cortex 189 (2025) 107–130

Available online at www.sciencedirect.com

ScienceDirect

Journal homepage: www.elsevier.com/locate/cortex

Single Case Report

Cross-language interaction during sequential anomia treatment in three languages: Evidence from a trilingual person with aphasia

Claudia Peñaloza [a,b,c,*], *Manuel Jose Marte* [d], *Anne Billot* [d,e,f] and *Swathi Kiran* [d]

[a] *Department of Cognition, Development and Educational Psychology, University of Barcelona, Barcelona, Spain*
[b] *Institute of Neurosciences, University of Barcelona, Barcelona, Spain*
[c] *Cognition and Brain Plasticity Unit, Bellvitge Biomedical Research Institute (IDIBELL), L'Hospitalet de Llobregat, Barcelona, Spain*
[d] *Center for Brain Recovery, Boston University, Boston, MA, USA*
[e] *Harvard Medical School, Cambridge, MA, USA*
[f] *Department of Neurology, Massachusetts General Hospital, Boston, MA, USA*

ARTICLE INFO

Article history:
Received 17 January 2025
Revised 6 May 2025
Accepted 9 May 2025
Action editor Holly Robson
Published online 4 June 2025

Keywords:
Trilingual aphasia
Anomia treatment
Cross-language generalization
Cross-language interference
Basal ganglia

ABSTRACT

Language rehabilitation research has reported mixed evidence in bilinguals with aphasia suggesting that therapy can benefit the treated language alone or additionally result in cross-language generalization to the untreated language, while cross-language interference effects are less common. However, treatment effects in multilinguals with aphasia (MWA) have been less frequently investigated, and examining cross-language interactions during therapy may help to better understand their treatment response in each language. This study reports on P1, a trilingual person with severe aphasia with extensive damage to cortical language regions and the basal ganglia, who received sequential semantic-based treatment for anomia in her L3 French, L1 Spanish and L2 English. Overall, significant treatment gains in the treated language were restricted to her L3 French, the weakest language, while her treatment response was limited across languages likely due to severe language impairment and extensive damage to the language processing network. Cross-language generalization effects were absent and P1 showed cross-language interference in her L2 English during treatment in her L3 French. Cross-language intrusions were observed between languages, more frequently in her L2 English (the least available language in treatment) than in her L1 Spanish (the strongest language). The absence of cross-language generalization and presence of cross-language interference in P1 were likely due to damage in the basal ganglia and executive deficits reflecting damage to the language control network. Severe language processing and language control impairments can hinder the balance between activation and inhibition mechanisms necessary to support response to language treatment in MWA.

* *Corresponding author*. Department of Cognition, Development and Educational Psychology, Faculty of Psychology, University of Barcelona Passeig de la Vall d'Hebron, 171, 08035, Barcelona, Spain.
E-mail address: claudia_penaloza@ub.edu (C. Peñaloza).
https://doi.org/10.1016/j.cortex.2025.05.017

1. Introduction

Aphasia is an acquired language disorder that results from damage to brain regions supporting the language processing system. Aphasia in multilingual speakers[1] is a complex phenomenon given the simultaneous activation of languages in the multilingual brain (Colomé, 2001; Costa et al., 2006), their patterns of interaction given their relative degrees of proficiency and dominance (van Hell & Tanner, 2012), and the control processes governing speech production in the appropriate language according to context demands (Green & Abutalebi, 2013; Green, 1998). In consequence, the study of treatment-induced language recovery in multilingual persons with aphasia (MWA) requires a comprehensive approach examining factors inherently related to multilingualism and the integrity of brain regions supporting language processing and control to further our understanding of rehabilitation outcomes (Peñaloza & Kiran, 2019). The present study sought to examine the effects of a semantic-based anomia treatment sequentially provided in French, Spanish and English, on the lexical retrieval abilities of a trilingual person with aphasia presenting with damage to brain regions involved in language processing and control.

In MWA, acquired brain injury may lead to different degrees and patterns of impairment across their languages affecting a variety of language processing domains including lexical access. Moreover, recovery in MWA may not always be similar across languages (Peñaloza & Kiran, 2019) and differences in post-injury language processing abilities in their native (L1), second (L2) and other languages may further reflect differences in their premorbid achieved proficiency (Peñaloza, Barrett, & Kiran, 2019). As anomia is a hallmark residual feature across all aphasic syndromes, most rehabilitation research with MWA has focused on the recovery of lexical retrieval in one or more languages (see Ansaldo & Saidi, 2014; Faroqi-Shah et al., 2010; Kohnert, 2009 for reviews and Goral et al., 2023; Lee & Faroqi-Shah, 2024 for meta-analyses). The existing evidence indicates that significant treatment gains can be generally expected in the treated language regardless of whether the treatment language is the L1 or a later acquired language (Faroqi-Shah et al., 2010; Goral et al., 2023; Lee & Faroqi-Shah, 2024). Additionally, cross-language generalization effects in the untreated language can also be observed in MWA although they are less frequent and often smaller than those observed in the treated language (Goral et al., 2023; Lee & Faroqi-Shah, 2024). Some studies have shown language specific gains with no cross-language generalization (Galvez & Hinckley, 2003; Meinzer et al., 2007; Miller Amberber, 2012) while others have demonstrated cross-language therapy gains although these benefits depend on multiple factors not well understood. Recent meta-analysis research has shown that cross-language generalization is modulated by the age of acquisition of the treatment language (Goral et al., 2023) while other factors including differences in proficiency across languages, language distance (Goral et al., 2023) and treatment language (Goral et al., 2023; Lee & Faroqi-Shah, 2024) do not appear to significantly contribute to cross-language generalization effects in the untreated language. However, it is also recognized that multiple aspects of bilingualism and the bilingual language background of MWA contribute to their response to language therapy (Goral et al., 2023). Cross-language generalization effects have been reported for cognates from L1 (Spanish) to L2 (English) (Kohnert, 2004) and for languages with close structural distance, from L3 (French) to L2 (English) but not to L1 (German) (Miertsch et al., 2009). Cross-language generalization has also been reported in MWA with comparable premorbid proficiency across languages, with L2 to L1 transfer effects from Frulian to Italian (Marangolo et al., 2009) and from English to French (Kiran & Roberts, 2010) as well as L1 to L2 transfer effects from Spanish to English (Edmonds & Kiran, 2006; Peñaloza et al., 2021). In turn, unbalanced MWA show more individual variability in the direction of cross-language generalization. Transfer effects have been reported from the premorbidly least proficient to the most proficient language including L2 to L1 generalization from English to Russian (Iakupova & Kiran; 2011) and from English to Spanish (Kiran et al., 2013), as well as from Spanish to English when both languages are acquired from birth (Edmonds & Kiran, 2006). Nonetheless, some MWA can show transfer effects from the premorbidly most proficient L1 (Spanish) to the least proficient L2 (English) (Kiran et al., 2013).

It is worth noting that cross-language interference effects have also been described in MWA as the absence of cross-language generalization (Kurland & Falcon, 2011), decreased post-therapy naming performance in the untreated language (Abutalebi et al., 2009), increased language mixing (Kurland & Falcon, 2011) and cross-language intrusions from the treated into the untreated language (Abutalebi et al., 2009; Keane & Kiran, 2015). Notably, interference effects can also affect grammatical accuracy in the untreated language at the sentence level (Goral et al., 2013). Altogether, these findings underscore the need of examining treatment effects following therapy in each language in MWA (Lee & Faroqi-Shah, 2024) while the direction in which generalization and interference effects occur in this population also deserves further research.

Importantly, positive treatment effects have been frequently observed in MWA after semantic-based treatments for anomia (Ansaldo & Saidi, 2014) leading to improvements in treated items, untreated semantically related items and their

[1] The term multilingual is used here to refer to individuals who speak two or more languages and use them regularly in their everyday lives (Grosjean, 2021).

corresponding untreated translations (Edmonds & Kiran, 2006; Kiran et al., 2013; Kiran & Roberts, 2010; Peñaloza et al., 2021; Scimeca et al., 2023). Semantic-based interventions are assumed to stimulate the conceptual properties of treated words in the semantic system to increase their activation (and that of untreated words with shared semantic features) at post-semantic levels to facilitate their lexical retrieval (Quique et al., 2019). Following theoretical models of language production in multilinguals, the semantic system spreads simultaneous activation to the lexical nodes of the speakers' languages via connections of varying strengths reflecting their relative degrees of achieved proficiency (Kroll & Stewart, 1994). Hence, treatment-induced spreading activation in the semantic network would result in increased activation of the lexical representations of treated words and their semantically related neighbors in the treated language, as well as their translations in the other speaker's untreated languages facilitating within and cross-language generalization (Kiran et al., 2013).

However, spreading activation is not the only mechanism needed for functional lexical retrieval in multilingual speakers. Inhibitory control mechanisms are required to effectively manage language co-activation to access the intended lexical representation in the target language while suppressing the non-target language to prevent interference (Green, 1998). Inhibitory mechanisms have been proposed to prevent cross-language generalization in MWA as treatment-induced increased activation in the treated language may increase the suppression of the untreated language (Goral & Lerman, 2020). Moreover, increased suppression of the untreated (often strongest) language to control its potential interference on the treated (often weakest) language during treatment may result in the lingering suppression of the untreated language after treatment explaining cross-language interference effects (Goral & Lerman, 2020). Therefore, functional language control mechanisms may be required to regulate levels of activation and inhibition in each language in a balanced manner to support cross-language generalization treatment effects in MWA (Goral & Lerman, 2020; Keane & Kiran, 2015).

Language control in multilinguals relies on a cortico-subcortical network in which the basal ganglia support the control of the two languages while keeping track of the target language (Abutalebi & Green, 2007; Calabria et al., 2018). Absence of cross-language generalization and increased cross-language interference following language treatment have been reported in MWA with basal ganglia damage (Abutalebi et al., 2009; Keane & Kiran, 2015). These findings provide evidence for negative treatment effects in MWA as resulting from damage to brain regions involved in language control networks (Peñaloza & Kiran, 2017).

Examining anomia treatment effects across the treated and the untreated languages in MWA is highly relevant as it can help us gain a better understanding of the extent to which cross-language generalization effects can be expected and the factors related to multilingualism and brain damage that may facilitate or hinder language treatment response. Crucially, single case experimental designs with treatment provided sequentially in each language spoken by MWA provide a unique opportunity to compare treatment effects across languages targeted during therapy (Kurland & Falcon, 2011). Such comparisons can help to establish if there is an optimal language to target in treatment given a specific combination of relevant factors defining individual therapy outcomes. Moreover, sequential treatment designs allow to better characterize any cross-linguistic interactions that may arise between the treated and the untreated languages and any language selection difficulties during lexical retrieval that may impact naming performance during therapy.

The present study used this approach to examine the effects of a semantic-based anomia treatment provided sequentially in the three languages (i.e., French treatment phase, Spanish treatment phase and English treatment phase) of P1, a trilingual person with post-stroke aphasia. The comprehensive examination of P1's response to treatment can contribute to gaining a better understanding of anomia therapy response in MWA in different ways. First, rehabilitation research with MWA is limited, and studies providing sequential semantic treatment in all the spoken languages of MWA have rarely been conducted. Hence, this study can help to address current gaps in knowledge regarding how lexical retrieval deficits in MWA change in response to treatment in each language and how languages interact over the course of language-specific interventions. Second, the multilingual background, degree of language impairment and lesion characteristics of P1 offer a unique opportunity to address relevant open questions in multilingual aphasia rehabilitation. P1's linguistic background combines a profile of L2 and L3 acquisition during early adolescence, high exposure and use of all languages in early adult life leading to intermediate to high proficiency across languages, with a drastic drop of L3 exposure and use for several years in her later adult life prior to stroke. Notably, P1 presented severe impairment in all three languages after her stroke. P1's profile of language background and impairment facilitates the understanding of multilingual treatment outcomes in the context of (i) L1 and L2 deficits which mainly reflect the effects of neural injury to the language processing system (rather than low pre-stroke proficiency which can play a confounding role, Peñaloza, Barrett, & Kiran, 2019) and (ii) L3 deficits which capture the effects of neural damage on a previously proficient yet attrited language. Finally, P1's brain injury to cortical perisylvian regions and the basal ganglia further allow to interpret her treatment response across languages in consideration of her damage to both the language processing and the language control networks.

Our first aim was to assess the effects of therapy provided in each language treatment phase including (i) primary treatment outcomes: direct treatment effects on treated items in the treated language (relative to control items) and cross-language generalization effects on their untreated translations in the untreated languages (relative to control translations) (ii) and secondary treatment outcomes: effects on standardized tests of lexical retrieval. Although we expected semantic therapy to lead to superior treatment gains in the treated relative to the untreated language regardless of the language targeted in each treatment phase (Goral et al., 2023; Lee & Faroqi-Shah, 2024), we also anticipated that treatment effects might be attenuated by P1's aphasia severity across all languages. We also expected limited cross-language generalization to untreated languages given the presence of basal ganglia damage (Abutalebi et al., 2009; Keane & Kiran, 2015). Our second aim was to determine whether therapy in each language would lead to overall superior gains across all target items (i.e., treated items and their untreated translations) relative to all non-target items (i.e., control items and control translations). Our third aim was to characterize P1's patterns of language selection for lexical retrieval during naming probes measuring the effects of each treatment phase (e.g.,: if accurate naming responses were restricted to English when P1 was probed in English, or if her responses experienced intrusions from the non-target languages French and Spanish when probed in English). We focused only on accurate naming responses to isolate any potential errors in language selection during lexical retrieval from incorrect responses reflecting other types of errors in lexical access. Cross-language intrusions in lexical retrieval have been reported in MWA with basal ganglia lesions (Keane & Kiran, 2015) impairing the inhibitory control of non-target languages (Abutalebi & Green, 2007). Hence, given the presence of basal ganglia damage and the intermediate-high linguistic competence of P1 in her three languages, we expected to observe cross-language intrusions in naming probes in all language treatment phases, possibly reflecting errors in language selection during word retrieval. Our final aim was to determine the presence of nonverbal executive dysfunction in P1. We expected to observe executive deficits as reported in MWA with basal ganglia damage presenting with cross-language interference effects in language therapy (Keane & Kiran, 2015).

In summary, our study departed from the assumption that MWA with neural damage restricted to the language processing network are likely to show positive treatment gains since semantic anomia therapy facilitates activation propagation within the language processing system across languages (Kiran et al., 2013). Moreover, cross-language generalization should be expected for languages with relatively close structural distance (Miertsch et al., 2009) and comparable pre-stroke proficiency across languages (Edmonds & Kiran, 2006; Kiran & Roberts, 2010; Marangolo et al., 2009; Peñaloza et al., 2021). In this scenario, treatment gains would be mainly constrained by the amount of neural damage to the language processing network reflected in the speaker's aphasia severity across languages. However, in MWA who present additional damage to the language control network such as P1, one would expect that any positive effects of semantic therapy would be largely outweighed by the detrimental effects of basal ganglia damage on the functionality of the language control network (Abutalebi et al., 2009; Keane & Kiran, 2015) leading to limited or no cross-language generalization. While this study is not suited to tease apart the facilitation effects from semantic therapy on the language processing network and the interference effects from damage to the language control network, it does seek to provide insights about the interplay between these two mechanisms in the rehabilitation of anomia in MWA. This is highly relevant since language processing and control impairments often coexist after brain injury.

2. Methods

2.1. *Participant*

P1 was a right-handed female trilingual speaker of Spanish, English and French. She completed 16 years of formal education, obtaining a bachelor's degree in business. Prior to her stroke, she worked as a real estate agent. She experienced a left hemisphere stroke in 2013 at the age of 62 and she was 68 years old (77 months post-stroke onset) at the time of study enrollment. She self-referred to our research laboratory to receive anomia therapy under a personalized protocol developed for this study. She had corrected-to-normal vision and normal hearing and demonstrated sufficient ability to understand study procedures. She did not present with severe psychiatric or neurological illness other than stroke. She provided her informed written consent to participate in this study following procedures approved by the Boston University Charles River Campus Institutional Review Board (reference number: 3309E/1927E).

2.2. *Pre-stroke language background*

P1 completed the Language Use Questionnaire (LUQ, Kastenbaum et al., 2019) which collected the following metrics for each language separately. *Age of Acquisition* (AoA) was measured as the age of L2 and L3 learning onset. *Family proficiency* represented P1's ratings on her mother, father and siblings' proficiency in each language. *Educational history* indicated the percentage of her use of each language across different educational levels. *Daily use* reflected the percentage of time P1 and her conversation partners spent using each language during weekdays and weekends. *Lifetime exposure* indicated the average percentage of time that P1

Table 1 – Multilingual background of P1 at pre and post stroke as assessed by the Language Use Questionnaire.

	Age of Acquisition (years)			Family Proficiency (%)			Educational History (%)			Lifetime Exposure (%)			Lifetime Confidence (%)			Daily Language Use (%)			Language Ability Rating (5 max)		
	L1	L2	L3	L1	L2	L3	L1	L2	L3	L1	L2	L3	L1	L2	L3	L1	L2	L3	L1	L2	L3
Pre-stroke	0	15	15	100	8.33	0	81.33	9.33	9.33	50	29	21	100	35	45	41	59	0	5	5	3
Post-stroke																43	57	0	3.25	3	1.25

L1 = Spanish; L2 = English; L3 = French. Age of Acquisition, Family Proficiency and Educational History are general metrics of language background that require only one assessment. Lifetime Exposure and Lifetime Confidence are evaluated over the entire lifetime spanning the age at stroke onset. Daily Use and Language Ability Rating are evaluated before and after stroke. Language ability rating was measured on a 5-point scale (1 = non-fluent; 5 = native fluent).

heard, spoke, and read each language over her lifetime. *Lifetime confidence* reflected her percentage of confidence in hearing, speaking, and reading each language over her lifetime. Finally, *Language ability rating* represented her self-rated level of proficiency in each language averaging her ability to speak, listen, read, write and her overall fluency. Table 1 summarizes P1's pre- and post-stroke trilingual language background on the LUQ.

P1 was born in Venezuela to monolingual Spanish parents. While living in Venezuela, she remained a Spanish monolingual speaker until age 15, completing elementary and high school education in this language. She first received English and later French instruction (a few classroom hours per week) between the ages of 15 and 18 but she did not use these languages outside of school. She moved to Quebec, Canada at the age of 21, attended college with instruction in English between the ages of 21 and 25, and used both English and Spanish when speaking to peers, while also being highly exposed to French in the community. During her college years, she received predominant exposure to French, followed by English and Spanish, while she stayed fully confident in Spanish and less confident in French and English. After completing her studies, she lived and worked in Quebec until the age of 37. During this time, she reported more use of, exposure to, and confidence in French relative to English, while Spanish was used only to speak with family and friends. After living in Canada for 16 years, she moved to the US, where she predominantly used and was exposed to English more than Spanish and French. However, she still used French to help her daughter with homework as she attended a trilingual school. Thus, between the ages of 37 and 50, her language exposure, use and confidence in English became largely superior relative to French. At the age of 50 her French was fully dropped since her daughter graduated school. Hence, between the ages 50 and 62 (time of stroke onset) English became more dominant in exposure, use and confidence relative to Spanish, and her confidence and ability in French decreased. Prior to her stroke, she would use English for work-related and leisure activities, both English and Spanish at work and with her daughter, and Spanish only to talk to relatives in Venezuela. Her pre-stroke self-rated proficiency indicated native fluency for Spanish and English, whereas her proficiency for French reflected the ability to comprehend the language and communicate primarily in concrete sentences with correct but simplified grammar. As per the LUQ, Spanish was defined as P1's native, first-acquired language (L1), English as her second-acquired language (L2) and French as her third-acquired language (L3). Spanish and English showed similar high dominance, superior to French prior to her stroke.

2.3. Post-stroke language background

P1's stroke prevented her return to work, which impacted her language use patterns. Between the ages of 62 (stroke onset) and 68 (study enrollment), her exposure to and use of Spanish became predominant relative to English, and she stayed more confident in English relative to French. Her daily language use was reduced to mainly interactions in both English and Spanish with her family at home, leisure activities in both languages and reading mainly in English. Her self-rated post-stroke proficiency for Spanish and English indicated ability to comprehend both languages and use them to communicate primarily in simple sentences, while her overall ability in French reflected non-fluent levels.

2.4. Clinical background

As per medical records, P1 experienced a stroke in 2013. Her clinical MRI scans revealed an acute left middle cerebral artery stroke with multiple infarcts. Her structural brain images were further reviewed by an experienced neurologist who determined the stroke affected the left frontal, posterior parietal and temporal lobes including the anterior temporal areas, in addition to left subcortical regions including the external capsule and the basal ganglia (Fig. 1). At hospital discharge, she presented with receptive and expressive

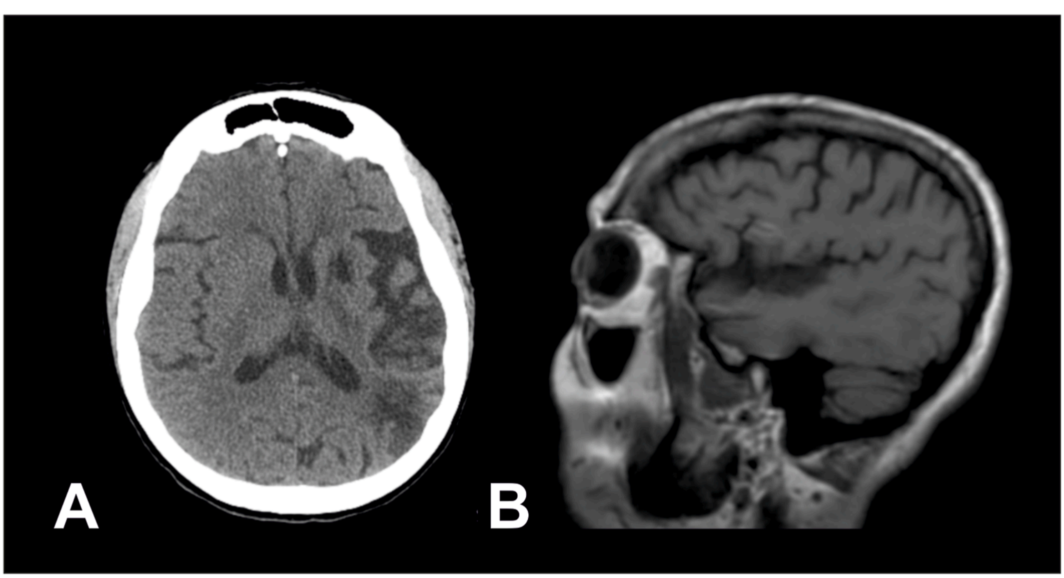

Fig. 1 – Lesion location. Axial CT (A) and sagittal MRI (B) brain scans showing P1's extensive brain lesion in the left hemisphere including frontal, temporal and posterior parietal regions, as well as the external capsule and the basal ganglia. The CT scan is shown on radiological convention (left hemisphere is displayed on the right).

aphasia and right hemiparesis. She received physical, occupational and speech and language therapy (once per week for four years). At study enrollment, she reported persistent difficulty with verbal expression, repetition and word finding, while her comprehension abilities in hearing and reading were more spared.

2.5. Baseline language and executive function assessments

P1 underwent language assessments in her L1, L2 and L3 separately. We used the Western Aphasia Battery–Revised (WAB-R) in English (Kertesz, 2006), Spanish (Kertesz et al., 1990) and a French translated version to evaluate the presence and severity of aphasia in each language. We also used the English, Spanish and Quebec-French versions of the Bilingual Aphasia Test (Paradis & Libben, 1987) to evaluate her comprehension and production abilities in each language (BAT-B) and translation abilities across languages (BAT-C). Two subtests of the Psycholinguistic Assessment of Language Processing in Aphasia (PALPA) in English (Kay et al., 1992), Spanish (EPLA; Valle & Cuetos, 1995) (i.e., PALPA 47/EPLA 45 and PALPA 48/EPLA 46) and a French translated version were employed to assess her lexical-semantic processing abilities, while the three-picture version of the Pyramids and Palm Trees (PAPT; Howard & Patterson, 1992) was used to assess her nonverbal semantic knowledge. Her lexical access was assessed in her three languages using a 60-item naming screener (Peñaloza, Grasemann, Dekhtyar, Miikkulainen, & Kiran, 2019) consisting of concrete, high frequency items, the Boston Naming Test (BNT) in English (Kaplan et al., 2001), Spanish (Kohnert et al., 1998) and French (Roberts & Doucet, 2011), and verbal fluency tasks including semantic fluency (animals, food, clothing) and phonemic fluency (Spanish: letters P, M, R; English: letters F, A, S; French: letters P, F, L) (Benton & Hamsher, 1976; Peña-Casanova et al., 2009; St-Hilaire et al., 2016). The French versions of the WAB-R and PALPA in English reported here were carefully translated in our laboratory by a native French speech and language pathologist (A.B.) and reviewed by a second native French speaker with a B.A. degree in English. Additionally, we used the Raven's Colored Progressive Matrices (RCPM) of the WAB-R (Kertesz, 2006) to evaluate her nonverbal executive function abilities (Gilmore et al., 2019). Table 2 summarizes P1's baseline clinical aphasia profile, severity and overall language and executive function performance.

All baseline assessments were conducted in person in 2-h sessions by highly proficient Spanish-English (C.P. and M.J.M.) and French-English (A.B.) bilingual researchers trained to complete the clinical assessment and treatment protocols of this study. All tests were administered on alternating English-only, Spanish-only and French-only sessions to avoid interference between languages.

2.6. Selection of treatment stimuli

We used the Item Selection Naming Test (ISNT), a comprehensive picture naming screener including 273 items (excluding cognates) across 13 semantic categories with validated semantic features for treatment (Peñaloza et al., 2020). The ISNT was administered in each language separately to identify 30 unique picture exemplars that the patient failed to name in her three languages. Of these, 15 were selected as *treated items* and 15 as *control items* (for a total of 45 treated items and 45 control items across all three languages).

The first set of stimuli (target items) included 15 *treated items* (e.g.,: English: “coat”) and their corresponding 15 *untreated translations* (e.g.,: Spanish: “abrigo”; French: “manteau”). Target items allowed to assess direct treatment effects on the treated language and indirect treatment effects on the untreated language (e.g.: cross-language generalization). The second set of stimuli (non-target items) included 15 *control items* divided into 10 monitored items and 5 unmonitored items (e.g.: English: “fox”) and their *control translations* (e.g.,: Spanish: “zorro”; French: “renard”). Non-target items allowed to examine change on untreated items that were semantically unrelated to the treated items. The 15 treated items and 15 control items in each language rotated during each treatment phase targeting one of P1's languages. For instance, given the example above for the English treatment phase, “abrigo” was the treated item in the Spanish treatment phase (“coat” and “manteau” were the untreated translations in English and French respectively) and “zorro” was the control item (“fox” and “renard” were the control translations in English and French respectively).

All treated and control items had 1 to 3 syllables and were comparable in their lexical frequency per million (CLEARPOND database, Marian et al., 2012) across languages (mean rank for treated items in English = 24.07; Spanish = 21.67; French = 23.27, $H(2)=.260$, $p=.878$; mean rank for control items in English = 22.20; Spanish = 24.60; French = 22.20, $H(2)=.334$, $p=.846$) (see the Appendix for the full list of all target and non-target items across the three languages).

2.7. Primary outcome measures

Naming probe scores constituted the primary outcome measures to evaluate the effects of each language treatment phase. Naming probes including treated items, control items, untreated translations and control translations were administered before, during and after treatment to evaluate the effects of each language treatment phase on all treated and untreated items across languages. Three *baseline naming probes* (per language) were conducted across three different sessions (9 naming probes in total) prior to treatment in each language to establish stability in naming performance (≤30% variability). Twelve *treatment naming probes* (per language) were administered to measure change in naming performance over the course of therapy, one at the beginning of every second session of a week of treatment (36 treatment naming probes in total). Three *post-treatment naming probes* (per language) were completed after treatment on three different sessions to assess treatment outcomes in the treated and untreated languages (6 post-treatment naming probes in total).[2] *Post-treatment naming probes* also served as *baseline naming probes* prior to each following language treatment phase. Baseline and post-treatment naming probes included

[2] Post-treatment naming probes were not completed after treatment in English since P1 presented with health difficulties that did not allow her to continue with additional testing at the end of this study.

Table 2 – Baseline performance of P1 on language and cognitive tests conducted in Spanish, English and French.

Language/cognitive assessment	Max. score	Spanish (L1)	English (L2)	French (L3)
Western Aphasia Battery (WAB-R)				
Spontaneous speech	20	9	7	7
Auditory verbal comprehension				
Yes-No questions	60	51	48	42
Auditory word comprehension	60	33	32	20
Sequential commands	80	35	37	28
Repetition	100	71	69	61
Naming				
Object naming	60	17	25	7
Word fluency	20	0	1	1
Sentence completion	10	0	1	0
Responsive speech	10	0	2	0
Aphasia Quotient (WAB-AQ)	100	44.1	45.5	36.8
Aphasia syndrome		Broca's	Broca's	Broca's
Raven's colored progressive matrices (RCPM) (executive functions)[a]	36	13	NA	NA
Bilingual Aphasia Test (BAT) – part B				
Pointing	10	10	10	10
Semi-complex commands	10	6	4	4
Complex commands	20	0	1	0
Verbal auditory discrimination	18	12	11	0
Syntactic comprehension	87	47	42	30
Semantic categories	5	1	3	1
Synonyms	5	2	1	1
Antonyms	5	0	0	1
Antonyms II	5	1	3	1
Grammaticality judgments	10	3	4	2
Semantic acceptability	10	7	6	3
Word repetition (repetition only)	30	20	22	25
Word repetition (judgement of nonwords)	30	23	27	17
Sentence repetition	7	6	3	2
Series	3	0	0	0
Verbal fluency (total)	U	1	0	0
Naming	20	5	4	2
Sentence construction	5	0	0	0
Semantic opposites	10	1	0	2
Derivational morphology	10	0	0	0
Morphological opposites	10	1	3	0
Description	3	1	1	1
Mental arithmetic	15	0	0	0
Listening comprehension	5	1	2	0
Reading (words)	10	5	2	1
Reading (sentences)	10	0	0	0
Reading text (comprehension)	6	0	0	0
Copying words	5	MD	4	3
Dictation (words)	5	MD	0	0
Dictation (sentences)	5	MD	0	0
Reading comprehension for words	10	8	8	6
Reading comprehension for sentences	10	7	7	3
Bilingual Aphasia Test (BAT) – part C				
Word recognition[b]	5	4 (S–E)	5 (E–S)	5 (F–S)
	5	4 (S–F)	4 (E-F)	5 (F-E)
Translation of words[b]	10	0 (S–E)	0 (E–S)	0 (F–S)
	10	0 (S–F)	1 (E-F)	0 (F-E)
Translation of sentences[b]	18	0 (S–E)	2 (E–S)	1 (F–S)
	18	0 (S–F)	0 (E-F)	0 (F-E)
Grammaticality judgements	16	5	2	5
Psycholinguistic Assessment of Language Processing in Aphasia (PALPA)				
Spoken word-picture matching (PALPA 47)	40	32	33	34
Written word-picture matching (PALPA 48)	40	31	30	24
Other language tests				
Pyramids and Palm Trees test (PAPT) (3-picture version)[a]	52	40	NA	NA
60-Item naming screener	60	3	2	0
Boston Naming test	60	5	6	2

(*continued on next page*)

Table 2 – (*continued*)

Language/cognitive assessment	Max. score	Spanish (L1)	English (L2)	French (L3)
Semantic verbal fluency				
Food	U	4	0	0
Clothing	U	2	0	0
Animals	U	2	0	0
Phonemic verbal fluency				
Letter P (French); Letter P (Spanish); Letter F (English)	U	1	0	1
Letter F (French); Letter M (Spanish); Letter A (English)	U	0	0	0
Letter L (French); Letter R (Spanish); Letter S (English)	U	0	0	0
Baseline naming probes [c]				
French Treatment phase (treated items and translations)	100%	6%[d]	11%[d]	0%[e]
French Treatment phase (control items and translations)	100%	10%[f]	13%[f]	0%[g]

[a] Nonverbal tests administered only in L1 (Spanish).
[b] Translation tests from French to Spanish (F–S); French to English (F-E); Spanish to English (S–E); Spanish to French (S–F); English to Spanish (E–S) and English to French (E-F).
[c] Baselines are reported only for the French treatment phase since this was the first language targeted in the intervention. Each baseline naming probe included 25 items (15 treated items, 10 control items) in French and their untreated translations in Spanish and English (% accuracy across three baseline naming probes is provided).
[d] Untreated translations.
[e] Treated items.
[f] Control translations.
[g] Control items. MD = Missing data not recorded during administration; NA = Not administered; U = Unlimited.

all treated and both monitored and unmonitored control items (30 items per language), whereas treatment naming probes included all treated items and monitored control items only (25 items per language).[3] For consistency in the statistical analyses, this study only considered the monitored control items and monitored control translations as they were assessed systematically in all naming probes in each language treatment phase. Hence, monitored items are referred to as "control items" and "control translations" hereafter.

Each naming probe presented the corresponding item pictures in pseudo-randomized order avoiding the sequential presentation of items of the same semantic category to minimize item-to- item carry-over effects. Naming probes were administered using a language-blocked design such that all items were tested in one language first and tested again in the other two languages separately. Also, each language was tested by a different examiner for consistency of language use during naming assessment. To facilitate the transition between languages across naming probes, P1 was briefly pre-exposed to the target language using educational or news videos which were discussed with the examiner using the target language only (e.g.,: English naming probe 1 was followed by video and conversation in Spanish prior to Spanish naming probe 1). Language order for naming probes was counterbalanced across sessions.

As regards the scoring system, naming probes credited 1 point for each clear and intelligible naming response, allowing for acceptable dialectal variations, self-corrected responses and a deviation of a phoneme in an otherwise correct response. We used two scoring methods: (i) a *language-dependent scoring method* which required naming responses to be produced in the target language being tested (e.g.,: correct responses retrieved in French on French naming probes) and (ii) a *language-independent scoring method*, which computed a naming response as accurate regardless of language selection for retrieval (e.g.,: correct responses retrieved in either French, English, or Spanish on French naming probes).

[3] Unmonitored control items were probed only at baseline and after treatment to have a control item set with reduced multiple-testing effects available in case this was necessary for statistical analyses involving control items.

2.8. *Secondary outcome measures*

Secondary outcome measures involved tests of lexical access in all three languages including the Boston Naming Test (Kaplan et al., 2001; Kohnert et al., 1998) and semantic and phonemic verbal fluency tests (Benton & Hamsher, 1976; Peña-Casanova et al., 2009; St-Hilaire et al., 2016) to determine whether P1 would show transfer effects to items not targeted during therapy. These tests were completed after each language treatment phase, except for the final English treatment phase due to unexpected health difficulties in P1. All tests were administered separately for each language on alternating single-language sessions to avoid cross-language interference (Table 3).

2.9. *Multilingual anomia treatment*

Fig. 2 provides a schematic summary of the assessment and treatment procedures involved in this study. P1 received a semantic feature-based treatment with demonstrated effectiveness for anomia in bilinguals with aphasia (Kiran et al., 2013; Peñaloza et al., 2021). The same treatment was provided sequentially in each of P1's languages across three treatment phases: French treatment phase, Spanish treatment phase and English treatment phase. The order of languages targeted in treatment followed previous evidence reporting cross-language generalization effects from the premorbidly least to most proficient language in MWA (Edmonds & Kiran, 2006; Kiran et al., 2013). In this study, we targeted French (L3) first since it was weaker in premorbid proficiency and use relative to the other languages. Spanish (L1) was targeted next since despite comparable premorbid proficiency (i.

Table 3 – P1's pre- and post-treatment raw scores on lexical retrieval tests in French, Spanish and English and change scores for French and Spanish treatment.

Language Test	Pre-Treatment (French)	Post-Treatment (French)[a]	Change from pre to post-treatment (French)	Post-Treatment (Spanish)	Change from pre to post-treatment (Spanish)
Boston Naming test (French)	2	4	**+ 2**	3	−1
Boston Naming test (Spanish)	5	7	**+ 2**	2	−5
Boston Naming test (English)	6	4	−2	2	−2
Semantic fluency (French)					
Food	0	0	= 0	0	= 0
Clothing	0	0	= 0	0	= 0
Animals	0	1	**+ 1**	0	−1
Semantic fluency (Spanish)					
Food	4	NA	NA	NA	NA
Clothing	2	NA	NA	NA	NA
Animals	2	0	−2	0	= 0
Semantic fluency (English)					
Food	0	NA	NA	NA	NA
Clothing	0	NA	NA	NA	NA
Animals	0	0	= 0	0	= 0
Phonemic fluency (French)					
Letter P	1	2	**+ 1**	2	= 2
Letter F	0	0	= 0	0	= 0
Letter L	0	1	**+ 1**	0	−1
Phonemic fluency (Spanish)					
Letter P	1	1	= 1	3	**+ 2**
Letter M	0	1	**+ 1**	0	−1
Letter R	0	1	**+ 1**	0	−1
Phonemic fluency (English)					
Letter F	0	0	= 0	0	= 0
Letter A	0	0	= 0	0	= 0
Letter S	0	0	= 0	0	= 0

Summary of raw scores for pre and post-treatment standardized measures of lexical retrieval in each language, and change scores to measure treatment effects on all three languages after the French and Spanish treatment phases (assessments after the English treatment phase were not conducted since P1 presented with health difficulties that precluded additional testing).

[a] Test scores following the French treatment phase served as pre-treatment scores to compute change scores following the Spanish treatment phase. Change scores indicate improvement (+), decreased (−) or equal (=) performance relative to the previous assessment. Post-treatment gains are marked in bold. NA = Not assessed.

e., LUQ language ability rating) relative to English (L2), Spanish showed lower premorbid language use (Table 1).

Each 2-h treatment session presented a maximum of 15 treated items in the targeted language in a randomized order across treatment sessions. A trained clinician guided P1 through 6 treatment steps which emphasized the semantic feature attributes of each treated item. Briefly, treatment steps included (i) *naming* a visual exemplar of the treated item, (ii) *feature classification* involving two components: *feature selection*: deciding whether each one of 15 semantic features applied or not to the treated item, and *feature assignment*: classifying the item selected features as describing the item's function, characteristics, physical attributes, location or superordinate category, (iii) *association*: thinking of something else that was associated with the treated item, (iv) *yes/no questions*: deciding whether or not each of 15 semantic features matched that particular item, (v) *naming* the same visual exemplar of the treated item and (vi) *sentence production*: creating a short sentence with the treated word. Each completed step was followed by feedback (i.e., correct response) provided by the clinician.

The French and English treatment phases followed a planned schedule of 24 treatment sessions (twice per week), for a total of 48 h of treatment across 12 weeks. The Spanish treatment phase with a similar planned schedule, was interrupted for 2 months as P1 presented with acute illness after 18 treatment sessions. Once recovered, she completed 3 baseline probes to determine stability in naming performance and received the last 6 treatment sessions to complete her treatment schedule. Overall, P1 received a total of 144 h/36 weeks of treatment across the entire course of the intervention, completing 12 treatment naming probes during each treatment. The full intervention was delivered at home via videoconference since P1 could not attend in-person sessions due to geographic constraints, stroke-related difficulties, and restrictions placed during the beginning of the COVID-19 pandemic. Videoconference treatment followed a well-established protocol with evidence for high reliability and equivalence of treatment gains relative to in-person therapy in bilinguals with aphasia (Peñaloza et al., 2021).

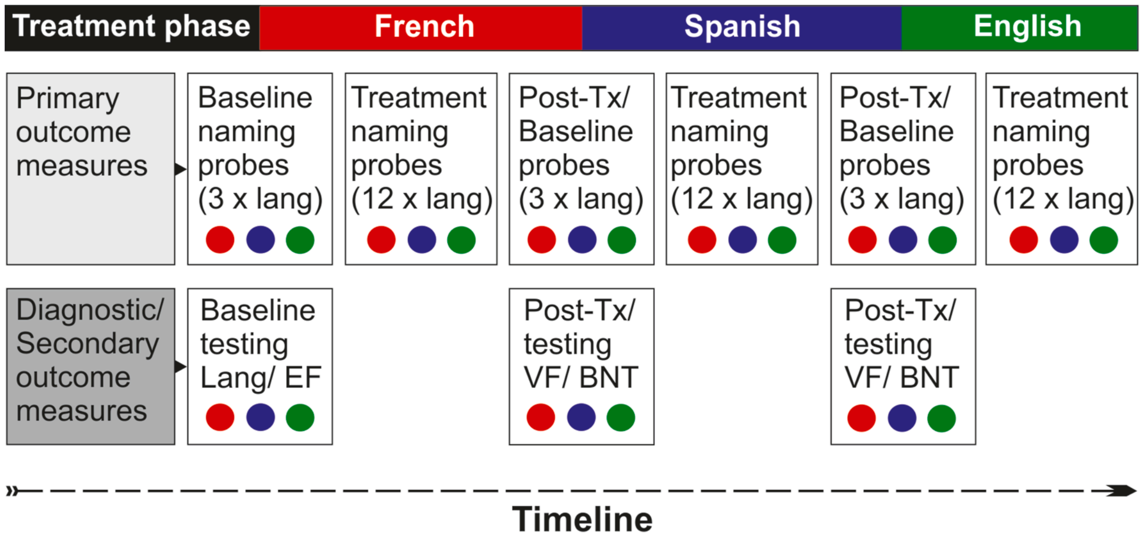

Fig. 2 – Summary of assessment and treatment procedures. P1 received anomia treatment delivered sequentially in her three languages (French treatment phase, Spanish treatment phase and English treatment phase). P1 completed 3 baseline, 12 treatment and 3 post-treatment naming probes (primary treatment outcomes) per language (colored dots: red = French, blue = Spanish, green = English). Post-treatment naming probes after French and Spanish treatment also served as baseline naming probes for the next treatment phase. P1 also completed baseline diagnostic assessments for language (Lang) and executive functions (EF), as well as verbal fluency tasks (VF) and the Boston Naming Test (BNT) (secondary outcome measures) after each treatment phase. Note that testing was discontinued after the final treatment phase due to acute health difficulties.

2.10. *Statistical analyses*

All statistical analyses were conducted using the R Statistical Software (R Core Team, 2021) unless reported otherwise. To address our first aim seeking to establish treatment effectiveness we constructed nine generalized linear mixed effects models (GLMMs) using the ‘lme4’ package (Bates et al., 2015) with a logistic link function examining the effect of treatment in each language (i.e., the direct effect of French, Spanish, and English treatment on treated items in the treated language and cross-language generalization effects on the untreated translations in each untreated language, respectively). The dependent variable was binary representing naming accuracy (language-dependent scoring method) examined across baseline, treatment, and post-treatment naming probes for each of the three languages separately. For the first three GLMMs evaluating the direct treatment effects of each language treatment phase, the dependent variable was naming accuracy on the 15 treated items and the 10 control items in the treated language. The categorical fixed effect Set, represented the status of the item (treated/control). For the next six GLMMs evaluating cross-language generalization effects, the dependent variable was naming accuracy on the 15 untreated translations and the 10 control translations in each of the untreated languages. The categorical fixed effect Set, represented the status of the item (untreated translation/control translation). In all GLMMs, Time, denoting the sequential administration of naming probes during treatment (i.e., 18 baseline, treatment, and post-treatment naming probes collected in the French and Spanish treatment phases separately, and 15 baseline and treatment naming probes collected in the English treatment phase, for a total of 51 naming probes), was a fixed effect. Further, the two-way interaction between Time and Set was included to examine the effect of treatment on treated items over the course of each language treatment. Finally, a random intercept for item was included in the model. Additionally, as done previously (Peñaloza et al., 2021), treatment effects on secondary outcome measures of lexical retrieval were assessed by computing change scores (post-treatment score − pre-treatment score) for the BNT and verbal fluency tasks in each language following each language treatment phase.

To address aim 2 seeking to examine the broad effects of treatment provided in each language separately we implemented three GLMMs which were specifically tailored to model the French, Spanish, and English treatment phases, respectively. The dependent variable was binary representing naming accuracy (language-dependent scoring method) on the target items (15 treated items and their 30 untreated translations) and the non-target items (10 control items and their 20 control translations) in all three languages (French, Spanish, and English) examined across baseline, treatment, and post-treatment naming probes. Similar to the analyses addressing aim 1, two fixed effects were factored into the models: the categorical fixed effect Set, representing the status of the item (target/non-target), and a continuous fixed effect represented by Time. Each model included each fixed effect, the two-way interaction between them as well as a random intercept for items.

To address aim 3, regarding language selection for word retrieval relative to probe language during each treatment phase, we again implemented three separate GLMMs modeling the French, Spanish, and English treatment phases, respectively. The dependent variable was binary, representing naming accuracy (language-independent scoring method) on both target items (treated items and their untreated translations) and non-target items (control items and their control translations) in the three languages examined. Three fixed effects were factored into the models. Each model included a three-level categorical fixed effect, Probe Language, denoting the language in which the naming probe was administered. This variable was sum-coded (recoded into two variables, each with levels −1, 0, and 1) enabling us to capture and compare the effects of each probed language against the grand mean (the average response across all probed languages). All models also included a three-level categorical fixed effect, Retrieval Language, representing the language used to retrieve an item on a naming probe, also sum-coded in the same manner as Probe Language. Lastly, models also included a continuous fixed effect represented by Time, denoting the numbered sequence of naming probes collected throughout each language treatment phase. The models included each fixed effect and all possible two- and three-way interactions between fixed effects, as well as a random intercept for the items.

After constructing the GLMMs, post-hoc pairwise contrasts were conducted using the "emmeans" package in R (Lenth et al., 2023) to calculate a given model's estimated marginal means and trends. These values drawn from the fitted model, represent the average response for each factor level while adjusting for the presence of other variables in the model. We performed pairwise comparisons using Tukey's method to adjust the *p*-values and correct for family-wise error rate. Lastly, we used the DHARMa package (Hartig & Lukas, 2022) to carry out residual diagnostics to assess the goodness-of-fit of our models. All GLMMs referenced yielded non-significant results across dispersion, outlier, and Kolmogorov–Smirnov diagnostic tests, suggesting good model fit. All GLMMs are summarized in Supplementary Tables 1–15 as per best-practice reporting guidelines (Meteyard & Davies, 2020).

Finally, to address aim 4, we conducted a case–control comparison to analyze the performance of P1 on the RCPM (Kertesz, 2006) using the SingleBayes_Es.exe software (Crawford & Garthwaite, 2007). This method, widely used in neuropsychology research, allowed us to employ a Bayesian approach to compare P1's RCPM score with the scores obtained by a control sample consisting of 19 healthy controls reported elsewhere (Kertesz & McCabe, 1975).

3. Results

3.1. *Effectiveness of each language treatment phase as measured by primary and secondary treatment outcomes*

We first evaluated the direct effects of each language treatment phase on the treated items relative to the control items in the treated language (Supplementary Tables 1–3). The first model evaluating the effects of the French treatment phase revealed significant main effects of Time ($\beta = .096$, $SE = .024$, 95% CI [.049, .143], $z = 4.000$, $p < .001$) and Set ($\beta = -3.713$, $SE = 1.755$, 95% CI [−7.152, −.273], $z = -2.115$, $p = .034$) on naming accuracy, indicating that both treated ($\beta = .095$, $SE = .023$) and control ($\beta = .054$, $SE = .108$) items improved over time, and that treated items were named with higher accuracy than control items (Fig. 3). However, the Time × Set interaction was not significant ($p = .710$), suggesting these relationships did not vary over the course of treatment. The model assessing the effects of the Spanish treatment phase revealed no significant effects (all p values $\geq .369$), except for the effect

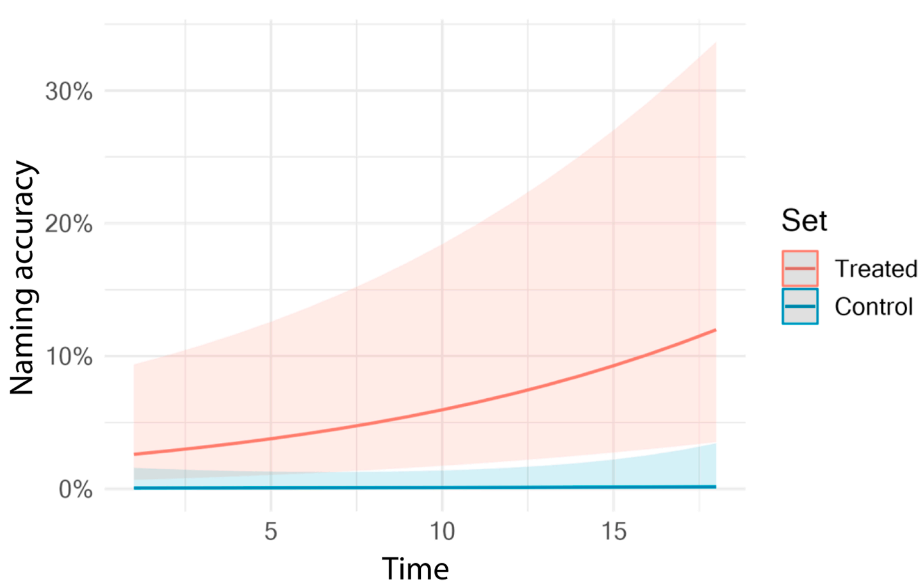

Fig. 3 – Progression of naming accuracy during the French treatment phase. Naming accuracy (language-dependent scoring method) trends for P1 across naming probes over the course of treatment (Time) during the French treatment phase contrasting treated items (red line) versus control items (blue line). 95% confidence intervals are shown as shaded regions around lines. Significant main effects of Time and Set demonstrated improvements on both types of items over the course of treatment in French, with higher accuracy noted for treated versus control items.

of Set approaching statistical significance (β = −2.018, SE = 1.042, 95% CI [−4.061, .025], z = −1.936, p = .053). Similarly, the model assessing direct treatment effects during the English treatment phase yielded no significant effects (all p values ≥ .437). Hence, only the French treatment phase led to significant general improvements in naming accuracy, while the Spanish and English treatment phases showed no significant treatment effects.

We also assessed cross-language generalization effects of each language treatment phase on each one of the untreated languages (Supplementary Tables 4–9). In models evaluating the cross-language generalization effects of the French treatment phase, a significant Time × Set interaction was observed for English (β = −.137, SE = .058, 95% CI [−.25, −.023], z = −2.361, p = .018). Post-hoc contrasts of estimated marginal trends revealed a significant contrast (p = .018) between the slopes of untreated translations (β = .000) and control translations (β = −.137). This suggests that as the treatment progressed, naming accuracy did not change for the English untreated translations of the French treated items and decreased for the English control translations of French control items (Fig. 4). All other cross-language generalization effects for Spanish and English were non-significant (p values ≥ .199).

Importantly, we found no significant Time × Set interactions reflecting cross-language generalization effects during the Spanish and English treatment phases. In the models assessing the cross-language generalization effects of the Spanish treatment phase, we only found a significant main effect of Set for English (β = −2.244, SE = 1.076, 95% CI [−4.353, −.136], z = −2.086, p = .037), indicating that overall, the English untreated translations were named with higher accuracy than the English control translations. All other effects for French and English were non-significant (p values ≥ .204). Similarly, in the English treatment models, only the main effect of Set for French approached significance (β = −4.684, SE = 2.709, 95% CI [−9.994, .625], z = −1.729, p = .084). All other effects for cross-language generalization following French and Spanish treatment were non-significant (p values ≥ .240). In summary, while no cross-language treatment benefits were detected in P1 in any of the language treatment phases, the improvement in French treated items during the French treatment phase was accompanied by cross-language interference effects as reflected by a decline in naming accuracy for the English control translations.

As regards to secondary treatment outcomes (Table 3), the French treatment phase resulted in minimal but positive score changes on the BNT and verbal fluency tasks for French and Spanish (except for decreased semantic fluency in Spanish). No change was observed on verbal fluency in English, although BNT performance decreased in this language following treatment in French. In turn, the Spanish treatment phase was followed by decreased BNT scores across all languages, with no change or minimally decreased verbal fluency in French and English.

3.2. *General effects of each treatment phase on all three languages*

Three GLMMs were conducted to assess the effects of each treatment phase (Time) on all three languages, contrasting naming accuracy for all target items (treated items and untreated translations) versus all non-target items (control items and control translations) (Set) (See Supplementary Tables 10–12).

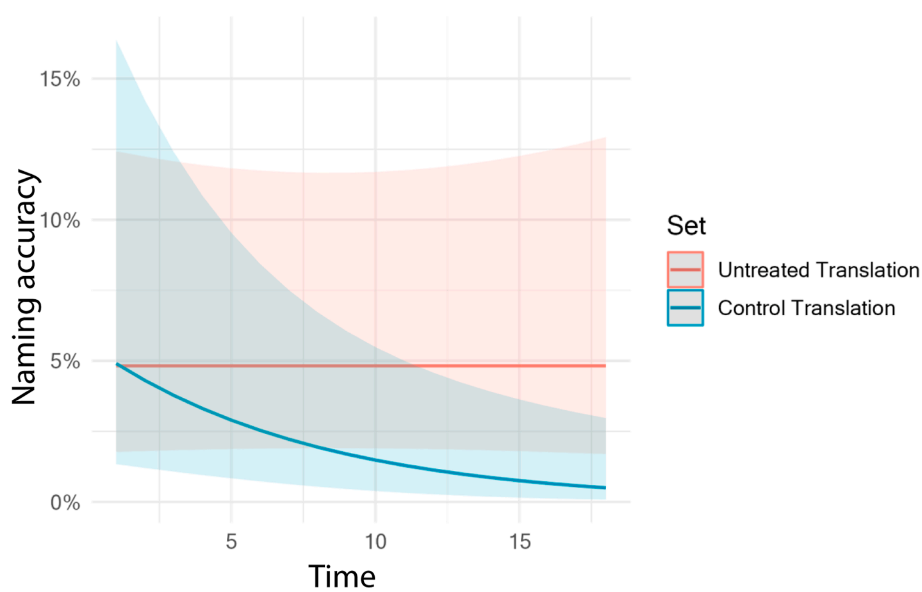

Fig. 4 – Cross-language treatment effects on English untreated translations following the French treatment phase. The significant Time × Set interaction on naming accuracy (language-dependent scoring method) for English translations during the French treatment phase is shown, contrasting untreated translation items (red line) versus control translation items (blue line). 95% confidence intervals are shown as shaded regions around lines. Naming accuracy remained unchanged for English untreated translations of French treated items (absence of cross-language generalization), while it decreased for English control translations (cross-language interference).

The first model evaluating the effects of treatment provided in French, revealed a statistically significant effect of Time on naming accuracy (β = .069, SE = .027, 95% CI [.016, .121], z = 2.573, p = .010) as treatment progressed. Although no initial significant effect of Set on naming accuracy was found (β = 1.049, SE = .661, 95% CI [−.248, 2.345], z = 1.586, p = .113), a significant Time × Set interaction emerged (β = −.245, SE = .067, 95% CI [−.375, −.115], z = −3.683, p < .001), indicating differential effects on naming accuracy for target versus non-target items over the course of treatment. Post-hoc contrasts of estimated marginal trends revealed a significant contrast (p = .0002) between the slopes of target (β = .069) and non-target items (β = −.176) suggesting that as the treatment progressed, naming accuracy improved for target items but decreased for non-target items (Fig. 5).

Contrastingly, the second model assessing the effects of treatment provided in Spanish indicated no significant effects of Time on naming accuracy (β = −.020, SE = .022, 95% CI: −.063, .024, z = −.894, p = .372). We found a significant effect of Set on naming accuracy (β = −2.100, SE = .767, 95% CI: −3.603, −.597, z = −2.738, p = .006), suggesting superior naming performance for target versus non-target items. However, there was no significant Time × Set interaction (β = .055, SE = .055, 95% CI: −.053, .162, z = .995, p = .320).

The third model evaluating the effects of the English treatment phase demonstrated no significant effects of Time (β = .049, SE = .033, 95% CI [−.016, .115], z = 1.460, p = .144) or Set (β = −1.851, SE = .973, 95% CI [−3.759, .056], z = −1.902, p = .057) on naming accuracy. The Time × Set interaction also proved non-significant (β = .057, SE = .066, 95% CI [−.072, .188], z = .870, p = .385). In summary, only the French treatment phase resulted in a significant overall improvement on naming accuracy across all three languages over time, with naming improvement for target items and decreased naming performance for non-target items.

3.3. *Language selection for lexical retrieval during treatment provided in each language and across all three treatment phases*

We further evaluated P1's language selection for lexical retrieval across all treatment phases. For each treatment phase, we evaluated (i) whether the targeted language in treatment was the most frequently selected language for lexical retrieval (Retrieval language) leading to superior naming accuracy regardless of the language being tested (Probe language) (main effect of Retrieval language) and (ii) whether this pattern of language selection changed over the course of treatment (interaction Time x Retrieval language). We also assessed (iii) the extent to which P1's language selection for lexical retrieval (Retrieval language) led to superior naming accuracy in the target language in which she was probed (Probe language) (e.g.,: being able to accurately retrieve items in the target language under assessment) relative to other non-target languages (e.g.,: being able to retrieve items in the non-target language reflecting cross-language intrusions into the language being probed) (interaction Probe language x Retrieval language) and (iv) whether the pattern of language selection leading to accurate retrieval in relation to the language being assessed changed over the course of treatment (interaction Time x Probe language x Retrieval language). Three GLMMs were separately constructed for treatment provided in French, Spanish, and English (See Supplementary Tables 13–15). These analyses revealed various relevant findings about the cross-language interactions taking place over the course of each treatment phase, described in the following sections.

3.3.1. *French treatment phase*

The French treatment phase model showed a main effect of Retrieval language on naming accuracy (β = −.717, SE = .348,

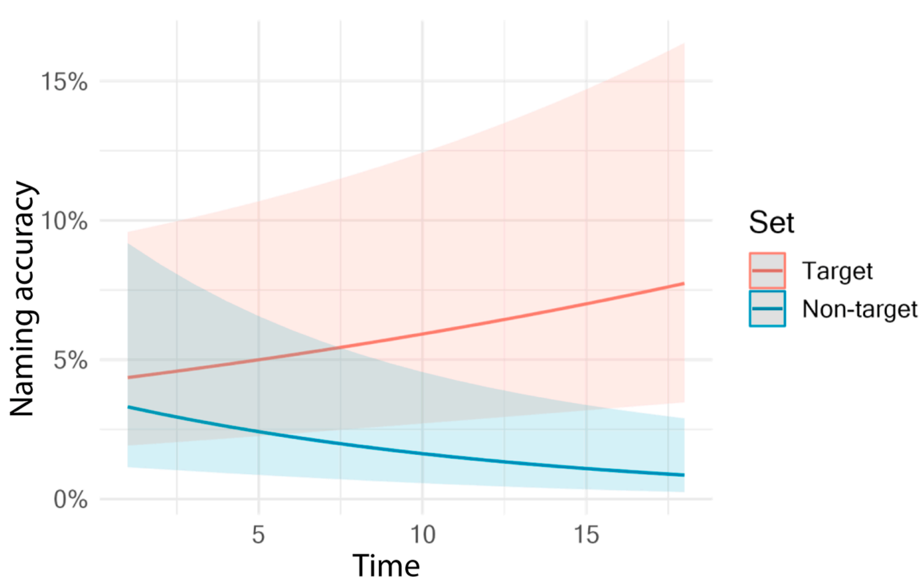

Fig. 5 – Overall effects of the French treatment phase on naming accuracy. The significant Time × Set interaction on naming accuracy (language-dependent scoring method) during the French treatment phase is shown contrasting target items (red line) and non-target items (blue line). 95% confidence intervals are shown as shaded regions around lines. While accuracy improved for target items over time, it decreased for non-target items.

95% CI [−1.400, −.034], z = −2.058, p = .040). Post-hoc marginal means contrasts revealed that Spanish naming responses were significantly more accurate than English naming responses (p = .003). No other contrasts reached significance (all p values ≥ .151) including the contrast between French and Spanish, indicating that neither French nor Spanish was used more than the other for accurate lexical retrieval despite treatment being provided in French. Importantly, the interaction between Time and Spanish as the Retrieval language was significant (β = .075, SE = .030, 95% CI [.017, .133], z = 2.523, p = .012). Post-hoc marginal trends contrasts demonstrated that the slope of Spanish (β = −.029) and English (β = −.002) responses over time were negative, whereas the slope of French responses was positive (β = .096). Moreover, the contrast between French and Spanish slopes was significant (p = .020), suggesting that P1's selection of French for accurate lexical retrieval increased over the course of treatment in French to the detriment of retrieval in Spanish. Overall, these findings indicate that although L1 Spanish was predominantly selected as the Retrieval language during treatment in L3 French, its selection decreased over the course of treatment in French (Fig. 6).

The interaction between French as the Probe language and French as the Retrieval language was significant (β = 1.167, SE = .440, 95% CI [.305, 2.029], z = 2.654, p = .008), suggesting superior naming accuracy when French was both the Probe and the Retrieval language (Fig. 7A). Post-hoc contrasts revealed that, when French was the Probe language, French as the Retrieval language resulted in superior naming accuracy relative to both English and Spanish (both p ≤ .003), whereas the English-Spanish contrast was not significant (p = .821). This indicates that when treated in French and probed in French, selecting the target French for lexical retrieval led to higher naming accuracy in P1 relative to her other non-target languages. Hence, her retrieval ability in French did not experience significant cross-language intrusion from the untreated Spanish or English.

When Spanish was the Probe language, the contrast between Spanish and French as the Retrieval language yielded no significant differences (p = .999). However, the contrasts between retrieval in Spanish and retrieval in French and English resulted in significantly larger estimates for Spanish (both p ≤ .0001), indicating that retrieval in Spanish led to superior naming accuracy. This suggests that when treated in French and probed in Spanish, selecting the target Spanish for lexical retrieval led to higher naming accuracy in P1 relative to her other non-target languages. Therefore, her retrieval ability in Spanish did not undergo significant cross-language intrusion from the treated language French or the untreated language English.

Finally, when English was the Probe language, there were no significant differences for English relative to French or Spanish as Retrieval language (p values ≥ .431), although significantly greater naming accuracy was observed for retrieval in Spanish relative to French (p = .043). This suggests that when treated in French and probed in English, selecting the target English for lexical retrieval did not lead to higher naming accuracy in P1 relative to retrieving items in her other non-target languages. Thus, retrieval in English likely underwent cross-language intrusion from both the treated French and the untreated Spanish.

All Time x Probe language x Retrieval language three-way interactions were non-significant (all p values ≥ .053), suggesting that the patterns of language selection for lexical retrieval reported above did not significantly change over the course of the French treatment phase in relation to the Probe language.

3.3.2. *Spanish treatment phase*

Different from the French treatment phase, we did not find a significant main effect of Retrieval language (p = .119) and the interaction Time by Retrieval language was not significant (p = .305) despite a positive slope of Spanish (β = .021) and negative slopes of French (β = −.181) and English (−.071).

Fig. 7B shows the results of the interaction Probe language x Retrieval language. A significant interaction was found between Spanish Probe language and Spanish Retrieval language (β = 1.370, SE = .450, 95% CI [.488, 2.253], z = 3.043, p = .002), indicating superior naming accuracy when Spanish was both the Probe and the Retrieval language. Post-hoc contrasts revealed similar findings to those identified during the French treatment phase. When the Probe language was Spanish, retrieval in Spanish resulted in higher naming accuracy relative to retrieval in English and French (both p ≤ .0001), while no significant differences were observed for retrieval in English versus French (p = .459). Similarly, when French was the probe language, retrieval in French resulted in superior naming accuracy relative to retrieval in English and Spanish (both p ≤ .0001), while differences for retrieval in Spanish versus English were not significant (p = .350). This suggests that when treated in Spanish and probed in Spanish and French, selecting the target language for lexical retrieval led to superior naming accuracy in Spanish and French respectively relative to retrieving items accurately in her non-target languages. Therefore, P1's retrieval ability in Spanish did not undergo significant cross-language intrusion from her untreated languages. Likewise, her retrieval ability in French did not experience significant cross-language intrusion from the treated Spanish or the untreated English. However, when English was the probe language, there were no significant differences in naming accuracy among the three retrieval languages (all p values ≥ .209). This suggests that when treated in Spanish and probed in English, selecting the target English for lexical retrieval did not lead to superior naming accuracy in P1 relative to retrieving items in her other non-target languages. Thus, retrieval in English likely underwent significant cross-language intrusion from both the treated Spanish and the untreated French.

All Time x Probe language x Retrieval language three-way interactions were non-significant (all p values ≥ .456), suggesting that the reported patterns of language selection for lexical retrieval as they relate to the Probe language did not significantly change over the course of the Spanish treatment phase.

3.3.3. *English treatment phase*

Similar to the Spanish treatment phase, we did not find a significant main effect of Retrieval language (p = .288) and the interaction between Time and Retrieval language was not significant (p = .330) despite a positive slope of English (β = .051) and negative slopes of French (β = −.093) and Spanish (−.006).

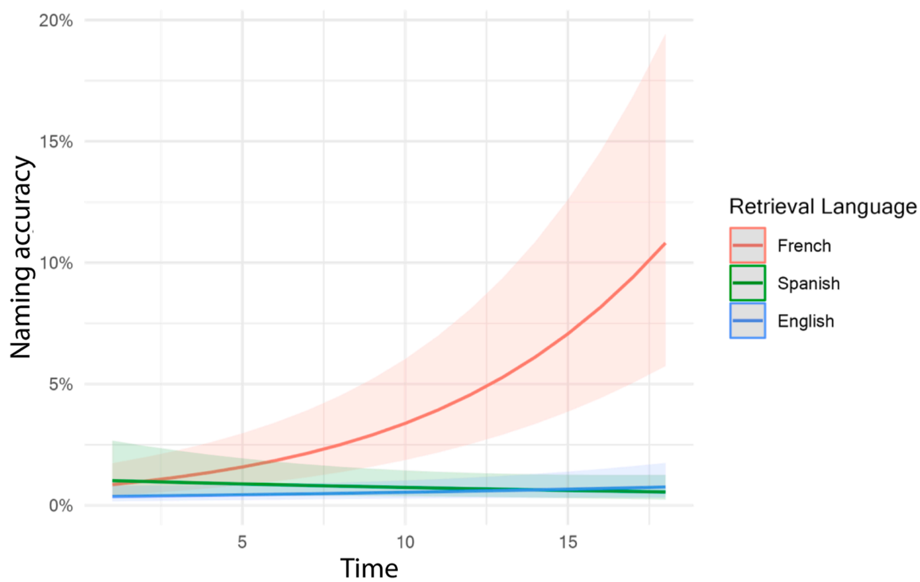

Fig. 6 — Dynamics of language selection for lexical retrieval during the French treatment phase. The Time × Language interaction on naming accuracy (language-independent scoring method) during the French treatment phase is shown, contrasting the languages selected by P1 for lexical retrieval on naming probes: French (red), Spanish (green), and English (blue). 95% confidence intervals are shown as shaded regions around lines. During the French treatment phase, P1 showed superior naming accuracy in Spanish over English (main effect of retrieval language) with no notable differences between lexical retrieval in French versus Spanish. However, she shifted her retrieval language from Spanish to French as treatment in French progressed (Time × Spanish interaction) suggesting an increased selection of the treated language for lexical retrieval.

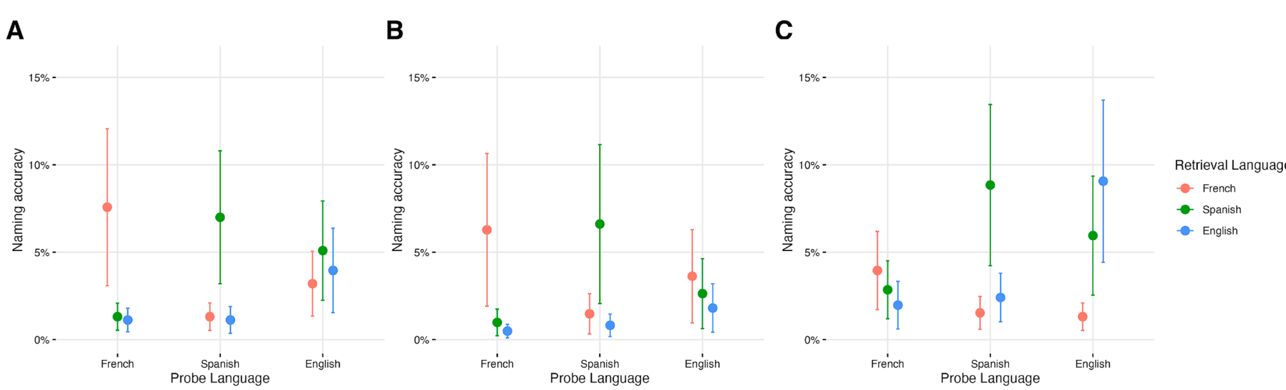

Fig. 7 — Language selection for lexical retrieval during the French, Spanish and English treatment phases. The Probe Language x Retrieval Language interaction on naming accuracy (language-independent scoring method) in each treatment phase is shown. The contrast between languages selected by P1 for lexical retrieval on naming probes is depicted (colored dots with error bars): French (red), Spanish (green), and English (blue). The French treatment phase (A) and the Spanish treatment phase (B) revealed that naming in French on French naming probes and naming in Spanish on Spanish naming probes yielded superior naming accuracy without significant cross-language intrusion from other languages, whereas naming probes in English showed cross-language intrusions from both French and Spanish. The English treatment phase (C) showed that selecting English and Spanish for lexical retrieval during English naming probes resulted in comparable levels of higher naming accuracy relative to naming in French (suggesting cross-language intrusion from the untreated Spanish into the treated English during naming probes). Selecting Spanish for lexical retrieval during Spanish naming probes showed superior accuracy compared to French and English, while no differences across languages were observed for naming probes in French (suggesting cross-language intrusion from both the treated English and the untreated Spanish into French during naming probes).

The statistical model revealed a significant interaction between Probe language and Retrieval language ($\beta = 2.646$, $SE = 1.168$, $z = 2.265$, $p = .024$) (Fig. 7C). Post-hoc contrasts revealed that when English was the Probe language, both English and Spanish as Retrieval languages led to significantly higher naming accuracy compared to French (both $p \leq .0016$), although no significant difference in naming accuracy was observed between English and Spanish ($p = .242$). Conversely, when Spanish was the probe language, retrieval in Spanish resulted in superior naming accuracy relative to both French and English (both $p \leq .0002$), although no significant difference in naming accuracy was observed between English and Spanish as retrieval languages ($p = .242$). These results suggest that when treated in English and probed in English, retrieval in the treated English likely underwent cross-language intrusion from the untreated Spanish but not from the untreated French. However, when Spanish was the probe language, P1's retrieval ability in Spanish led to superior naming accuracy despite receiving some degree of cross-language intrusion from the treated English (but not from the untreated French). Finally, when French was the probe language, post-hoc contrasts showed that the differences between each pair of retrieval languages were not statistically significant (all $p \geq .170$), suggesting that P1's ability to retrieve items in French underwent cross-language intrusion from the treated English and the untreated Spanish.

All Time x Probe language x Retrieval language three-way interactions were non-significant (all p values $\geq .066$), suggesting that the reported patterns of language selection for lexical retrieval as they relate to the Probe language did not significantly change over the course of the English treatment phase.

3.4. *Case-control comparison on executive function performance*

We compared P1's performance on the RCPM (13/36) to the performance of 19 healthy controls (age M = 60.6, SD = 5.3) on this test (M = 24.8, SD = 6.6) reported elsewhere (Kertesz & McCabe, 1975). The analysis indicated the presence of nonverbal executive deficits as P1's score was significantly below the mean performance of the healthy controls [$t = -1.743$, $p = .04$, one-tailed; effect size (Z-CC) for the difference between case and controls = −1.788 (95% CI: −2.511 to −1.044); Bayesian point estimate of the percentage of control population falling below case's score = 4.9256 (95% CI: .6014−14.8190)].

4. Discussion

This study sought to evaluate the effects of a semantic-based anomia treatment provided to P1, a trilingual person with aphasia who presented with post-stroke impairment in her three languages after extensive damage to language cortical regions and the basal ganglia. The same semantic treatment targeting anomia was sequentially provided in her L3 French, L1 Spanish and L2 English. Our study aimed (i) to assess the direct treatment effects and cross-language generalization effects of each language treatment phase on each language, (ii) to evaluate the overall effects of each treatment phase across languages, (iii) to characterize P1's patterns of language selection for lexical retrieval on naming probes taking place during treatment and (iv) to determine the presence of executive function deficits to better understand P1's treatment response. Our main findings indicate that P1 showed significant overall treatment effects restricted to the weakest L3 French, while her strongest L1 Spanish and L2 English did not show significant treatment benefits, and the latter also presented with cross-language interference effects. P1's treatment response can be best understood in the context of severe post-stroke impairment in her three languages, executive deficits and damage to critical regions in both the language processing and the language control network. These findings are discussed in more detail in the following sections.

4.1. *Treatment effects on the treated language*

To address our first aim, we assessed the direct effects of each treatment phase on treated items relative to control items (primary outcome measures) and P1's change in lexical retrieval on standardized language tests (secondary outcome measures) in the treated language. These analyses considered a language-dependent scoring method (i.e., accurate naming responses in the target language).

We found no significant treatment gains in the treated language on any of P1's three languages when contrasting change on treated versus control items after each treatment phase. Only the first treatment phase in French yielded a significant improvement for both treated and control items in French over the course of treatment, suggesting a generalized effect on naming ability in this language. Additionally, we identified small increments in P1's BNT and verbal fluency scores following the French treatment phase suggesting small improvements in other measures of lexical retrieval in this language. In turn, we found decreased performance in these tests in Spanish following treatment in this language (lexical retrieval performance in English could not be assessed since testing was discontinued after the English treatment phase due to acute illness). Altogether, only the weakest L3 French but not the strongest L1 Spanish or L2 English showed significant overall improvements after treatment in the targeted language.

The observed treatment responsiveness in L3 French is consistent with past research showing that non-native languages have potential for treatment-induced recovery in MWA (Faroqi-Shah et al., 2010). Of note, although P1 used French for several years in adulthood, this language only achieved intermediate proficiency and it likely underwent attrition prior to stroke due to discontinued exposure and lack of use since the age of 50 (her stroke occurred at the age of 62). Thus, her overall positive response to treatment in French supports past evidence for positive treatment gains in MWA in languages acquired during adulthood with extensive use despite lower proficiency relative to earlier-acquired languages (Goral et al., 2023) and in attrited languages with incomplete mastery prior to stroke (Lerman et al., 2023).

Importantly, treatment response in MWA depends on multiple factors including age of language acquisition (Goral et al., 2023), pre-stroke language abilities, post-stroke language impairment and lesion characteristics (Goral & Lerman, 2020; Peñaloza & Kiran, 2019). Language proficiency indexes the depth of encoding of linguistic knowledge and deeply encoded languages may be less vulnerable to brain insult (Nadeau, 2019). As recovery tends to be better for the strongest (Kohnert, 2004; Nadeau, 2019) and acquired-from-birth languages (Goral et al., 2023) one may have expected that the L1 Spanish and L2

English showed better responsiveness to therapy given their high pre-stroke proficiency and continued language exposure and use after stroke. However, the absence of significant treatment gains in these languages suggests that other factors including aphasia severity, lesion characteristics and cross-language interaction may have determined P1's response to treatment in these languages.

P1 presented with extensive damage to left fronto-temporo-parietal regions including the basal ganglia resulting in severe deficits across languages (i.e., comparably low WAB-AQ scores). This severe breakdown in the language processing system likely resulted in substantially increased thresholds for activation across languages. Moreover, although semantic-based treatment for anomia is assumed to enhance activation at semantic representational levels to increase spreading activation in the language system and make linguistic representations more accessible (Kiran et al., 2013), P1 also presented with semantic processing deficits (i.e., below the cut-off PAPT scores) which may have limited the positive effects typically observed after semantic-based therapy. Additionally, the order in which languages were targeted in treatment may have impacted treatment response in each language differently. Although the weakest L3 French may have presented with low activation levels due to pre-stroke attrition, and lack of language exposure and use at post stroke combined with severe impairment, French was the first language targeted in treatment. Hence, while it may have received interference from Spanish and English being P1's other stronger languages, interference was not due to heightened levels of activation resulting from earlier intervention in these languages. In contrast, while the L1 Spanish and L2 English may have presented with better (yet insufficient) baseline levels of activation than French due to higher language exposure and use at pre and post stroke, they still presented with severe impairment and may have received interference from previously treated languages. In such case, the native L1 Spanish may have received interference from a weaker L3 French with heightened levels of activation resulting from recent treatment, whereas the L2 English treated at the end, may have received cumulative interference from a weaker L3 French and the strongest L1 Spanish, both with heightened levels of activation resulting from recent treatment. This is consistent with the finding that only the first treatment phase (free from cumulative interference from previously treated languages) was generally effective across languages (see section 4.3).

4.2. *Cross-language treatment effects on the untreated languages*

We further addressed our first aim assessing whether cross-language generalization effects occurred in each treatment phase by contrasting naming accuracy for untreated translations relative to control translations in the untreated languages (primary outcome measures) and examining P1's change in lexical retrieval on standardized language tests on those languages (secondary outcome measures). Again, these analyses considered a language-dependent scoring method (i.e., accurate naming responses in the target language).

The contrast between untreated and control translations in each language treatment phase revealed no significant cross-language generalization effects in the untreated language across all three language treatment phases. Notably however, the improvement in the L2 French during the French treatment phase was accompanied by cross-language interference effects in the L2 English with a significant decline in naming accuracy in the English control translations. These findings are consistent with P1's performance on secondary outcome measures in (i) treatment in French, with no change or small decrease on her BNT and verbal fluency scores in English, and unstable performance on these tests in Spanish and (ii) treatment in Spanish, with no change or decrease on her BNT and verbal fluency scores in both French and English. Our findings align with previous evidence for treatment gains being most often prominent in a treated language relative to non-treated languages (Faroqi-Shah et al., 2010; Goral et al., 2023; Lee & Faroqi-Shah, 2024), evidence for the absence of cross-language generalization in severe aphasia (Keane & Kiran, 2015; Kurland & Falcon, 2011; Meinzer et al., 2007) and evidence of cross-language interference effects with decreased performance in the untreated language (Abutalebi et al., 2009; Goral et al., 2013).

Previous reviews have shown mixed results regarding the occurrence and direction of cross-language generalization in MWA (Ansaldo & Saidi, 2014; Faroqi-Shah et al., 2010; Kohnert, 2009). Cross-language generalization is more likely to occur in balanced bilinguals regardless of the treated language and may occur more frequently from the least to the most proficient language in unbalanced bilinguals (Ansaldo & Saidi, 2014). However, when cross-language generalization does not occur and interference effects are observed, they may result from faulty inhibitory mechanisms whereby increased activation of the treated language increases the suppression of the untreated language to minimize its potential interference (Goral & Lerman, 2020). In some cases, this operation may lead to the lingering suppression of the untreated language even after treatment (Goral et al., 2013). As found in previous research, the absence of cross-language generalization and presence of interference effects in P1 may result from her basal ganglia damage (Abutalebi et al., 2009; Keane & Kiran, 2015).

4.3. *Effects of each treatment phase on all three languages*

Our second aim was complementary to the analyses discussed above in assessing the overall effects of each treatment phase on all three languages comprehensively across all target items (treated items and untreated translations) relative to non-target items (control items and control translations). As in previous analyses, our approach considered a language-dependent scoring method (i.e., accurate naming responses in the target language).

This examination revealed that any significant improvements in naming accuracy observed across all three languages were restricted to the French treatment phase. These findings suggest that overall, only the first treatment phase was generally effective while the effects of the subsequent treatment phases in Spanish and English were attenuated. As P1 showed extensive neural damage and severe baseline impairment across languages, limited responsiveness to therapy was expected. However, as discussed previously, it is likely that during the French treatment phase, languages were free from cumulative interference from previously treated languages yielding this first treatment phase more effective relative to the following ones (see section 4.1).

4.4. *Language selection for accurate lexical retrieval during treatment*

Our third aim was to characterize P1's patterns of language selection during naming probes in each treatment phase (i.e., if accurate naming responses were provided in the probed language or in non-target languages reflecting cross-language intrusions). These analyses considered the language-independent scoring method (i.e., accurate naming responses regardless of the target language). Our analyses revealed three main findings. First, the French treatment phase revealed that overall and regardless of the probe language, the strongest untreated L1 Spanish competed with the weakest treated L3 French for language selection leading to accurate naming performance. However, as treatment progressed, using the L3 French increasingly led to accurate naming responses as P1's reliance on her L1 Spanish decreased. In turn, retrieval in the treated language during the Spanish and English treatment phases did not result in significantly more accurate responses relative to retrieval in the untreated languages. Hence, only the L3 French became more available for lexical retrieval as a result of treatment, consistent with the overall significant gains restricted to this language (section 4.3).

Secondly, when examining P1's language selection leading to accurate naming in the probe language, we obtained similar results for the French and Spanish treatment phases. During the French treatment phase, the treated L3 French and the untreated L1 Spanish predominantly led to accurate retrieval in French and Spanish naming probes. During the Spanish treatment phase, the results were similar for the treated L1 Spanish and the untreated L3 French, suggesting that these two languages did not undergo significant cross-language intrusions during these treatment phases.

In contrast, using the L2 English did not result in more accurate naming responses on English probes across the two initial therapy phases, suggesting that English was overall less available for lexical retrieval and underwent cross-language intrusions from French and Spanish. This is consistent with the observed cross-language interference effects from the treated French on the untreated English previously discussed (section 4.2). By the final English treatment phase, using English still did not lead to more accurate naming responses on English probes relative to other untreated languages suggesting continued cross-language intrusions (more so from Spanish than French). Notably, as in previous treatment phases, only the L1 Spanish continued to lead to more accurate naming responses on Spanish probes relative to retrieval in other languages. Conversely and different from previous treatment phases, using French during French naming probes no longer led to more accurate naming responses relative to using other languages, suggesting reduced availability of French for lexical retrieval towards the end of the intervention. Finally, these patterns of language selection for lexical retrieval did not significantly change over the course of each treatment phase.

In summary, only the strongest L1 Spanish remained consistently available for language selection leading to accurate retrieval in the probed language across treatment phases regardless of the treated language and suffered less cross-language intrusions. The L2 English remained largely unavailable for language selection across all language treatment phases, even when it was treated several months after showing cross-language interference from treatment in French. This was the language undergoing more cross-language intrusions. In turn, while French showed an overall positive response in the first treatment phase and remained available for language selection during the second phase, treatment effects may have declined towards the third therapy phase making it less available for accurate retrieval and more prone to cross-language intrusions.

Notably, involuntary pathological language switching and cross-language intrusions have been reported in MWA with basal ganglia lesions reflecting a difficulty to inhibit non-target languages for appropriate lexical retrieval in the target language (Abutalebi et al., 2009; Ansaldo et al., 2010; Keane & Kiran, 2015; Kurland & Falcon, 2011). Damage to language control regions may prevent the correct selection among highly competing lexical units across languages (Ansaldo & Saidi, 2014) or may result in one language being more accessible than others for use in lexical retrieval regardless of the treated language (Kurland & Falcon, 2011). Additionally, cross-language intrusions can occur from the strongest L1 into the weakest L2 in the absence of treatment and from the weakest yet treated L2 into the untreated strongest L1 (Abutalebi et al., 2009) as well as from the treated into the untreated language in sequential treatments targeting the L2 and L3 (Keane & Kiran, 2015). P1's naming performance further suggests that a treated language can receive intrusions from the untreated languages and that cross-language influences likely occur both within and across treatment phases in sequential language interventions. In consideration of this previous evidence, the presence of basal ganglia damage and concomitant executive function deficits, P1's cross-language intrusions may reflect a language control deficit. In this view, ineffective language inhibitory mechanisms may result in failure to activate the correct lexical representations according to the linguistic context demands, thus negatively impacting treatment effectiveness as measured by language-dependent scoring methods.

Alternatively, P1's cross-language intrusions may reflect a compensatory response that emerged during treatment to meet a communicative purpose when the target representation in the intended language was not fully available. Although her language switching patterns prior and after her stroke were not formally evaluated, this behavior was not clearly evident prior to treatment. Also, all items selected for P1's treatment were items she was initially unable to name accurately in her three languages. Despite the lack of significant improvements, it is possible that some item representations may have become more (yet inconsistently) available across languages as a result of treatment targeting the same items across the three treatment phases. For instance, the treated word form “coat” may not have been available for retrieval in L2 English when probed in this language, but its availability across languages as “abrigo” (L1 Spanish) and “manteau” (L3 French) may have increased since this item's conceptual representation was also targeted across treatments in her L3 French and L1 Spanish. Cross-language intrusions led to successful naming reflecting increased availability of targeted representations albeit in the non-target language as captured by language-independent scoring methods.

The possibility that cross-language intrusions emerge as a compensatory strategy is in line with recent research with MWA reporting conscious intentional switches between languages to substitute or aid the retrieval of a target word in the intended language (Mooijman et al., 2025). Cross-linguistic intrusions resulting in accurate lexical retrieval also align with observations from neurotypical multilinguals in dense-code switching contexts. As proposed by the Adaptive Control Hypothesis languages can operate in a cooperative opportunistic manner to achieve successful lexical retrieval, allowing speakers to bypass the inhibitory control demands of contexts that require retrieval in a single language. In this view, our results are consistent with previous evidence showing that MWA can benefit from their lexical knowledge in an available language during tasks that do not constrain word retrieval to a single language, showing superior performance relative to single language retrieval tasks (Carpenter et al., 2020). Although it is difficult to determine the exact nature of the cross-language intrusions observed in P1, they may result from the dysfunctional interaction between impaired language control and language processing networks and the increased yet inconsistent availability of lexical representations as a result of sequential treatment targeting the same items across languages.

4.5. *Executive function deficits and basal ganglia damage*

As regards to our final aim, we found that P1 presented with executive function deficits as measured by the RCPM. The frontal lobes are essential for a variety of executive operations and the basal ganglia have a modulatory role in these cognitive functions via fronto-striatal connections (Middleton & Strick, 2000). Basal ganglia lesions can disrupt these connections resulting in persistent deficits of lexical retrieval that depend on executive processes (Peñaloza et al., 2014). As languages compete for selection in multilinguals, functional executive control mechanisms are required to manage language co-activation during lexical retrieval, helping the selection of the target language by suppressing interference from non-target languages (Green, 1998). Language control in multilinguals is supported by a cortico-subcortical network comprising several brain regions including (i) the anterior cingulate cortex involved in conflict and error monitoring during language selection and switching; (ii) the prefrontal cortex supporting conflict resolution entailing the selection of the target language and the inhibition of the non-target language; (iii) the inferior parietal lobule directing attentional orienting for language selection; and crucially (iv) the basal ganglia supporting the control of the two languages while keeping track of the target language (Abutalebi & Green, 2007; Calabria et al., 2018). In MWA, basal ganglia damage can impair language control mechanisms and account for differential and selective patterns of impairment and recovery across languages (Aglioti & Fabbro, 1993; Verreyt et al., 2013), pathological mixing and switching (Abutalebi et al., 2000; Ansaldo et al., 2010) and cross-language interference (Keane & Kiran, 2015).

Importantly, nonverbal inhibitory control performance predicts response to semantic treatment in MWA (Bihovsky et al., 2023), and the potential for cross-language generalization effects depends on the integrity of the cognitive control network (Ansaldo & Saidi, 2014). Indeed, neuroimaging studies have shown that language recovery in MWA is associated with increased functional connectivity between language processing and executive control regions during naming (Abutalebi et al., 2009; Radman et al., 2016). Hence, our findings align with past research (Keane & Kiran, 2015) and suggest that the absence of cross-language generalization and the occurrence of interference effects during language therapy in P1 may result from basal ganglia damage disrupting the functionality of the language control network and its interaction with the language processing network.

4.6. *An integrative view of anomia treatment effects in severe multilingual aphasia*

Semantic-based treatments are effective to improve lexical access deficits in MWA (Ansaldo & Saidi, 2014). They are assumed to boost activation of the conceptual representations of words at the semantic level, facilitating increased spreading activation at post-semantic levels (Quique et al., 2019) for the treated and the untreated languages (Kiran et al., 2013). MWA with neural damage restricted to the language processing network are likely to show lexical access improvement following semantic anomia therapy, with benefits being generally expected in the treated language (Faroqi-Shah et al., 2010; Goral et al., 2023; Lee & Faroqi-Shah, 2024). Additionally, cross-language generalization may be expected if they speak languages which are relatively close in structural distance (Miertsch et al., 2009) and show similar pre-stroke proficiency across languages (Edmonds & Kiran, 2006; Kiran & Roberts, 2010; Marangolo et al., 2009; Peñaloza et al., 2021). For these individuals, treatment gains may be mainly constrained by the amount of neural damage in their language processing system reflected in their aphasia severity across languages. Severe aphasia can result from extensive neural damage to regions in the language processing network and it can manifest with both lexical access and semantic processing deficits, and increased activation thresholds that limit accessibility to word representations across languages. Hence, damage to the lexical-semantic system as in P1, may result in reduced stimulation of the conceptual representations of the treated words and limited spreading activation to their lexical representations in the treated and the untreated languages, resulting in limited treatment effects across languages. This is consistent with computational modeling evidence showing that substantial semantic and phonetic damage to a bilingual language neural network (Peñaloza et al., 2019) to simulate severe semantic and naming deficits across languages in bilingual aphasia also prevents significant language recovery across languages during neural network retraining (Grasemann et al., 2021).

Importantly, when brain damage also extends to key brain regions within the language control network such as the basal ganglia, severe breakdown in the bilingual language processing system may interact with a dysfunctional language control system responsible for inhibiting non-target languages to facilitate the selection of the target language for accurate lexical retrieval according to linguistic context demands. Hence, the dysfunctional interaction between the impaired language processing and language control systems may break the balance between increased spreading activation promoted via treatment and increased suppression to control interference from non-target languages necessary for effective treatment leading to cross-language generalization (Goral & Lerman, 2020; Kiran et al., 2013). This broken balance between activation and suppression mechanisms may result in cross-language interference effects which can occur within treatment and across treatments targeting different languages in the same individual. Cross-language interference may affect one language more than others yielding it less accessible for lexical retrieval (e.g., English in P1), whereas languages less affected by interference from stronger languages previously activated via therapy may be more responsive when targeted in treatment (e.g., French in P1).

P1's evidence suggests that cross-language intrusions (i.e., lexical retrieval of a target word in the non-target language) can occur during multilingual anomia treatment regardless of the treated and probed language and in varying directions between weaker and stronger languages. Although the causal nature of cross-language intrusions remains unclear, two currently debated accounts include language control deficits and lexical compensation (see Mooijman et al., 2025 for an overview). The first account is supported by P1's evidence of cross-language interference effects during treatment, presence of neural damage to the language control system and concurrent executive function deficits. The second account is supported by evidence of P1's intrusions resulting in the correct retrieval of probed items in non-target languages, particularly when the probed language showed less availability (i.e., English). Although evidence from P1 cannot differentiate between these accounts, her findings suggest they might not be mutually exclusive. Findings from Mooijman et al. (2025) suggest that lexical compensation is (i) beneficial for MWA with mild or no language control deficits for whom using the non-target language can be a voluntary strategy, (ii) less likely in MWA with severe language control deficits for whom using the non-target language is rather involuntary and detrimental. Hence, lexical compensation may emerge to varying extents along a continuum of language control impairment severity with P1's profile reflecting an intermediate point along this continuum. Her basal ganglia injury may have led to language control deficits resulting in cross-language interference effects in English with less availability of this language for lexical retrieval. However, a compensation strategy may have emerged (likely from treatment) to occasionally allow for the opportunistic yet inconsistent access to lexical representations in her more available French and Spanish during naming probes in English.

4.7. *Limitations and future research*

A few study limitations are worth noting. Unexpected acute illness in P1 occurred twice during this study, interrupting treatment in Spanish for 2 months and prior to post English treatment assessments. The first interruption may have reduced treatment responsiveness in Spanish while the second did not allow to fully evaluate the effects of the English treatment phase. However, given the presence of interference effects in English and its lower availability for lexical retrieval, it is unlikely that results would have differed substantially. Also, our time-demanding study design required us to minimize the use of additional tests that may have helped to characterize P1's language impairment in a more fine-grained manner and to evaluate treatment effects more extensively. The use of a language switching questionnaire would have allowed us to assess the presence and frequency of code-switching patterns in P1 prior and after her stroke and get a better understanding of her cross-language intrusions. As regards to treatment, anecdotal feedback from P1's family suggests she showed increased spontaneous communication and more fluent speech over the course of therapy, although these changes may have not been captured by our measures. Discourse, other language production and functional communication measures should be included in future rehabilitation research with MWA.

Additionally, we did not examine her performance on inhibitory control tasks as the extensive assessment of her three languages was already highly demanding for P1. Although the RCPM may have been sensitive to detect executive deficits, inhibitory control tasks would have enabled us to assess the functionality of her language control network in a more reliable manner. Future studies should employ these measures as they tap more directly on the control mechanisms used in lexical retrieval by multilinguals. Using the language-dependent and language-independent scoring approaches employed here may be beneficial for future research, helping to separately measure cross-language intrusions from fully correct naming responses in the target language and identify cross-language interaction effects in lexical retrieval. Finally, as our treatment did not directly target language selection deficits, these were not expected to change during therapy. Other therapy approaches designed to improve executive control processes involved in language selection or decrease control demands may help to improve treatment response and promote cross-language generalization in MWA (Ansaldo et al., 2010).

4.8. *Conclusions*

This study assessed the effects of semantic treatment sequentially provided in three languages on the lexical retrieval abilities of P1, a trilingual person with severe aphasia. Our findings suggest that P1's severe language impairment due to extensive damage to the language processing network may have limited treatment effectiveness across all her treated languages. The overall treatment gains restricted to the weaker L3 French may have benefited from the absence of cross-linguistic interference from previously treated languages. Additionally, damage to the basal ganglia, a key structure in the language control network, may have resulted in difficulty to properly maintain the balance between activation and inhibition mechanisms for language selection during treatment (Goral & Lerman, 2020; Kiran et al., 2013). Dysfunctional control mechanisms in P1 may explain the absence of cross-language generalization, and the cross-language interference from the treated L3 French to the untreated L2 English, while the stronger L1 showed better resistance to cross-linguistic influences. We propose that severe impairment and extensive brain damage affecting the interplay between the language processing and the language control networks may hinder significant treatment-induced recovery across languages in MWA. Additionally, cross-language intrusions reflecting the use of the non-target language for lexical retrieval may follow sequential multilingual anomia treatment although their underlying causal mechanism deserves more research. The comprehensive examination of cross-language interaction during treatment in MWA may help to better understand their language impairment and develop more effective interventions to maximize treatment outcomes after acquired brain injury.

CRediT authorship contribution statement

Claudia Peñaloza: Writing – review & editing, Writing – original draft, Visualization, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Manuel Jose Marte:** Writing – review & editing, Writing – original draft, Visualization, Investigation, Formal analysis, Data curation. **Anne Billot:** Writing – review & editing, Methodology, Investigation, Data curation, Conceptualization. **Swathi Kiran:** Writing – review & editing, Writing – original draft, Methodology, Funding acquisition, Conceptualization.

Scientific transparency statement

DATA: Some raw and processed data supporting this research are publicly available, while some are subject to restrictions: https://osf.io/d6ys2/

CODE: All analysis code supporting this research is publicly available: https://osf.io/d6ys2/

MATERIALS: Some study materials supporting this research are publicly available, while some are subject to restrictions: https://osf.io/d6ys2/

DESIGN: This article reports, for all studies, how the author(s) determined all sample sizes, all data exclusions, all data inclusion and exclusion criteria, and whether inclusion and exclusion criteria were established prior to data analysis.

PRE-REGISTRATION: No part of the study procedures was pre-registered in a time-stamped, institutional registry prior to the research being conducted. No part of the analysis plans was pre-registered in a time-stamped, institutional registry prior to the research being conducted.

For full details, see the *Scientific Transparency Report* in the supplementary data to the online version of this article.

Funding

This work was supported by the National Institutes of Health/National Institute on Deafness and Other Communication Disorders grant U01DC014922 awarded to Swathi Kiran. Claudia Peñaloza is supported by grant RYC2021-034561-I funded by MCIN/AEI/10.13039/501100011033 and the European Union NextGenerationEU/PRTR. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Declaration of competing interest

Swathi Kiran owns ownership stock and is an advisor for Constant Therapy Health with no scientific overlap with the present work. There are no other conflicts to disclose.

Acknowledgments

The authors thank P1 and her family for their committed participation in this study. We also thank Dr. Andreas Charidimou for his expert opinion on our study participant's brain MRI data.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cortex.2025.05.017.

Item	Treated items and Untreated translations[1]			Control items and Control translations[2]		
	Spanish (L1)	English (L2)	French (L3)	Spanish (L1)	English (L2)	French (L3)
1	Bolsa	Bag	Sac	Medias	Pantyhose	Bas
2	Camarón	Shrimp	Crevette	Aguja	Needle	Aiguille
3	Conejo	Rabbit	Lapin	Mosca	Fly	Mouche
4	Pato	Duck	Canard	Zorro	Fox	Renard
5	Gallo	Rooster	Coq	Cohete	Rocket	Fusée
6	Queso	Cheese	Fromage	Tobillo	Ankle	Cheville
7	Caja	Box	Boîte	Espada	Sword	Épée
8	Bufanda	Scarf	Écharpe	Horno	Oven	Four
9	Cuchillo	Knife	Couteau	Tronco	Log	Bûche
10	Piña	Pineapple	Ananas	Tiburón	Shark	Requin
11	Abrigo	Coat	Manteau	Tetera	Teapot	Théière
12	Ratón	Mouse	Souris	Pelo	Hair	Cheveux
13	Lentes	Glasses	Lunettes	Esposas	Handcuffs	Menottes
14	Cubeta	Bucket	Seau	Tina	Bathtub	Baignoire
15	Cuchara	Spoon	Cuillère	Delantal	Apron	Tablier

[1] Target items (expected to change over the course of treatment), [2] Non-target items (expected to remain relatively stable over the course of treatment). Control items 1–10 are monitored items included in every treatment naming probe. Control items 11–15 are unmonitored items included only in baseline naming probes and post-treatment naming probes.

Appendix

Items employed across the three language treatment phases

REFERENCES

Abutalebi, J., & Green, D. (2007). Bilingual language production: The neurocognition of language representation and control. *Journal of Neurolinguistics, 20*(3), 242–275. https://doi.org/10.1016/j.jneuroling.2006.10.003

Abutalebi, J., Miozzo, A., & Cappa, S. F. (2000). Do subcortical structures control 'language selection' in polyglots? Evidence from pathological language mixing. *Neurocase, 6*(1), 51–56. https://doi.org/10.1080/13554790008402757

Abutalebi, J., Rosa, P. A. D., Tettamanti, M., Green, D. W., & Cappa, S. F. (2009). Bilingual aphasia and language control: A follow-up fMRI and intrinsic connectivity study. *Bilingualism: Acquistion of Linguistic and Non-Linguistic Skills in Bilinguals, 109*(2), 141–156. https://doi.org/10.1016/j.bandl.2009.03.003

Aglioti, S., & Fabbro, F. (1993). Paradoxical selective recovery in a bilingual aphasic following subcortical lesions. *Neuroreport, 4*(12).

Ansaldo, A. I., & Saidi, L. G. (2014). Aphasia therapy in the age of globalization: Cross-linguistic therapy effects in bilingual aphasia. *Behavioural Neurology, 2014.* , Article 603085. https://doi.org/10.1155/2014/603085

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

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software,* 67(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Benton, A. L., & Hamsher, K. (1976). *Multilingual aphasia examination* (2nd ed.). AJA Associates.

Bihovsky, A., Ben-Shachar, M., & Meir, N. (n.d.). Semantic Feature Analysis (SFA) treatment in L1 and L2 in bilingual aphasia: Effects of cognitive and language factors. Aphasiology, 1–29. https://doi.org/10.1080/02687038.2023.2226359.

Calabria, M., Costa, A., Green, D. W., & Abutalebi, J. (2018). Neural basis of bilingual language control. *Annals of the New York Academy of Sciences, 1426*(1), 221–235. https://doi.org/10.1111/nyas.13879

Carpenter, E., Rao, L., Peñaloza, C., & Kiran, S. (2020). Verbal fluency as a measure of lexical access and cognitive control in bilingual persons with aphasia. *Aphasiology*, 1–22. https://doi.org/10.1080/02687038.2020.1759774

Colomé, À. (2001). Lexical activation in bilinguals' speech production: Language-specific or language-independent? *Journal of Memory and Language, 45*(4), 721–736. https://doi.org/10.1006/jmla.2001.2793

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

Crawford, J. R., & Garthwaite, P. H. (2007). Comparison of a single case to a control or normative sample in neuropsychology: Development of a Bayesian approach. *Cognitive Neuropsychology, 24*(4). https://doi.org/10.1080/02643290701290146. Article 4.

Edmonds, L. A., & Kiran, S. (2006). Effect of semantic naming treatment on crosslinguistic generalization in bilingual aphasia. *Journal of Speech, Language, and Hearing Research: JSLHR, 49*(4), 729–748. https://doi.org/10.1044/1092-4388(2006/053

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

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

Gilmore, N., Meier, E. L., Johnson, J. P., & Kiran, S. (2019). Nonlinguistic cognitive factors predict treatment-induced recovery in Chronic Poststroke Aphasia. *Archives of Physical Medicine and Rehabilitation, 100*(7), 1251–1258. https://doi.org/10.1016/j.apmr.2018.12.024

Goral, M., & Lerman, A. (2020). Variables and mechanisms affecting response to language treatment in multilingual people with aphasia. *Behavioral Sciences, 10*(9). https://doi.org/10.3390/bs10090144

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

Goral, M., Norvik, M. I., Antfolk, J., Agrotou, I., & Lehtonen, M. (2023). Cross-language generalization of language treatment in multilingual people with post-stroke aphasia: A meta-analysis. *Brain and Language, 246*, Article 105326. https://doi.org/10.1016/j.bandl.2023.105326

Grasemann, U., Peñaloza, C., Dekhtyar, M., Miikkulainen, R., & Kiran, S. (2021). Predicting language treatment response in bilingual aphasia using neural network-based patient models. *Scientific Reports, 11*(1), Article 10497. https://doi.org/10.1038/s41598-021-89443-6

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. Cambridge Core.

Green, D. W., & Abutalebi, J. (2013). Language control in bilinguals: The adaptive control hypothesis. *Journal of Cognitive Psychology, 25*(5), 515–530. https://doi.org/10.1080/20445911.2013.796377

Grosjean, F. (2021). *Life as a bilingual: Knowing and using two or more languages*. Cambridge University Press.

Hartig, F., & Lukas, L. (2022). *Package 'DHARMa'*.

Howard, D., & Patterson, K. (1992). *Pyramids and palm trees test*. San Antonio, TX: Pearson.

Kaplan, E., Goodglass, H., & Weintraub, S. (2001). *Boston naming test* (2nd ed.). Philadelphia, PA: Lea & Febiger.

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: Language and Cognition, 22*(2), 300–330. https://doi.org/10.1017/S1366728918000366. Cambridge Core.

Kay, J., Coltheart, M., & Lesser, R. (1992). *Psycholinguistic assessments of language processing in aphasia*. Hove, England: Psychology Press.

Keane, C., & Kiran, S. (2015). The nature of facilitation and interference in the multilingual language system: Insights from treatment in a case of trilingual aphasia. *Cognitive Neuropsychology, 32*(3–4), 169–194. https://doi.org/10.1080/02643294.2015.1061982

Kertesz, A. (2006). *Western Aphasia battery - Revised*. San Antonio, TX: Psychological Corporation.

Kertesz, A., & McCabe, P. (1975). Intelligence and aphasia: Performance of aphasics on Raven's coloured progressive matrices (RCPM). *Brain and Language, 2*, 387–395. https://doi.org/10.1016/S0093-934X(75)80079-4

Kertesz, A., Pascual-Leone Garcia, A., & Pascual-Leone, G. (1990). *Batería de afasias "Western" [Western Aphasia Battery en versión en versión y adaptación castellana]*. Valencia, Spain: Nau Llibres.

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

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

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

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

Kohnert, K. J., Hernandez, A. E., & Bates, E. (1998). Bilingual performance on the Boston naming test: Preliminary norms in Spanish and English. *Brain and Language, 65*(3), 422–440. https://doi.org/10.1006/brln.1998.2001

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

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

Lee, S., & Faroqi-Shah, Y. (2024). A meta-analysis of anomia treatment in bilingual aphasia: Within- and cross-language generalization and predictors of the treatment outcomes. *Journal of Speech, Language, and Hearing Research: JSLHR, 67*(5), 1558–1600. https://doi.org/10.1044/2024_JSLHR-23-00026

Lenth, R. V., Bolker, B., Buerkner, P., Giné-Vásquez, I., Herve, M., Jung, M., Love, J., Miguez, F., Riebl, H., & Singmann, H. (2023). *Package 'emmeans'*.

Lerman, A., Goral, M., & Obler, L. K. (2023). Rehabilitating an attrited language in a bilingual person with aphasia. *Clinical Linguistics & Phonetics, 37*(3), 272–290. https://doi.org/10.1080/02699206.2022.2052966

Marangolo, P., Rizzi, C., Peran, P., Piras, F., & Sabatini, U. (2009). Parallel recovery in a bilingual aphasic: A neurolinguistic and fMRI study. *Neuropsychology, 23*(3), 405–409. https://doi.org/10.1037/a0014824

Meinzer, M., Obleser, J., Flaisch, T., Eulitz, C., & Rockstroh, B. (2007). Recovery from aphasia as a function of language therapy in an early bilingual patient demonstrated by fMRI. *Neuropsychologia, 45*(6), 1247–1256. https://doi.org/10.1016/j.neuropsychologia.2006.10.003

Meteyard, L., & Davies, R. A. I. (2020). Best practice guidance for linear mixed-effects models in psychological science. *Journal of Memory and Language, 112*, Article 104092. https://doi.org/10.1016/j.jml.2020.104092

Middleton, F. A., & Strick, P. L. (2000). Basal ganglia and cerebellar loops: Motor and cognitive circuits. *Brain Research Reviews, 31*(2), 236–250. https://doi.org/10.1016/S0165-0173(99)00040-5

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

Miller Amberber, A. (2012). Language intervention in French–English bilingual aphasia: Evidence of limited therapy transfer. *Clinical and Research Applications of the Bilingual Aphasia Test, 25*(6), 588–614. https://doi.org/10.1016/j.jneuroling.2011.10.002

Mooijman, S., Schoonen, R., Goral, M., Roelofs, A., & Ruiter, M. B. (2025). Why do bilingual speakers with aphasia alternate between languages? A study into their experiences and mixing patterns. *Aphasiology*, 1–28. https://doi.org/10.1080/02687038.2025.2452928

Nadeau, S. E. (2019). Bilingual aphasia: Explanations in population encoding. *Journal of Neurolinguistics, 49*, 117–143. https://doi.org/10.1016/j.jneuroling.2018.10.002

Paradis, J., & Libben, G. (1987). *The assessment of bilingual aphasia: The Bilingual Aphasia test*. Hove, UK: Psychology Press.

Peña-Casanova, J., Quiñones-Úbeda, S., Gramunt-Fombuena, N., Quintana-Aparicio, M., Aguilar, M., Badenes, D., Cerulla, N., Molinuevo, J. L., Ruiz, E., Robles, A., Barquero, M. S., Antúnez, C., Martínez-Parra, C., Frank-García, A., Fernández, M., Alfonso, V., Sol, J. M., & Blesa, R., for the NEURONORMA Study Team. (2009). Spanish multicenter normative studies (NEURONORMA Project): Norms for verbal fluency tests. *Archives of Clinical Neuropsychology, 24*(4), 395–411. https://doi.org/10.1093/arclin/acp042

Peñaloza, C., Barrett, K., & Kiran, S. (2019). The influence of prestroke proficiency on poststroke lexical-semantic performance in bilingual aphasia. *Aphasiology*, 1–18. https://doi.org/10.1080/02687038.2019.1666082

Peñaloza, C., Dekhtyar, M., Scimeca, M., Carpenter, E., Mukadam, N., & Kiran, S. (2020). Predicting treatment outcomes for bilinguals with aphasia using computational modeling: Study protocol for the PROCoM randomised controlled trial. *BMJ Open, 10*(11), Article e040495. https://doi.org/10.1136/bmjopen-2020-040495

Peñaloza, C., Grasemann, U., Dekhtyar, M., Miikkulainen, R., & Kiran, S. (2019). BiLex: A computational approach to the effects of age of acquisition and language exposure on bilingual lexical access. *Brain and Language, 195*, Article 104643. https://doi.org/10.1016/j.bandl.2019.104643

Peñaloza, C., & Kiran, S. (2017). Neuroimaging evidence in the treatment of bilingual/multilingual adults with aphasia. *Perspectives of the ASHA Special Interest Groups, 2*(2), 126–131. https://doi.org/10.1044/persp2.SIG2.126

Peñaloza, C., & Kiran, S. (2019). Recovery and rehabilitation patterns in bilingual and multilingual aphasia. In *The handbook of the neuroscience of multilingualism* (pp. 553–571). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781119387725.ch27.

Peñaloza, C., Rodríguez-Fornells, A., Rubio, F., De Miquel, M. A., & Juncadella, M. (2014). Language recovery and evidence of residual deficits after nonthalamic subcortical stroke: A 1 year follow-up study. *Journal of Neurolinguistics, 32*, 16–30. https://doi.org/10.1016/j.jneuroling.2014.08.001

Peñaloza, C., Scimeca, M., Gaona, A., Carpenter, E., Mukadam, N., Gray, T., Shamapant, S., & Kiran, S. (2021). Telerehabilitation for word retrieval deficits in bilinguals with aphasia: Effectiveness and reliability as compared to in-person language therapy. *Frontiers in Neurology, 12*. https://www.frontiersin.org/articles/10.3389/fneur.2021.589330.

Quique, Y. M., Evans, W. S., & Dickey, M. W. (2019). Acquisition and generalization responses in aphasia naming treatment: A meta-analysis of semantic feature analysis outcomes. *American Journal of Speech-Language Pathology, 28*(1S), 230–246. https://doi.org/10.1044/2018_AJSLP-17-0155

R Core Team. (2021). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. URL https://www.R-project.org/.

Radman, N., Mouthon, M., Di Pietro, M., Gaytanidis, C., Leemann, B., Abutalebi, J., & Annoni, J.-M. (2016). The role of the cognitive control system in recovery from bilingual aphasia: A multiple single-case fMRI study. *Neural Plasticity, 2016*. , Article 8797086. https://doi.org/10.1155/2016/8797086

Roberts, P. M., & Doucet, N. (2011). Performance of French-speaking Quebec adults on the Boston Naming Test. *Canadian Journal of Speech-Language Pathology and Audiology, 35*(3), 254–267.

Scimeca, M., Peñaloza, C., & Kiran, S. (2023). Multilevel factors predict treatment response following semantic feature-based intervention in bilingual aphasia. *Bilingualism: Language and Cognition*, 1–17. https://doi.org/10.1017/S1366728923000391. Cambridge Core.

St-Hilaire, A., Hudon, C., Vallet, G. T., Bherer, L., Lussier, M., Gagnon, J.-F., Simard, M., Gosselin, N., Escudier, F., Rouleau, I., & Macoir, J. (2016). Normative data for phonemic and semantic verbal fluency test in the adult French–Quebec population and validation study in Alzheimer's disease and depression. *The Clinical Neuropsychologist, 30*(7), 1126–1150. https://doi.org/10.1080/13854046.2016.1195014

Valle, F., & Cuetos, F. (1995). *EPLA: Evaluación del procesamiento lingüístico en la afasia [EPLA: Evaluation of Language Processing in Aphasia]*. Hove, UK: Lawrence Erlbaum associates Ltd.

van Hell, J. G., & Tanner, D. (2012). Second language proficiency and cross-language lexical activation. *Language Learning, 62*(s2), 148–171. https://doi.org/10.1111/j.1467-9922.2012.00710.x

Verreyt, N., De Letter, M., Hemelsoet, D., Santens, P., & Duyck, W. (2013). Cognate effects and executive control in a patient with differential bilingual aphasia. *Applied Neuropsychology Adult, 20*(3), 221–230. https://doi.org/10.1080/09084282.2012.753074