BMJ Open

Predicting treatment outcomes for bilinguals with aphasia using computational modeling: Study protocol for the PROCoM randomised controlled trial

Claudia Peñaloza, Maria Dekhtyar, Michael Scimeca, Erin Carpenter, Nishaat Mukadam, Swathi Kiran

To cite: Peñaloza C, Dekhtyar M, Scimeca M, *et al*. Predicting treatment outcomes for bilinguals with aphasia using computational modeling: Study protocol for the PROCoM randomised controlled trial. *BMJ Open* 2020;**10**:e040495. doi:10.1136/bmjopen-2020-040495

► Prepublication history for this paper is available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2020-040495).

Received 14 May 2020
Revised 15 October 2020
Accepted 26 October 2020

Aphasia Research Laboratory, Department of Speech, Language and Hearing Sciences, Boston University, Boston, Massachusetts, USA

Correspondence to
Dr Claudia Peñaloza; penaloza@bu.edu

ABSTRACT

Introduction Bilinguals with aphasia (BWA) present varying degrees of lexical access impairment and recovery across their two languages. Because both languages may benefit from therapy, identifying the optimal target language for treatment is a current challenge for research and clinical practice. Prior research has demonstrated that the BiLex computational model can accurately simulate lexical access in healthy bilinguals, and language impairment and treatment response in bilingual aphasia. Here, we aim to determine whether BiLex can predict treatment outcomes in BWA in the treated and the untreated language and compare these outcome predictions to determine the optimal language for rehabilitation.

Methods and analysis The study involves a prospective parallel-group, double-blind, randomised controlled trial. Forty-eight Spanish–English BWA will receive 20 sessions of semantic treatment for lexical retrieval deficits in one of their languages and will complete assessments in both languages prior and after treatment. Participants will be randomly assigned to an experimental group receiving treatment in the optimal language determined by the model or a control group receiving treatment in the language opposite to the model's recommendation. Primary treatment outcomes include naming probes while secondary treatment outcomes include tests tapping additional language domains. Treatment outcomes will be compared across the two groups using 2×2 mixed effect models for repeated measures Analysis of variance (ANOVA) on metrics of treatment effects commonly employed in rehabilitation studies (ie, effect size and percentage change).

Ethics and dissemination All procedures included in this protocol (protocol number 29, issue date: 19 March 2019) were approved by the Boston University Charles River Campus Institutional Review Board at Boston, Massachusetts (reference number: 4492E). The results of this study will be published in peer-reviewed scientific journals and will be presented at national and international conferences.

Trial registration number NCT02916524.

Strengths and limitations of the study

- This study uses computational modeling to predict language therapy outcomes in bilinguals with aphasia accounting for their bilingual background and language impairment.
- The computational model can accurately simulate lexical access in healthy bilinguals and lexical access deficits and treatment response in bilinguals with aphasia.
- We will evaluate the potential of computational modeling to guide clinical practice with bilinguals with aphasia by predicting the most beneficial target language for rehabilitation.
- The effects of semantic treatment on word finding deficits in both languages will be examined in a large sample including 48 bilinguals with aphasia.
- The use of a non-stratified randomisation approach may limit comparisons between treatment groups regarding baseline characteristics that may influence prognosis.

INTRODUCTION

Aphasia, a common speech and language disorder after brain damage, is a crucial target for bilingual rehabilitation given its contribution to the long-term disability and impact on the quality of life of stroke survivors[1] in bilingual societies worldwide. Bilinguals with aphasia (BWA) experience deficits that affect one or both of their languages with varying degrees and patterns of impairment and recovery,[2 3] resulting in communicational limitations across multiple sociocultural and linguistic contexts. Speech and language therapy is effective for poststroke aphasia,[4] however, most existing evidence on its effectiveness in BWA comes from single case studies or studies with small heterogeneous samples. Current reviews addressing language rehabilitation outcomes in BWA suggest that

both languages show potential for treatment-induced recovery[5–7] although evidence is mixed and remains inconclusive. For instance, some studies have shown that BWA can present treatment gains in the treated language and generalisation effects to the untreated language.[8–13] Yet, other studies have shown benefits in just the treated language but lack of generalisation to the untreated language,[14 15] and even cross-language interference effects.[16 17] Several factors can modulate individual differences in treatment outcomes in BWA[18] including prestroke factors such as bilingual background[19 20] and poststroke factors such as language impairment[10] and lesion location.[17] Inter-individual variability may further be influenced by differences in the methodological approach and interventions employed across studies and the language combinations examined across participants, thus limiting the generalisation of prior findings. Overall, the existing behavioural research has contributed to our current understanding of language dysfunction and treatment-induced recovery in BWA. However, the above-mentioned limitations have resulted in a lack of evidence-based guidelines that can help clinicians develop optimal rehabilitation plans for BWA[21] and crucially, inform them which language should be targeted in treatment to observe maximum treatment gains across the treated and the untreated language.

Predictive models of rehabilitation outcomes in aphasia can help evaluate the contribution of relevant factors to language therapy outcomes.[22 23] In BWA, such models may further allow identifying the optimal treatment language and developing treatment plans that increase the likelihood of observing maximum benefits. However, making reliable predictions of treatment effects is challenging as collecting representative patient datasets requires large scale longitudinal studies that are highly demanding and often unfeasible. Computational modeling can offer an alternative promising approach to the prediction of treatment outcomes in BWA while accounting for relevant prestroke and poststroke factors that influence language recovery. Computational models have been used to examine multiple aspects of second language processing in healthy bilinguals[24–26] and simulate language impairment[27] and treatment response in BWA.[28] Building on our prior research,[25 27 28] we developed BiLex, a neural network model that can accurately simulate lexical access in Spanish–English healthy bilinguals in both their native (L1) and second (L2) language, while accounting for individual differences in their L2 age of acquisition (AOA), and the relative amounts of lifetime use and exposure to each language.[29] The BiLex model has been further employed to simulate lexical access impairment and treatment outcomes in a retrospective sample of 13 Spanish–English BWA.[30] This simulation was successfully achieved by (1) modeling prestroke naming ability for each BWA accounting for their age at testing, L2 AoA, and prestroke exposure and use of each language, (2) modeling their L1 and L2 naming impairment by implementing damage to the semantic and phonetic systems of the neural network simulating the bilingual lexicon of each individual and (3) retraining the neural network to simulate treatment effects in each language when therapy is provided in just one language. These simulations demonstrated that BiLex can accurately capture post-stroke lexical access impairment and treatment outcomes in each language, thus validating its use to predict treatment response in BWA.[30]

The primary objective of this randomised control trial (RCT) is to determine the capacity of our computational model to predict language therapy outcomes in a sample of 48 Spanish–English BWA. Specifically, we aim to examine whether BiLex can accurately determine the optimal treatment language that will result in maximum gains in the treated and the untreated language for each patient by comparing simulated treatment effects when treatment is provided in one language versus the other. To this aim, we will randomly allocate participants to a model-prescribed experimental group receiving therapy in the language defined as optimal by the model, or a model-opposite control group receiving therapy in the opposite language. We hypothesise that BWA treated following the model's recommendation will show greater treatment outcomes than those who are trained opposite to the model's recommendation. This RCT involves semantic feature analysis[31] treatment with proved effectiveness in BWA.[8 11 32] Because this treatment entails strengthening semantic representations which are shared across languages in bilinguals[33] we expect positive treatment effects in the treated language and some degree of generalisation to the untreated language as shown in prior research.[8–13] Thus, we will examine treatment effects on primary and secondary treatment outcome measures in the treated language and will employ the same metrics to determine the degree of cross-language transfer effects in the untreated language. Secondary objectives include providing a computational explanation for different patterns of cross-language generalisation arising from facilitation and inhibitory mechanisms[11 32] and expanding the architecture and function of the model to Chinese–English to account for a different bilingual combination.

METHODS

Design

The study will employ a prospective parallel-group, double-blind, RCT design to test the efficacy of the BiLex computational model to predict treatment gains in both languages in each BWA. Figure 1 depicts all procedures for participants throughout the trial, which are reported here following the Standard Protocol Items: Recommendations for Interventional Trials guidelines for RCTs.[34]

Patient and public involvement

Patients or the general public were not involved in the design, conduct, reporting or dissemination of the research.

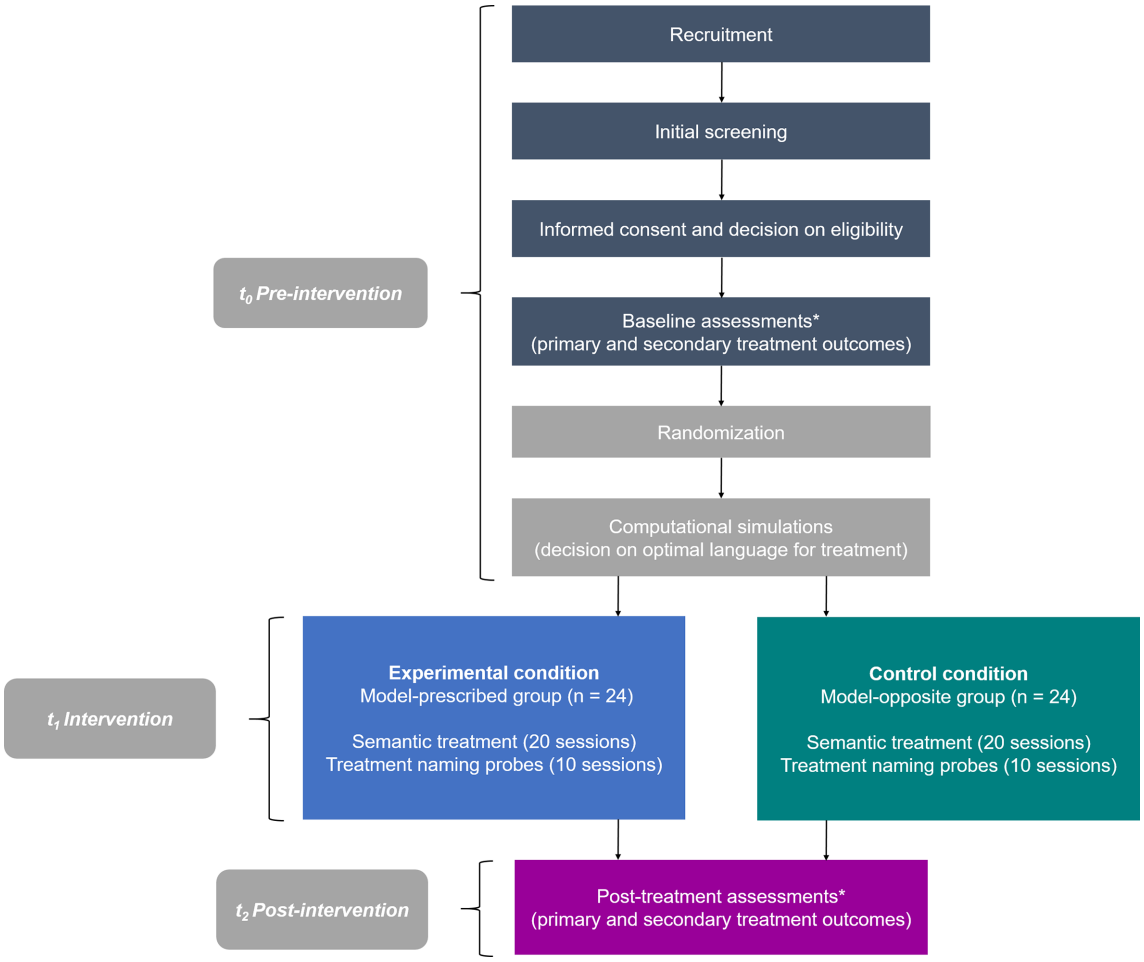

Figure 1 Flowchart for participants throughout the study. Asterisks denote three assessments for primary outcomes during pre-treatment and post-treatment.

Participants and setting

Participants will be 48 Spanish–English BWA equally distributed across the model-prescribed and model-opposite groups (see the Statistical analysis section for sample size estimation). Participants will be actively recruited from hospitals in Massachusetts, and bilingual research and rehabilitation centres in other regions including San Francisco, California and Austin, Texas (see ClinicalTrials.gov for information about recruitment locations). Participants will also be recruited via referrals from neurologists, speech and language pathologists and other clinicians across the USA. The study will also be advertised so that potential participants can self-refer to the research team.

Eligibility and enrollment

As part of the initial screening process, recruiting sites will request information relevant to the eligibility of BWA to the referring clinician, institution, or caregivers to determine whether they are potential study participants. Trained clinicians will provide detailed information about study participation in the patient's preferred language, explain all related procedures and obtain written consent from patients willing to participate in the RCT. Full eligibility will be determined according to inclusion and exclusion criteria (table 1) after clinicians obtain the signed consent form during the first session. To ensure eligibility, the patient and caregiver will fill out a patient history form collecting information about demographics, stroke history, medical history, current medications and an overview of current language difficulties. They will also complete a Language Use Questionnaire (LUQ)[35] to examine the patient's bilingual background and a HIPAA release form to request stroke-related medical records from treating physicians, hospitals and rehabilitation facilities to ensure fulfilment of inclusion criteria (table 1). Prospective enrollment will last 5 years between 2018 and 2023 and will include 32 participants in addition to the planned sample size to minimise potential patient attrition.

Baseline pre-treatment assessments

Initial assessments[29 35–49] will be conducted during 2 hour sessions over 7 days or more if needed (table 2). Variation in number of sessions across patients is expected reflecting individual differences in pace, degree of severity and time availability.

Stimuli

All patients will be tested on the Item Selection Naming Test (ISNT) developed to examine naming ability on a

Table 1 Inclusion and exclusion criteria

Inclusion criteria	Exclusion criteria
► Male or female (age range between 18 and 85)	► Premorbid history of a speech/language disorder
► Any number of years of education	► Proficiency in more than just Spanish and English
► Bilingual speaking Spanish and English with any degree of language proficiency prior to stroke	► Overt, behaviorally noticeable, attentional limitations that interfere with completing the experimental tasks
► Aphasia* secondary to a left hemisphere stroke (diagnosed by a neurologist on the basis of clinical CT/MRI imaging or medical reports)	► Active medical disease that may compromise participation†
► Stroke leading to aphasia must have occurred at least 6 months prior to the initiation of study participation	► Diagnosed with mental illness other than active depression
► Naming deficits must be present with concurrent lexical/semantic impairment	► Neurological condition other than that which resulted in aphasia
► Visual and auditory acuity must be sufficient for all assessment and treatment procedures	► Current medications known to exert significant effects on cognitive processes (ie, neuroleptics, steroids, anticholinesterase inhibitors and others)
► Ability to understand and follow study procedures for the entire length of the study	► Current drug or alcohol use or dependence that, in the opinion of the principle investigator, may interfere with adherence to study requirements
	► Inability or unwillingness of individual to give written informed consent

*Aphasia severity and clinical profile will be determined using the WAB-R English[44] and Spanish.[45]
†Participants will be excluded if acute exacerbations of stable disease occur.
WAB-R, Western Aphasia Battery–Revised.

large set of 273 unique items of 13 broad semantic categories, including semantically related and unrelated noncognates (ie, <50% cross-language phonemic overlap). The test will be administered in each language separately to identify items that each patient fails to retrieve correctly in both languages. These items will be used to create six stimuli sets per patient, three sets in English and three sets of their Spanish direct translations. Each set will include 15 items, for a total of 45 unique words and their translations (table 3). SET 1 will include treated words and their translations, SET 2 will include words that are semantically related to those in SET 1 (ie, semantic category coordinates) and their translations, and SET 3 will include words that are semantically unrelated to those in SET 1 and their translations. Word pairs in SET 1 and SET 2 will be sister terms as verified on the online lexical database Wordnet available at https://wordnet.princeton.edu/. If the ISNT does not provide enough items that the patient fails to name in both languages to create the stimuli sets, clinicians can select items from the 60-item naming screener[29] or test the patient on an additional naming screener with novel items. All stimuli sets will be unique to each patient reflecting their particular naming deficits across both languages.

Naming probes

All 45 pictures corresponding to the 45 words in English and their Spanish translations will be included in baseline naming probes, treatment naming probes and post-treatment naming probes, which are similar in structure but are labelled differently according to the time point of their administration. Three baseline naming probes will be conducted on three different sessions (maximum gap of 1 week between probes) prior to treatment to establish individual baseline naming performance. Baselines will be considered stable when they show no more than 30% variability.[8] Two additional baseline probes will be administered for patients showing larger performance variability. Ten treatment naming probes will be conducted, one at the beginning of every second session of a week of treatment. Three post-treatment naming probes will be completed after treatment on three different sessions (maximum gap of 1 week between probes) to examine the maintenance of treatment effects once treatment is interrupted.

Each naming probe will present all 45 pictures with a language-blocked design (ie, all items tested in one language first and tested again in the other language) and language order will be counterbalanced across sessions. The order of presentation will be pseudo-randomised to avoid the sequential order of items from the same semantic category. Pictures will be presented on a computer and clinicians will record responses without providing accuracy feedback. All naming probes will credit one point for each response produced in the target language that involves the target word, an acceptable variation of it or a production that deviates from the correct response by one phoneme. Incorrect responses including instances of language mixing will be analysed following an

Table 2 Primary and secondary outcome measures

Outcome measure	Assessments	Language of administration	Measurement variable
Primary: lexical access	Naming probes*	Spanish and English	% accuracy (ES and PC)
Secondary: lexical access	Boston Naming Test (BNT)[36 37]†	Spanish and English	% accuracy
	Category generation task[38]†	Spanish and English	Number of words produced % accuracy
	Controlled Oral Word Association Test (COWAT)[39 40]†	Spanish and English	Number of words produced % accuracy
Secondary: non-verbal semantic knowledge	Pyramid and Palm Trees(PAPT): picture modality[41]†	Preferred language	% accuracy
Secondary: lexical-semantic processing	Psycholinguistic Assessments of Language Processing in Aphasia (PALPA)[42 43]†: subtests of spoken and written word–picture matching, auditory and written synonym judgments and word semantic association	Spanish and English	% accuracy
Secondary: reading	Psycholinguistic Assessments of Language Processing in Aphasia (PALPA)[42 43]†: subtests of letter-length readingc[41 42]	Spanish and English	% accuracy
Secondary: aphasia severity	Western Aphasia Battery–Revised (WAB-R)[44 45]†	Spanish and English	WAB-AQ Score
Secondary: production and comprehension	Bilingual Aphasia Test (BAT–B)[46]†	Spanish and English	% accuracy
Secondary: translation	Bilingual Aphasia Test: (BAT–C)[46]†	Spanish and English	% accuracy
Secondary: discourse	Aphasia bank[47]†	Spanish and English	CIU per minute
Secondary: perception of patient's communication abilities	Communicative Effectiveness Index (CETI)[48]†	Preferred language	Total score
Secondary: executive function	Cognitive Linguistic Quick Test (CLQT)[49]†	Preferred language	% accuracy

*Naming probes designed for each patient including treated and untreated items.
†Standardised tests of language and/ or cognition.
CIU, content information units; ES, effect size; PC, percentage change.

error scoring system reported elsewhere.[50] Naming probe scores will be recorded on a spreadsheet that graphs the trend lines across baseline, treatment and post-treatment naming probes for each patient.

Primary outcome measures

Naming probes will serve as primary treatment outcomes (table 2). The percentage of accuracy on treatment naming probes will serve as the primary dependent measure to monitor change over time in word finding ability during treatment. Specifically, SET 1 will allow assessing treatment effects on treated words and untreated translations, SET 2 will allow evaluating treatment effects on semantically related untreated words and translations and SET 3 will allow examining change in unrelated/untreated control words (table 3). Additionally, the percentage of accuracy on baseline and post-treatment naming probes will help evaluating change in word finding ability post-treatment relative to pre-treatment in both languages (see the Statistical analysis section).

Secondary outcome measures

Treatment steps encompassed in our intervention may target additional language domains other than naming

Table 3 Example of the six stimuli sets created to measure the effects of treatment in each language

Stimuli set	Description	Outcome measured	Example *
SET 1: treatment words	15 words used as treatment items	Treatment gains in the treated language	SET 1— English 'apple'
SET 1: treatment translations	15 translations of treatment words	Cross-language generalisation to untreated translations	SET 1— Spanish 'manzana'
SET 2: semantically related words	15 words semantically related to treatment words in SET 1 (ie, category coordinates)	Within-language generalisation to semantically related untreated words in the same language	SET 2— English 'orange'
SET 2: semantically related translations	15 translations of the semantically related words in SET 2 (ie, category coordinates)	Cross-language generalisation to semantically related words in the untreated language	SET 2— Spanish 'naranja'
SET 3: semantically unrelated control words	15 words semantically unrelated to treatment words in SET 1	Reference set for gains in the treated language	SET 3— English 'ship'
SET 3: semantically unrelated control translations	15 translations of the semantically unrelated words in SET 3	Reference set for gains in the untreated translations	SET 3— Spanish 'barco'

All word sets for each patient will be similar in average word frequency and length.
*Example of sets created for a patient for whom English is the treated language.

ability. Although BiLex will not predict treatment responsiveness as measured by secondary outcome measures, treatment effects will also be reported using secondary outcome measures of semantic knowledge, lexical-semantic processing, comprehension, single word and sentence production, and executive function (see table 2 for standardised tests and specific measurement variables). In all cases, pre-treatment and post-treatment will be the specific measurement time points of interest for analysis. Change from pre-treatment to post-treatment will be considered the metric for participant-level analysis and mean group values will be computed as method of aggregation. All outcome measures are considered reproducible, relevant and sensitive to measure change in post-stroke aphasia and have been used in prior bilingual aphasia rehabilitation research.[6 8–17 21 32]

Computational simulations

The BiLex computational model[29] will be employed to conduct 50 independent simulations per patient, with each simulation leading to a predicted treatment outcome in each language separately (ie, percentage of items incorrectly named in the baseline naming probes that were correctly named after treatment). Next, BiLex will compare treatment outcomes (eg, outcomes in the treated and untreated language when treatment is provided in English, vs outcomes in the treated and untreated language when treatment is provided in Spanish) by adding up the treatment effects in both treated and untreated language for each simulation and computing the number of simulations that predict a larger overall effect for treatment in English versus treatment in Spanish. This comparison will enable BiLex to predict the optimal language for treatment (ie, the language with ≥10% of all 50 simulations showing superior outcomes when treatment is provided in one language over the other). When BiLex has no clear recommendation (ie, the difference in simulated outcomes for treatment in English vs treatment in Spanish is <10%), it will compute the median treatment effect for all 50 simulations for each treatment language and will predict the optimal language of treatment using a different criterion (ie, the language showing a median difference ≥5% relative to the other language will be determined as being more optimal for treatment). All simulations will use the same input data for each patient including their age at testing, LUQ metrics including L2 AoA, lifetime exposure to each language, and prestroke language use to simulate prestroke naming abilities in each language (the use of these LUQ metrics in simulations will be similar to our previous work described in detail elsewhere.[29] Also, their age at stroke onset, and poststroke semantic (ie, Pyramid and Palm Trees Test[41] Scores) and naming performance (ie, scores on the Boston Naming Test,[36 37] 60-item naming screener[29] and baseline naming probes) in each language will be used to simulate post-stroke language deficits and treatment response as previously reported.[30]

Randomisation and group allocation

The laboratory manager will use a pre-defined non-stratified randomisation module in REDCap[51] (concealed to all other team members involved in the RCT) to randomly assign each patient to one of the 25 slots available for each intervention group in the randomisation sequence: the experimental model-prescribed group receiving language therapy in the optimal language defined by the computational model, or the control model-opposite group receiving therapy in the opposite non-prescribed language (allocation ratio of 1:1). Following the order of enrollment and on completion

of baseline assessments, the computer scientist of the team will receive the relevant input data of each patient to conduct the simulations that will define his or her optimal language of treatment, and will inform the laboratory manager about the treatment language chosen for each patient. Next, the laboratory manager will enter the patient ID code into the randomisation module to assign the patient to the model-prescribed or the model-opposite group and will inform the treating clinician of the language chosen for treatment (eg, if the predicted language for a given patient is English and the randomisation module returns 'model-opposite group', the laboratory manager will inform the treating clinician that Spanish will be the treatment language and the patient will receive treatment accordingly). Because current clinical services for BWA depend on geographic location, availability of bilingual speech-language pathologists, the patient's language proficiency and personal preference, it is surmised that when all these factors are considered, patients may receive English or Spanish treatment at a 50% probability. Our experimental design reflects this 50% probability of receiving treatment in any one language by using a fully randomised allocation to the model-prescribed or model-opposite group. The same procedure schedule (ie, 10 weeks of treatment with testing prior and after treatment) will be followed by each group and the language targeted in treatment (ie, prescribed vs opposite) will be the only characteristic differing across groups. For consistency with the computational procedure, treatment will only be provided in one language whether or not cross-language generalisation is observed.

Blinding

This is a double-blind RCT. The researcher conducting the computational simulations determining the optimal language of treatment and the clinicians conducting assessments and treatment will be blind to each patient's group assignation. The principal investigator and other researchers will be blind to each patient's group allocation. Patients will know the language chosen for treatment but they will remain blind to their treatment group allocation.

Intervention

Patients will receive a semantic-feature analysis based treatment[31] with large evidence of effectiveness in improving word retrieval deficits in BWA.[8 11 32] Treatment will be provided over 20 sessions, two times per week during 2 hour sessions, and will be delivered in person or via videoconference following our teletherapy protocol if patients cannot attend one of the recruitment sites. Treatment will be conducted on a laptop computer using the Qualtrics survey software available at https://www.qualtrics.com. A Qualtrics survey will be developed for each patient's treatment session presenting the 15 treated items in the targeted language in randomised order. All sessions will be conducted using the Zoom communication software available at https://zoom.us/ to enable video-recording. Videos will be reviewed later for treatment fidelity and scoring of patient responses. Treatment steps emphasising semantic feature attributes of each treated item (table 4) will be guided by a trained clinician who will use a treatment key to provide response feedback (including requests to provide responses in the target language when instances of language mixing occur), and annotate all verbal responses.

Other interventions

Patients will not be engaged in other language interventions during treatment. No strict criteria are set for other concomitant interventions. Allowed interventions include the use of medications not known to exert significant effects on cognition.

Evaluation of progress and adherence to treatment

Individual progress during treatment will be discussed in weekly research meetings by reviewing treatment probe scores and corresponding trend lines across treatment sessions. To promote participant retention, all patients will have the opportunity to discuss their progress and concerns during each session. Adherence to treatment

Table 4 Treatment steps

Treatment steps *	Description
Step 1. Naming	Patient names the picture of the treated item on the screen. Clinician provides feedback.
Step 2. Feature selection	Patient reviews a list of 15 semantic features and selects the ones that do not apply to the treated item. Patient classifies the features that apply to the treated item according to their function, general characteristics, physical attributes, location and superordinate category.
Step 3. Association	Patient provides an association between the trained item and another item/ concept and explains how the two are related.
Step 4. Yes/ no questions	Patient reviews a list of 15 semantic features and determines whether they apply or do not apply to the treated item.
Step 5. Naming	Patient names the picture of the treated item on the screen.
Step 6. Sentence	Patient creates a short sentence with the trained word.

*Clinician provides feedback in all treatment steps (ie, correct response following patient's performance).

(ie, at least 80% of sessions attended) will be assessed in weekly research meetings. Patients who cannot make the required study visits will be withdrawn from the study.

Treatment fidelity

Clinicians will receive training to provide therapy according to our protocol. Treatment fidelity will be conducted by two trained bilingual research assistants who will watch treatment videos to determine the percentage of treatment steps performed correctly by clinicians on 25% of treatment sessions of each patient. Likewise, reliability will be conducted on 25% of the baseline and treatment naming probes and will also be performed on error analysis. Intra-class correlation statistics will be calculated for interrater reliability.

Post-treatment assessments

Three post-treatment naming probes and all primary and secondary outcome measures (table 2) will be administered to measure potential changes after treatment. Post-treatment assessments will be conducted during 2 hour sessions over 7 days or more if needed. Patients who drop out of the study for practical reasons (ie, transportation or health issues) will complete post-treatment assessments in an accelerated schedule.

Data management and confidentiality

Personal information and assessment and treatment data will be stored using REDCap, a secure web-based software platform for research studies.[51] Also, REDCap will allow to track study candidates who are contacted but not enrolled and record reasons for ineligibility and non-participation. To ensure confidentiality, patients will have a physical folder containing their signed consent form, test forms and other study materials stored in a locked cabinet at Boston University. Video-recorded sessions will be kept in password-protected computers files. Patient records will be identified only by a patient identification number to be used in all computer entry programs. Access to patient data in both paper and digital form will be restricted to authorised study research staff only.

Statistical analysis

The predictive capacity of the computational model will be confirmed when it demonstrates an accurate prediction of both optimal rehabilitation and sub-optimal rehabilitation outcomes. To this aim, we will compare primary treatment outcomes for the model-prescribed group with those of the model-opposite group using 2 (Group)×2 (Time-point: pre and post-treatment) mixed effect models for repeated measures Analysis of variance (ANOVA) on effect size (ES) and percentage change (PC) measuring treatment effects, using a random intercept and random slope for time model to account for within-subject variation over time.

ES and PC allow measuring the extent to which changes in primary treatment outcomes from prior to after treatment are statistically reliable. ES will be calculated as ((mean of post-treatment probes-mean of baseline probes)/SD of baseline).[52] In our previous work using the same treatment for BWA,[32] ES have ranged from 0 to 16.5 (Mean=9.0) for trained items, from 0 to 13.7 (Mean=3.1) for the semantically related items and from 0 to 10.6 (Mean=2.18) for translations. Control set effect sizes ranged from 0 to 6.8 (Mean=1.7). Thus, the average effect size based on our preliminary data is 3.4. Our RCT sample size (n=48) will have at least 80% power to detect large effect sizes (ie, ES ≥0.8) indicating significant change in primary treatment outcomes.[52] PC will be calculated as (mean of three post-treatment probes–mean of three pre-treatment probes) and will also be used to measure change in secondary treatment outcomes after treatment. Additionally, each model-prescribed patient will be matched to a model-opposite patient using propensity sampling in terms of age, L2 AoA, language exposure and degree of impairment in the two languages and we will compare differences between patient pairs in their time-series slope using simple paired t-tests. We expect that model-prescribed patients will show significantly higher slopes than the model-opposite patients. While we plan to complete a per-protocol analysis, secondary analysis of the data will also include intention-to-treat analysis to avoid selection bias.

Data monitoring and interim analyses

The Data and Safety Monitoring Board will include an external statistician who will review the treatment probe data and any available treatment outcome data on a quarterly basis. If no significant positive ES are reported for the first 12 patients, the statistician and the Data and Safety Monitoring Board will review the findings to change course in the experimental design. However, we do not intend to complete an interim analysis on the data as it influences the unsealing of the randomised patient assignment.

Quality monitoring

Clinicians collecting outcome measures will receive training on assessments and therapy including administration procedures, scoring criteria and data storing, and will demonstrate independent understanding and execution of correct study procedures. Important protocol modifications and protocol adherence will be discussed on weekly follow-up team supervision meetings.

Safety parameters and adverse events

No physical, social or legal risks are expected to arise as a direct consequence of trial participation (eg, intervention-related harms). No adverse event will be a solicited event. Adverse events will be recorded regardless of their relationship to the study intervention. The main adverse event is the occurrence of a second stroke or medical complications unrelated to their aphasia. In this case, participation will be terminated early so that patients can manage their health issues.

DISCUSSION

The last few decades of bilingual aphasia research have shown that BWA require assessment and treatment options that respond to their specific individual patterns of deficits and communicational needs in each language while considering their individual history of bilingualism. While there is evidence that both languages show potential for recovery in BWA,[5–7] the crucial question regarding which language should be targeted in their treatment to observe maximum therapy outcomes remains unanswered. This RCT will address this question employing a computational modeling approach[29] that accounts for factors known to determine recovery in BWA including prestroke bilingual background and poststroke impairment to predict individual treatment response.

The results of this RCT will contribute to the field of bilingual aphasia rehabilitation in several ways. Findings from this study may be of great clinical significance as they may inform clinicians which language to target in treatment and how much cross-language generalisation can be expected by treating one language versus the other considering the particular characteristics of each patient. If therapy in either language is available and cross-language transfer effects are to be expected, choosing the optimal language for therapy becomes crucial to improve treatment cost-benefits for BWA. If therapy is available in just one language, knowing which language to target in treatment can inform monolingual clinicians about when to seek assistance to treat a patient in the other language. Given its large patient cohort, this RCT may also inform about patterns of treatment-induced recovery according to specific bilingual profiles (ie, treatment response, extent of recovery and cross-language transfer for bilinguals with balanced vs unbalanced proficiency, and according to degree and type of impairment across languages). Our study may also contribute to our understanding of the mechanisms of language recovery in BWA during rehabilitation. Because parallel activation of the two language representation systems exists in bilinguals,[53] facilitation and inhibition mechanisms may provide an explanation for the presence or absence of generalisation effects in BWA[11 32] although they are not well understood. Retrospective computational simulations with patient treatment data may help understanding such mechanisms via a systematic variation of (1) facilitatory spreading activation of treated to untreated language representations within the treated language and across languages or (2) lateral inhibition between specific representations within and across languages. These simulations may provide a mechanistic account for the presence and absence of generalisation, and interference effects reported in the literature.[8–17] Another factor that may influence individual variation in treatment response across BWA is language distance.[54] Although our computational model simulates lexical access in Spanish–English speakers, it can expand its simulations to other language combinations by changing specific language representations while preserving its essential architecture. Thus, BiLex may predict treatment outcomes for BWA with different language combinations.

Finally, this work can contribute to the development of optimal rehabilitation protocols for BWA as comprehensively described therapy protocols for this population are rather uncommon. This protocol describes a RCT looking at the effects of semantic treatment for word finding deficits in the largest sample of BWA to be studied yet. Thus, it can guide future rehabilitation studies on study feasibility, documented procedures, selection of outcome measures and methods for delivering therapy. This RCT involves a few limitations including the fact that treatment targets specific concrete nouns and may present only limited generalisation to other word types and language measures, and the absence of follow-up assessments. Future research should address simulations of outcomes for different types of language therapy and include long-term assessments to evaluate the maintenance of treatment effects. To conclude, the present work will provide evidence of the feasibility of employing computational models to predict rehabilitation outcomes for BWA and will demonstrate the validity of using this approach in RCT studies. In the future, computational models may help optimise therapeutic interventions and reduce disability in BWA by considering an accurate characterisation of bilingual background and language deficits after brain damage.

Acknowledgements We would like to thank all participants for their time and commitment to the present study. We thank our collaborators Risto Miikkulainen and Uli Grasemann who provide their expertise in computational modeling that greatly assists the present study. We would also like to thank all clinicians involved in patient referral, recruitment and data collection and the research assistants providing support with treatment fidelity procedures.

Contributors SK conceptualised and designed the study, lead the project team and provided trial management and data analysis expertise. CP wrote the manuscript, assessment and treatment protocols, developed training procedures for clinicians and coordinated multisite data collection and patient referrals. CP and MD wrote the final study protocol. CP and MS implemented the treatment fidelity protocol and oversaw all related procedures. CP, MS, EC are currently in charge of ongoing recruitment, assessment and treatment procedures. MD and NM are in charge of patient randomisation, data handling and IRB procedures. All authors have reviewed and approved the manuscript.

Funding This work was supported by the National Institute on Deafness and Other Communication Disorders of the National Institutes of Health [grant U01DC014922] awarded to Swathi Kiran.

Disclaimer The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Competing interests SK serves as a consultant for The Learning Corporation with no scientific overlap with the present study. Claudia Peñaloza, Michael Scimeca and Erin Carpenter are currently employed by Boston University under NIH/NIDCD grant number U01DC014922 awarded to SK.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

ORCID iDs
Claudia Peñaloza http://orcid.org/0000-0003-1200-5936
Maria Dekhtyar http://orcid.org/0000-0002-7365-6508
Michael Scimeca http://orcid.org/0000-0002-1683-7767
Erin Carpenter http://orcid.org/0000-0002-7958-6803
Nishaat Mukadam http://orcid.org/0000-0001-9674-3519
Swathi Kiran http://orcid.org/0000-0003-1586-8913

REFERENCES

1 Hilari K, Needle JJ, Harrison KL. What are the important factors in health-related quality of life for people with aphasia? A systematic review. *Arch Phys Med Rehabil* 2012;93:S86–95.
2 Fabbro F. The bilingual brain: bilingual aphasia. *Brain Lang* 2001;79:201–10.
3 Paradis M. *A neurolinguistic theory of bilingualism*. Amsterdam: John Benjamins Publishing, 2004.
4 Brady MC, Kelly H, Godwin J, *et al*. Speech and language therapy for aphasia following stroke. *Cochrane Database Syst Rev* 2016;24:CD000425.
5 Ansaldo AI, Saidi LG. Aphasia therapy in the age of globalization: cross-linguistic therapy effects in bilingual aphasia. *Behav Neurol* 2014;2014:1–10.
6 Faroqi-Shah Y, Frymark T, Mullen R, *et al*. Effect of treatment for bilingual individuals with aphasia: a systematic review of the evidence. *J Neurolinguistics* 2010;23:319–41.
7 Kohnert K. Cross-language generalization following treatment in bilingual speakers with aphasia: a review. *Semin Speech Lang* 2009;30:174–86.
8 Edmonds LA, Kiran S. Effect of semantic naming treatment on crosslinguistic generalization in bilingual aphasia. *J Speech Lang Hear Res* 2006;49:729–48.
9 Goral M, Levy ES, Kastl R. Cross-language treatment generalisation: a case of trilingual aphasia. *Aphasiology* 2007;103:203–4.
10 Kiran S, Iakupova R. Understanding the relationship between language proficiency, language impairment and rehabilitation: evidence from a case study. *Clin Linguist Phon* 2011;25:565–83.
11 Kiran S, Roberts PM. Semantic feature analysis treatment in Spanish–English and French–English bilingual aphasia. *Aphasiology* 2010;24:231–61.
12 Kurland J, Falcon M. Effects of cognate status and language of therapy during intensive semantic naming treatment in a case of severe nonfluent bilingual aphasia. *Clin Linguist Phon* 2011;25:584–600.
13 Miertsch B, Meisel JM, Isel F. Non-treated languages in aphasia therapy of polyglots benefit from improvement in the treated language. *J Neurolinguistics* 2009;22:135–50.
14 Galvez A, Hinckley JJ. Transfer patterns of naming treatment in a case of bilingual aphasia. *Brain Lang* 2003;87:173–4.
15 Miller Amberber A. Language intervention in French–English bilingual aphasia: evidence of limited therapy transfer. *J Neurolinguistics* 2012;25:588–614.
16 Goral M, Naghibolhosseini M, Conner PS. Asymmetric inhibitory treatment effects in multilingual aphasia. *Cogn Neuropsychol* 2013;30:564–77.
17 Keane C, Kiran S. The nature of facilitation and interference in the multilingual language system: insights from treatment in a case of trilingual aphasia. *Cogn Neuropsychol* 2015;32:169–94.
18 Peñaloza C, Kiran S. Recovery and rehabilitation patterns in bilingual and multilingual aphasia. In: *The Handbook of the Neuroscience of Multilingualism*. *John Wiley & Sons* 2019:553–71.
19 Peñaloza C, Barrett K, Kiran S. The influence of prestroke proficiency on poststroke lexical-semantic performance in bilingual aphasia. *Aphasiology* 2020;34:1223–40.
20 Kuzmina E, Goral M, Norvik M, *et al*. What influences language impairment in bilingual aphasia? A meta-analytic review. *Front Psychol* 2019;10:445.
21 Roberts PM, Kiran S. Assessment and treatment of bilingual aphasia and bilingual anomia. In: *Speech and Language Disorders in Bilinguals*. Nova Science Publishers, 2007: 109–31.
22 El Hachioui H, Lingsma HF, van de Sandt-Koenderman MWME, *et al*. Long-Term prognosis of aphasia after stroke. *J Neurol Neurosurg Psychiatry* 2013;84:310–5.
23 Watila MM, Balarabe SA. Factors predicting post-stroke aphasia recovery. *J Neurol Sci* 2015;352:12–18.
24 Fang S-Y, Zinszer BD, Malt BC, *et al*. Bilingual object naming: a Connectionist model. *Front Psychol* 2016;7:644.
25 Miikkulainen R, Kiran S. Modeling the bilingual lexicon of an individual subject. In: Príncipe JC, Miikkulainen R, eds. *Advances in self-organizing maps*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009: 191–9.
26 Zhao X, Li P. Simulating cross-language priming with a dynamic computational model of the lexicon. *Biling: Lang Cogn* 2013;16:288–303.
27 Grasemann U, Sandberg C, Kiran S, *et al*. Impairment and rehabilitation in bilingual aphasia: A SOM-based model. In: Laaksonen J, Honkela T, eds. *Advances in self-organizing maps*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011: 207–17.
28 Kiran S, Grasemann U, Sandberg C, *et al*. A computational account of bilingual aphasia rehabilitation. *Biling* 2013;16:325–42.
29 Peñaloza C, Grasemann U, Dekhtyar M, *et al*. BiLex: a computational approach to the effects of age of acquisition and language exposure on bilingual lexical access. *Brain Lang* 2019;195:104643.
30 Peñaloza C, Grasemann U, Dekhtyar M, *et al*. A neural network model of naming impairment and treatment response in bilingual speakers with aphasia. 28th annual computational neuroscience meeting, 2019. *BMC Neuroscience* 2019;20:56.
31 Boyle M, Coelho CA. Application of semantic feature analysis as a treatment for aphasic dysnomia. *Am J Speech Lang Pathol* 1995;4:94–8.
32 Kiran S, Sandberg C, Gray T, *et al*. Rehabilitation in bilingual aphasia: evidence for within- and between-language generalization. *Am J Speech Lang Pathol* 2013;22:S298–309.
33 French RM, Jacquet M. Understanding bilingual memory: models and data. *Trends Cogn Sci* 2004;8:87–93.
34 Chan A-W, Tetzlaff JM, Altman DG, *et al*. Spirit 2013 statement: defining standard protocol items for clinical trials. *Ann Intern Med* 2013;158:200–7.
35 Kastenbaum JG, Bedore LM, Peña ED, *et al*. The influence of proficiency and language combination on bilingual lexical access. *Biling* 2019;22:300–30.
36 Kohnert KJ, Hernandez AE, Bates E. Bilingual performance on the Boston naming test: preliminary norms in Spanish and English. *Brain Lang* 1998;65:422–40.
37 Kaplan E, Goodglass H, Weintraub S. *Boston naming test*. 2nd Ed. Philadelphia, PA: Lea & Febiger, 2001.
38 Carpenter E, Rao L, Peñaloza C, *et al*. Verbal fluency as a measure of lexical access and cognitive control in bilingual persons with aphasia. *Aphasiology*;10:1–22.
39 Benton AL, Hamsher K. *Multilingual aphasia examination*. 2nd Ed. Iowa City, IA: AJA Associates, 1976.
40 Peña-Casanova J, Quiñones-Ubeda S, Gramunt-Fombuena N, *et al*. Spanish multicenter normative studies (NEURONORMA project): norms for verbal fluency tests. *Arch Clin Neuropsychol* 2009;24:395–411.
41 Howard D, Patterson K. *Pyramids and Palm Trees test*. San Antonio, TX: Pearson, 1992.
42 Kay J, Coltheart M, Lesser R. *Psycholinguistic assessments of language processing in aphasia*. Hove, England: Psychology Press, 1992.
43 Valle F, Cuetos F. *EPLA: Evaluación del procesamiento lingüístico en la afasia [EPLA: Evaluation of Language Processing in Aphasia]*. Hove, UK: Lawrence Erlbaum associates Ltd, 1995.
44 Kertesz A. *Western Aphasia Battery - Revised*. San Antonio, TX: Psychological Corporation, 2006.
45 Kertesz A, Pascual-Leone García A. *Batería de afasias "Western" [Western Aphasia Battery en versión y adaptación castellana]*. Valencia, Spain: Nau Llibres, 1990.
46 Paradis J, Libben G. *The assessment of bilingual aphasia: The Bilingual Aphasia Test*. Hove, UK: Psychology Press, 1987.
47 Macwhinney B, Fromm D, Forbes M, *et al*. AphasiaBank: methods for studying discourse. *Aphasiology* 2011;25:1286–307.
48 Lomas J, Pickard L, Bester S, *et al*. The communicative effectiveness index. *J Speech Hear Disord* 1989;54:113–24.
49 Helm-Estabrooks N. *Cognitive linguistic quick test: CLQT*. San Antonio, TX: Psychological Corporation, 2001.
50 Kiran S, Balachandran I, Lucas J. The nature of lexical-semantic access in bilingual aphasia. *Behav Neurol* 2014;2014:1–18.
51 Harris PA, Taylor R, Thielke R, *et al*. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform* 2009;42:377–81.
52 Beeson PM, Robey RR. Evaluating single-subject treatment research: lessons learned from the aphasia literature. *Neuropsychol Rev* 2006;16:161–9.
53 Costa A, HEIJ WLA, Navarrete E. The dynamics of bilingual lexical access. *Biling: Lang Cogn* 2006;9:137–51.
54 Conner PS, Goral M, Anema I, *et al*. The role of language proficiency and linguistic distance in cross-linguistic treatment effects in aphasia. *Clin Linguist Phon* 2018;32:739–57.