, 58*(4), 667–692. https://doi.org/10.1080/02724980443000142

Boyle, M. (2004). Semantic feature analysis treatment for anomia in two fluent aphasia syndromes. *American Journal of Speech-Language Pathology, 13*(3), 236–249. https://doi.org/10.1044/1058-0360(2004/025)

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. https://doi.org/10.1044/1058-0360.0404.94

Buchanan, E. M., Valentine, K. D., & Maxwell, N. P. (2019). English semantic feature production norms: An extended database of 4436 concepts. *Behavior Research Methods, 51*(4), 1849–1863. https://doi.org/10.3758/s13428-019-01243-z

Chen, T.-B., Lin, C.-Y., Lin, K.-N., Yeh, Y.-C., Chen, W.-T., Wang, K.-S., & Wang, P.-N. (2014). Culture qualitatively but not quantitatively influences performance in the Boston Naming Test in a Chinese-speaking population. *Dementia and Geriatric Cognitive Disorders Extra, 4*(1), 86–94. https://doi.org/10.1159/000360695

Cohen, J. (Ed.). (1988). *Statistical power analysis for the behavioural science* (2nd ed.). Routledge.

Connell, P. J., & Thompson, C. K. (1986). Flexibility of single-subject experimental designs. Part III: Using flexibility to design or modify experiments. *Journal of Speech and Hearing Disorders, 51*(3), 214–225. https://doi.org/10.1044/jshd.5103.214

Costa, A., La Heij, W., & Navarrete, E. (2006). The dynamics of bilingual lexical access. *Bilingualism, 9*(2), 137–151. https://doi.org/10.1017/S1366728906002495

Costa, A., Santesteban, M., & Caño, A. (2005). On the facilitatory effects of cognate words in bilingual speech production. *Brain and Language, 94*(1), 94–103. https://doi.org/10.1016/j.bandl.2004.12.002

Dai, E. Y. L., Kong, A. P. H., & Weekes, B. S. (2012). Recovery of naming and discourse production: A bilingual anomic case study. *Aphasiology, 26*(6), 737–756. https://doi.org/10.1080/02687038.2011.645013

de Groot, A. M. B. (1992). Determinants of word translation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*(5), 1001–1018. https://doi.org/10.1037/0278-7393.18.5.1001

Dekhtyar, M., Braun, E. J., Billot, A., Foo, L., & Kiran, S. (2020). Videoconference administration of the Western Aphasia Battery–Revised: Feasibility and validity. *American Journal of Speech-Language Pathology, 29*(2), 673–687. https://doi.org/10.1044/2019_AJSLP-19-00023

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. https://doi.org/10.1044/1092-4388(2006/053)

Edmonds, L. A., Nadeau, S. E., & Kiran, S. (2009). Effect of Verb Network Strengthening Treatment (VNeST) on lexical retrieval of content words in sentences in persons with aphasia. *Aphasiology, 23*(3), 402–424. https://doi.org/10.1080/02687030802291339

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

Faroqi-Shah, Y., Wood, E., & Gassert, J. (2010). Verb impairment in aphasia: A priming study of body-part overlap. *Aphasiology, 24*(11), 1377–1388. https://doi.org/10.1080/02687030903515362

Galvez, A., & Hinckley, J. J. (2003). Transfer patterns of naming treatment in a case of bilingual aphasia. *Brain and Language, 87*(1), 173–174. https://doi.org/10.1016/S0093-934X(03)00256-6

Gao, S. (1993). *Shiyuzheng* [Aphasia]. Peking Union Medical College.

Gentner, D. (2006). Why verbs are hard to learn. In K. Hirsh-Pasek & R. M. Golinkoff (Eds.), *Action meets word: How children learn verbs* (pp. 544–564). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195170009.003.0022

Gil, M., & Goral, M. (2004). Nonparallel recovery in bilingual aphasia: Effects of language choice, language proficiency, and treatment. *International Journal of Bilingualism, 8*(2), 191–219. https://doi.org/10.1177/13670069040080020501

Gilmore, N., Dwyer, M., & Kiran, S. (2019). Benchmarks of significant change after aphasia rehabilitation. *Archives of Physical Medicine and Rehabilitation, 100*(6), 1131–1139.e87. https://doi.org/10.1016/j.apmr.2018.08.177

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. https://doi.org/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. https://doi.org/10.1017/S1366728998000133

Grosjean, F. (1982). *Life with two languages: An introduction to bilingualism*. Harvard University Press.

Helm-Estabrooks, N. (2001). *Cognitive Linguistic Quick Test: CLQT*. The Psychological Corporation.

Hoeffel, E. M., Rastogi, S., Kim, M. O., & Hasan, S. (2012). The Asian population: 2010. US Department of Commerce, Economics and Statistics Administration, U.S. Census Bureau.

Howard, D., & Patterson, K. E. (1992). *The Pyramids and Palm Trees Test*.

Huang, C. T. J. (1989). Pro-drop in Chinese: A generalized control theory. In O. A. Jaeggli & K. J. Safir (Eds.), *The null subject parameter: Studies in natural language and linguistic theory* (pp. 185–214). Springer. https://doi.org/10.1007/978-94-009-2540-3_6

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. https://doi.org/10.1016/j.bandl.2005.10.001

Kaplan, E., Goodglass, H., & Weintraub, S. (2001). *Boston Naming Test*.

Kastenbaum, J. G., Bedore, L. M., Peña, E. D., Sheng, L., Mavis, I., Sebastian-Vaytadden, R., Rangamani, G., Vallila-Rohter, S., & Kiran, S. (2019). The influence of proficiency and language combination on bilingual lexical access. *Bilingualism, 22*(2), 300–330. https://doi.org/10.1017/S1366728918000366

Kertesz, A. (2007). *Western Aphasia Battery–Revised*. Pearson.

Kim, M., & Thompson, C. K. (2000). Patterns of comprehension and production of nouns and verbs in agrammatism: Implications for lexical organization. *Brain and Language, 74*(1), 1–25. https://doi.org/10.1006/brln.2000.2315

Kiran, S., & Roberts, P. M. (2010). Semantic feature analysis treatment in Spanish–English and French–English bilingual aphasia. *Aphasiology, 24*(2), 231–261. https://doi.org/10.1080/02687030902958365

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*(2), S298–S309. https://doi.org/10.1044/1058-0360(2013/12-0085)

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. https://doi.org/10.1080/02687038.2015.1049583

Kong, A. P. (2017). Speech-language services for Chinese-speaking people with aphasia (C-PWA): Considerations for assessment and intervention. *Perspectives of the ASHA Special Interest Groups, 2*(2), 100–109. https://doi.org/10.1044/persp2.SIG2.100

Lancaster, J. S., & Barsalou, L. W. (1997). Multiple organisations of events in memory. *Memory, 5*(5), 569–599. https://doi.org/10.1080/741941478

Lerman, A., Edmonds, L. A., & Goral, M. (2019). Cross-language generalisation in bilingual aphasia: What are we missing when we do not analyse discourse? *Aphasiology, 33*(9), 1154–1162. https://doi.org/10.1080/02687038.2018.1538493

Li, R., Faroqi-Shah, Y., & Wang, M. (2019). A comparison of verb and noun retrieval in Mandarin–English bilinguals with English-speaking monolinguals. *Bilingualism: Language and Cognition, 22*(5), 1005–1028. https://doi.org/10.1017/S1366728918000913

Li, R., & Kiran, S. (2023). Noun and verb impairment in single-word naming and discourse production in Mandarin-English bilingual adults with aphasia. *Aphasiology*, 1–29. https://doi.org/10.1080/02687038.2023.2189994

Li, R., Li, W., & Kiran, S. (2021). Effect of Mandarin Verb Network Strengthening Treatment (VNeST) in Mandarin–English bilinguals with aphasia: A single-case experimental design. *Neuropsychological Rehabilitation, 31*(8), 1224–1253. https://doi.org/10.1080/09602011.2020.1773278

MacWhinney, B. (2000). The CHILDES project: Tools for analyzing talk: Volume I: Transcription format and programs, volume II: The database.

MacWhinney, B., Fromm, D., Forbes, M., & Holland, A. (2011). AphasiaBank: Methods for studying discourse. *Aphasiology, 25*(11), 1286–1307. https://doi.org/10.1080/02687038.2011.589893

McRae, K., Hare, M., Elman, J. L., & Ferretti, T. (2005). A basis for generating expectancies for verbs from nouns. *Memory and Cognition, 33*(7), 1174–1184. https://doi.org/10.3758/BF03193221

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. https://doi.org/10.1016/j.jneuroling.2008.07.003

Paradis, M. (1987). *Bilingual Aphasia Test*. Lawrence Erlbaum Associates.

Paradis, M. (2001). Bilingual and polyglot aphasia. In R. S. Berndt (Ed.), *Handbook of neuropsychology: Language and aphasia* (pp. 69–91). Elsevier Science.

Peirce, J. W. (2007). PsychoPy–Psychophysics software in Python. *Journal of Neuroscience Methods, 162*(1–2), 8–13. https://doi.org/10.1016/j.jneumeth.2006.11.017

Peñaloza, C., Barrett, K., & Kiran, S. (2020). The influence of prestroke proficiency on poststroke lexical-semantic performance in bilingual aphasia. *Aphasiology, 34*(10), 1223–1240. https://doi.org/10.1080/02687038.2019.1666082

Sandberg, C., Gray, T., & Kiran, S. (2020). Development of a free online interactive naming therapy for bilingual aphasia. *American Journal of Speech-Language Pathology, 29*(1), 20–29. https://doi.org/10.1044/2019_AJSLP-19-0035

Thompson, C. K. (2012). *Northwestern Assessment of Verbs and Sentences* (NAVS).

Thompson, C. K., Lukic, S., King, M. C., Mesulam, M. M., & Weintraub, S. (2012). Verb and noun deficits in stroke-induced and primary progressive aphasia: The Northwestern Naming Battery. *Aphasiology, 26*(5), 632–655. https://doi.org/10.1080/02687038.2012.676852

Vigliocco, G., Vinson, D. P., Druks, J., Barber, H., & Cappa, S. F. (2011). Nouns and verbs in the brain: A review of behavioural, electrophysiological, neuropsychological and imaging studies. *Neuroscience & Biobehavioral Reviews, 35*(3), 407–426. https://doi.org/10.1016/j.neubiorev.2010.04.007

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. https://doi.org/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. https://doi.org/10.1080/02687038.2013.845739

Wang, H., & Thompson, C. K. (2016). Assessing syntactic deficits in Chinese Broca's aphasia using the Northwestern Assessment of Verbs and Sentences–Chinese (NAVS-C). *Aphasiology, 30*(7), 815–840. https://doi.org/10.1080/02687038.2015.1111995

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

Zeigler, K., & Camarota, S. A. (2019). *67.3 million in the United States spoke a foreign language at home in 2018*. Center for Immigration Studies.