Brain & Language 269 (2025) 105616

Contents lists available at ScienceDirect

Brain and Language

journal homepage: www.elsevier.com/locate/b&l

Research Article

Language mixing in people with aphasia: A cross-linguistic analysis using the 4M model

Aviva Lerman [a,b,*], Alina Bihovsky [c,d], Natalia Meir [c,e], Michal Ben-Shachar [e], Mira Goral [f]

[a] *Communications Disorders Department, Jerusalem Multidisciplinary College, Jerusalem, Israel*
[b] *Rehabilitation Department, Hadassah Hospital, Mt. Scopus, Jerusalem, Israel*
[c] *Department of English Literature and Linguistics, Bar-Ilan University, Ramat Gan, Israel*
[d] *Rehabilitation Hospital - Sheba Medical Center, Ramat Gan, Israel*
[e] *The Gonda Multidisciplinary Brain Research Center, Bar-Ilan University, Ramat Gan, Israel*
[f] *Speech-language-hearing Sciences, Lehman College and the Graduate Center, City University of New York, NY, USA*

ARTICLE INFO

Keywords:
4M model
Aphasia
Multilingual
Language mixing

ABSTRACT

Language mixing (LM) is common among individuals who speak two or more languages. LM in healthy multilingual individuals has been studied extensively, often by applying the Matrix Language Frame (MLF) and the 4M model, both developed by Myers-Scotton and colleagues. Our study examined LM in multilingual people with aphasia (PWA), to determine: (1) frequency of LM in the first-acquired language (L1) and later-acquired languages (L2/L3); (2) whether the 4M model accounts for LM patterns; and (3) how language-specific characteristics affect LM.

Narratives produced by 18 multilingual PWA (English-Hebrew; English-Spanish, English-Hebrew-Spanish and Russian-Hebrew speakers) with and without agrammatism were analyzed using the 4M model. More LM was observed in L2 contexts compared to L1 in participants with agrammatism, yet in participants without agrammatism no differences were found between L1 and L2. The patterns of LM in target and non-target languages were similar across L1 and L2: content morphemes were more likely to be embedded into the matrix language, aligning with previous research on healthy multilingual controls. Differences were found between the types of LM across the four studied languages.

Our findings indicate that multilingual PWA mix languages in a typical manner, similar to neurotypical multilingual individuals, aligning with the predictions of the 4M model. Furthermore, in multilingual PWA, as in healthy multilingual people, differences in LM across languages are related to the morphological richness of the matrix language and the congruency of the languages involved in LM. Clinically, based on our results, LM can be encouraged as a communication aid.

1. Introduction

Language mixing (LM) is a common practice among bilingual and multilingual speakers, as discussed in a seminal paper by Bentahila and Davies (1992). The motivation for and the characteristics of LM produced by neurologically healthy people have been studied extensively over the years (e.g., MacSwan, 2022; Muysken, 2000; Myers-Scotton, 1993; Myers-Scotton & Jake, 2000, 2017; Nguyen, 2024; Walters, 2005). In multilingual people who have aphasia, an acquired language impairment, LM has been a matter of controversy in the literature, with some researchers referring to it as pathological (Fabbro et al., 2000; Fyndanis, & Lehtonen, 2021; Perecman, 1984) and others suggesting it is used as a communication strategy (Goral et al., 2019; Lerman et al., 2019). In the literature on LM in aphasia, the frequency of mixing in connected speech has been studied (Muñoz et al., 1999; Paplikar, 2016), as has the distribution of word types most likely to be mixed (Bihovsky et al., 2023; Paplikar, 2016), but little has been reported about the internal structure of the mixed elements. The aim of our cross-linguistic study is to examine LM in a diverse sample of people with aphasia (PWA) who are speakers of typologically different languages, to

* Corresponding author at: Communications Disorders Department, Jerusalem Multidisciplinary College, Jerusalem, Israel.
E-mail address: avivale@edu.jmc.ac.il (A. Lerman).

https://doi.org/10.1016/j.bandl.2025.105616
Received 25 December 2024; Received in revised form 23 April 2025; Accepted 3 July 2025
Available online 24 July 2025

determine whether PWA use LM in a typical or atypical manner.

1.1. Language mixing in neurotypical multilingual people

One hallmark characteristic of neurotypical multilingual people is their ability to mix their languages or keep them separate, according to the communicative situation (Green & Abutalebi, 2013; Grosjean, 2001). Highly proficient multilingual people effortlessly switch their languages between utterances within a conversation and mix words and phrases from different languages within utterances when speaking with other people who share their languages (Grosjean, 2001; Muysken, 2000). The grammar characterizing LM and the pragmatic rules that determine the frequency and the directionality of it have been studied extensively (Alexiadou & Lohndal, 2018; MacSwan, 2022; Muysken, 2000; Myers-Scotton, 1993; Poplack, 1980; Walters, 2005). For example, the internal structure of LM has been examined, and models to explain it have been suggested (Myers-Scotton, 1993; Myers-Scotton & Jake, 2000, 2017).[1] Word-internal LM patterns across different languages have been discussed in numerous language pairs within the Matrix Language Frame (MLF) and the 4M models, both developed by Myers-Scotton and colleagues, to account for the types of morphemes that are likely to be mixed within utterances and to understand the linguistic relations between pairs of languages, which can predict acceptable mixing.

The MLF model aims to capture the structure of LM by identifying the matrix language of each utterance, which provides its main structure, and the embedded language or guest language, from which the mixed elements are inserted (Myers-Scotton, 2002). The embedded language may encompass a single lexical item (e.g., a noun, a preposition) or an island, comprising a phrase (e.g., prepositional phrase, embedded clause). The MLF distinguishes between two types of morphemes: content morphemes (e.g., a noun) and system morphemes (e.g., plural affix) (Myers-Scotton, 2002; Myers-Scotton & Jake, 2017).

The 4M model also identifies the matrix language and the embedded language but further classifies three specific types of system morphemes in the embedded language (Myers-Scotton & Jake, 2017). These morphemes include content morphemes (e.g., simple content morphemes such as singular nouns or zero-marked verbs; multimorphemic content morphemes such as plural nouns or inflected verbs), early system morphemes (e.g., determiners, conjunctions) and late system morphemes (e.g., late bridge system morphemes such as possessive markers, and outsider late system morphemes such as inflections, cases etc.) (Myers-Scotton & Jake, 2000; 2009; 2017). In multilingual people *without* aphasia, content morphemes are mixed more than early system morphemes which are mixed more than late system morphemes (Myers-Scotton & Jake, 2000; 2017). Furthermore, the matrix language has been observed to be driving the morpheme-related type of LM that is occurring, and not the embedded language, across an array of language pairs (Myers-Scotton & Jake, 2000; 2009).

Embedded constituents do not always appear as single lexical-item insertions; they can also occur as embedded islands, defined as "well-formed constituents in the embedded language that show internal structural relations" (Jake et al., 2002, p. 74). These embedded islands (e.g., determiner phrases like 'the box', adjectival phrases like 'the big box', and prepositional phrases like 'in the big box') are formed according to the grammar of the embedded language and are then inserted into the matrix language. The embedding of islands varies across speakers and languages. Jake et al. (2002) report that well-formed embedded islands vary across language pairs (e.g., approximately 10 % of English-Spanish LM data and 85 % of Swiss German-Italian LM data). The authors suggest that islands are more likely to occur in language pairs where the embedded material is incongruent with the matrix language. They propose that, in addition to phi-feature incongruence between German and Italian (i.e., person, number, and gender feature), case features also play a role.

Furthermore, cross-linguistic typological differences have been observed regarding other components of the matrix and embedded languages (Nguyen, 2024) as well as with respect to the categorization of morpheme class (Myers-Scotton & Jake, 2017). For example, pronouns can be categorized differently within and across languages, depending on the grammatical function within an utterance in any given language (Myers-Scotton & Jake, 2017). Thus, when analyzing data based on the 4M model, it is important to take language typology into account, and to look at language pairs with varying degrees of morphological complexity, both in neurotypical multilingual people and in multilingual PWA.

This study looked at the internal structure of within-utterance LM in languages with varying morphological richness (English, Hebrew, Russian, and Spanish). Whereas English has relatively sparse morphology, the other three languages have complex morphological systems, which include, for example, the rich case system in Russian, and overt markings of grammatical gender agreement in Hebrew, Russian, and Spanish. Analyses were conducted based on the 4M model, first by establishing the matrix language for each utterance, and then by analyzing the embedded components. However, rather than investigating neurotypical speakers of these languages, all participants in our study had aphasia.

1.2. Aphasia

Aphasia is an acquired language impairment resulting from a brain injury. Stroke-induced aphasia refers to aphasia resulting from a cerebrovascular accident in the language-dominant hemisphere of the brain, typically the left hemisphere. People who sustain a cerebrovascular accident resulting in aphasia experience difficulties in various linguistic abilities, including speaking, reading, writing, and comprehending spoken and/or sign language, to varying degrees (Berthier, 2005; Papathanasiou & Coppens, 2021). Aphasia may affect different linguistic processes and representations, as in the case of *agrammatism*, defined as a syntactic deficit that affects sentence comprehension, sentence production or both (Bastiaanse & Thompson, 2012; Faroqi-Shah, 2023; Goodglass & Menn, 1985). Many PWA improve their linguistic and communication abilities with time and with intervention, especially in the first weeks and months after the cerebrovascular accident. However, many sustain lasting impairments, chronically affecting their ability to produce and comprehend language (Raymer et al., 2008).

Those PWA who use more than one language in their daily communication prior to the onset of aphasia typically experience language difficulties in all their languages. However, the degree of impairment and path to improvement may differ across languages, depending on variables such as the degree of language proficiency prior to the stroke, the degree of language use and exposure before and after the stroke, and attitudes toward the languages, among others (Albert & Obler, 1978; Kiran & Gray, 2018; Paradis, 2004; Peñaloza et al., 2020). Particularly for multilingual people with agrammatic aphasia, the clinical expression of language impairment may be differential across languages depending on variations in syntactic structure and the relative

[1] We note that the terminology used in different models varies and is also somewhat contradictory across different fields. In linguistic models such as the MLF and 4M models, the term *Codeswitching* is used to define all instances of language mixing within a clause. However, in the aphasia literature *Codeswitching* is often used to denote the switching of languages across utterance boundaries, whereas *Codemixing* is used to denote the mixing of languages within an utterance, including within clauses within a given utterance. Furthermore, the term *Language mixing* is used in the aphasia literature more generally to encompass both codeswitching and codemixing, whereas in the field of linguistics is has been used more specifically for mixing within a single word. To avoid confusion, we do not use the term *Codeswitching* in this paper, rather, we use the more general term *Language mixing* to encompass all types of codeswitching and codemixing and we specify between-utterance language mixing and within-utterance language mixing where appropriate.

morphological richness of their languages (Abuom & Bastiaanse, 2012; Bastiaanse & Thompson, 2012; Menn et al., 1990). These dynamic patterns of relative impairment in the different languages of multilingual PWA are of great theoretical and clinical interest (Lerman et al., 2020; Peñaloza et al., 2020). In the current study, we analyzed LM in PWA with and without agrammatism, taking into account the participants' languages, relative proficiencies, and aphasia severity and type.

1.3. Language mixing in aphasia

Like all multilingual people, multilingual PWA mix their languages. However, evidence for excessive or unintentional mixing led scholars to argue that PWA may present with so called *pathological mixing*, consisting of inappropriate and/or involuntary mixing (Fabbro et al., 2000; Fyndanis, & Lehtonen, 2021; Perecman, 1984). Early reports of pragmatically inappropriate or unintentional between-utterance LM raised the possibility that aphasia affects pragmatic aspects of LM or that the brain lesion affects the control mechanisms that typically allow proper language selection (Green & Abutalebi, 2008). In contrast, in many instances, mixing – voluntary or not – has been used by PWA primarily as a communication strategy (Goral et al., 2019; Grosjean, 1985; Lerman et al., 2019). For example, if PWA cannot momentarily retrieve a word while speaking in one language, they may be able to access it in their other language, and thus avoid a communication breakdown.

Reports of LM in PWA focused not only on the pragmatic appropriateness of the mixing, but also on the frequency and the directionality of LM among multilingual speakers of varying linguistic abilities and language combinations. In terms of the frequency of LM, greater frequency of within-utterance LM in connected speech has been reported for PWA compared to people without aphasia (e.g., Muñoz et al., 1999; Paplikar, 2016). For example, Paplikar's (2016) data from 32 PWA and 32 people without aphasia, speakers of Kannada and English, showed greater numbers of words mixed in the production of PWA compared to people without aphasia. PWA also exhibited more between-utterance LM than did the neurologically healthy participants (Paplikar, 2016). Nevertheless, the data showed similar distribution of word types among the mixed words in the two participant groups, with nouns being the more frequently mixed word type. Finally, Paplikar (2016) found a similar pattern of mixing for the participants with and without aphasia in terms of the direction: both groups were more likely to have a higher LM ratio when the target language was L1-Kannada compared to when the target language was L2-English.

Whereas the directionality of the mixing in the Kannada-English participants is consistent with the pragmatic norms of their community (Paplikar, 2016), the directionality of LM in PWA can also pattern with relative proficiency or post-stroke abilities. People with and without aphasia have been observed to mix words from their stronger language while speaking in a language of lower abilities (Bihovsky et al., 2023; Goral et al., 2019). For example, Goral et al. (2019) reported LM data for 11 PWA speakers of various languages, and demonstrated that in all cases when there was a difference in the level of post-stroke abilities between the two languages, the LM ratio was higher when the target language was the weaker language. This was observed in both PWA with agrammatism ($n = 4$) and PWA without agrammatism ($n = 7$). A similar pattern was observed in LM data for 20 Russian-Hebrew PWA (Bihovsky et al., 2023). Consistently, people with more severe aphasia produced more instances of LM than people with milder aphasia (e.g., Bihovsky et al., 2023; Goral et al., 2019).

Little has been reported in the literature about the internal structure of the mixed elements in mixed language production of multilingual PWA. One exception includes Paplikar (2016), who used the MLF model to explain observed LM in her participants. She found that contrary to the general pattern of LM (more mixing from English while speaking Kannada), bound morphemes showed an opposite pattern. Thus, a higher LM ratio was observed for bound morphemes when the target language was L2-English compared to when the target language was L1-Kannada.

1.4. The current study

The aim of our study was to examine LM in multilingual PWA, speakers of varying language pairs, to determine whether within-utterance LM in PWA occurs in a typical or atypical manner relative to morphological structure. In the field of aphasia, the 4M model has previously been applied to language produced by monolingual people (Myers-Scotton & Jake, 2000). The authors found support for their hypotheses that the abstract level at which a morpheme is selected will affect the frequency of selection of one morpheme over another. This preferential selection could theoretically occur in either the same or a different language. However, while the MLF has been applied previously to LM data from multilingual PWA (Paplikar, 2016), the 4M model has not been directly applied to the output of this population. We examined connected language production elicited with and without visual material from 18 PWA, speakers of three different language combinations, using the 4M model as our framework. The study was designed to answer the following research questions about multilingual PWA with and without agrammatism:

I. Are there differences in the frequency of LM in L1 vs. L2?

We hypothesized that more LM would occur when speaking a weaker L2 than when speaking L1, and that the distribution of LM type within L1 and within L2 would be comparable, similar to previous studies on PWA and studies on neurotypical multilingual populations (Goral et al., 2019; Lerman et al., 2019; Myers-Scotton & Jake, 2000, 2017). Based on the literature to date on group studies of people with varied types of aphasia (i.e., with or without agrammatism), there is no reason to expect more LM for those with agrammatism than those without agrammatism (e.g., Goral et al., 2019), thus we predict similar frequencies of LM for each.

II. Does the 4M model account for within-utterance LM patterns?

We hypothesized that patterns of LM in multilingual PWA, with and without agrammatism, would be accounted for by the 4M model, following the findings for mixing patterns in people without aphasia (Myers-Scotton & Jake, 2017). Furthermore, we predicted that more LM would occur for the more preserved content morphemes than for the less preserved system morphemes, based on the findings reported in Myers-Scotton and Jake (2000) for monolingual people with agrammatic aphasia.

III. Do language-specific characteristics affect within-utterance LM?

We hypothesized that language-specific characteristics would affect LM, with LM being determined, at least in part, by the morphological characteristics of each language (Nguyen, 2024). Specifically, we predicted that morphologically complex words would be mixed as a complete unit, and therefore this would occur more in morphologically rich languages (e.g., Hebrew, Russian, Spanish) than in morphologically sparse languages (e.g., English). Alternately, we predicted that the morphological structure of English would allow for the mixing of single, morphologically simple words and that this would thus occur in English more than in the other languages.

2. Methodology

The materials, data, and analysis script for this study can be retrieved from the OSF depository: https://osf.io/jw8fy/?view_only=aa07e1d09f564e5abd2f00212f2c0c53.

2.1. Participants

Our study included 18 multilingual PWA whose ages ranged from 41 to 74 years old (Mean age = 61.94, SD = 8.97). Six participants spoke English and Hebrew, four participants spoke English and Spanish, one participant spoke English, Spanish and Hebrew, and seven participants spoke Russian and Hebrew. Some participants also spoke other languages to varying proficiencies, but they were not in everyday use at the time of testing. All participants were in the chronic stage of rehabilitation after a stroke or brain tumor. Demographics for each participant can be seen in Table 1, including age of acquisition of non-L1s, and relative reported pre-morbid proficiencies and post-morbid language abilities, based on the participants' self-report. L1 was considered the first-acquired language.

Participants were included in this study retrospectively from three separate labs, after having been recruited previously to take part in a variety of different studies on multilingual people with aphasia. In all previous studies, narrative data had been collected in each of the participants' daily-used languages before receiving any treatment in that particular study. Only participants who spoke English, Hebrew, Russian, and/or Spanish were included in this study. For two labs (English-Hebrew speakers, English-Spanish speakers, and an English-Spanish-Hebrew speaker), all participants who were able to produce narratives (i.e., more than one utterance together for the purpose of communicating one idea) were included in the study. For the third lab (Russian-Hebrew speakers), seven participants from a pool of 20 participants were quasi-randomly selected to be a part of this current study (every third participant in their data pool when the data pool was ordered by date of recruitment).

2.2. Procedure

Transcriptions of story sequences and personal narratives were collected from all participants. Story sequences comprised four, six or twelve drawings presented to the participants who were asked to tell the story depicted in the sequence; personal narratives were elicited by asking participants to talk about a given topic (e.g., tell me about a vacation you took) for a few minutes. The transcriptions were organized into utterances. Utterances were defined as an independent clause or sub-clausal unit, plus any associated subordinate clauses, also accounting for intonation and pauses to best represent units of spoken speech (Lerman, 2020).

First, the number of total verbal units (TVUs) were calculated for all narratives. TVUs included all words, parts of words that were clear part-word of a recognizable word (including repeated words and repeated phrases), and interjections, produced in any language. The number of TVUs produced in the non-target language was calculated as a percentage of all TVUs. Second, further analyses were conducted only on narratives with at least one instance of LM. For these narratives, we coded utterances relative to within-utterance LM, which was defined as producing morphemes, words, or phrases in a language other than the matrix language of the utterance (Myers-Scotton & Jake, 2017, 2000). Each utterance was coded as to whether it contained mixing or not, whether a matrix language could be identified, and whether that matrix language was the target language or not, yielding five categories: 1) having no within-utterance LM and produced in the target language (Non-LM-target), 2) having no within-utterance LM and produced in the non-target language (Non-LM-non-target), 3) having a matrix language (e.g., English, Hebrew, Russian, or Spanish) that is also the target language and included at least one instance of within-utterance LM (Matrix target), 4) having a matrix language that is a non-target language and included at least one instance of within-utterance LM (Matrix non-target), 5) having words that were produced in more than one language within the utterance but no matrix language was determined (Alternation).

Third, for all the utterances that were coded as having a matrix language and included at least one instance of within-language LM (Matrix-target and Matrix non-target), the number of incidences of LM were counted and analyzed. Analyses of within-utterance LM were conducted based on the 4M model (Myers-Scotton & Jake, 2017) as a framework, and we coded morphemes, words and/or phrases for simple content words (content-s), multimorphemic content words (content-m), early system morphemes (ESM), late bridge system morphemes, outsider late system morphemes, and complete and incomplete islands (see Table 2 for examples). Quantitative and qualitative analyses were then conducted on the data to identify the patterns of LM across language pairs, language dominance and/or severity of aphasia.

2.3. Statistical analysis

To analyze the LM data, we employed generalized linear mixed models using the glmmTMB (Generalized Linear Mixed Models using Template Model Builder) procedure in R (R Core Team, 2023). This approach effectively handled our sparse, continuous yet skewed data by utilizing a Gamma distribution with a log link function, allowing for robust estimation of fixed effects while accounting for Participant variability as a random effect.

For RQ1, we examined predictors of the percentage of LM included by fitting a model with Agrammatism (YES = 1, NO = 0), Target Language (L1 vs. L2), and their interaction as fixed effects, with Participant as a random effect. To further investigate LM across L1 and L2 in participants with and without agrammatism, we expanded the model to include Agrammatism, Target Language, and Utterance Type (Non-LM-Target, Non-LM-nonTarget, Matrix-Target, Matrix-non-Target, Alternation) as fixed effects, again incorporating Participant as a random effect.

For RQ2, we focused exclusively on Matrix-target and Matrix-non-target utterances and fitted a generalized linear mixed model in which the percentage of LM distribution served as the dependent variable. The fixed effects included Agrammatism, Target Language, and LM Type (content-s, content-m, ESM, complete island, incomplete island), along with their interaction, while Participant was included as a random effect to account for individual variability.

For RQ3, to analyze LM data across four matrix languages, we fitted another generalized linear mixed model, with Agrammatism, Matrix Language (English, Hebrew, Russian, and Spanish) and LM Type (content-s, content-m, ESM, complete island, incomplete island) as fixed effects, along with their interaction and Participant as a random effect. English was set as the reference category for the Matrix Language variable, and content-s was the reference for LM Type. Significant interactions were further examined using estimated marginal means (EMMs) with Bonferroni correction to adjust for multiple comparisons, ensuring a rigorous interpretation of the results. By leveraging glmmTMB, we ensured flexibility and precision in modeling our complex data structure, accounting for both fixed and random effects while effectively handling sparsity and skewness in the LM distribution.

3. Results

3.1. RQ1: Are there differences in the frequency of LM in L1 vs. L2 in multilingual PWA with and without agrammatism?

A total of 198 narratives were produced by the 18 multilingual PWA included in the study (L1: 96, L2: 102, see Table 3). The percentage of TVUs in the non-target language was 2 % while speaking in L1 and 7 % while speaking in L2 out of all 198 narratives. Ninety-nine narratives (i.e., 50 % of all narratives) contained at least one instance of LM (and therefore were included in the analyses of LM patterns): 33 out of the 96 narratives produced during L1 testing sessions (34 %), and 66 out of the 102 narratives produced during L2 testing sessions (65 %). In these 99 narratives, the majority of utterances were produced in the target language (70 % of 1032 utterances).

The model investigating the predictors of percentage of LM revealed

Table 1
Participant demographics.

P no.	Sex	Age (yrs)	Languages	AoA for non-L1s (yrs)	Type and severity[a] of aphasia in L1	Type and severity[a] of aphasia in L2	Pre-morbid prof (L1:L2)	Post-morbid abilities (L1:L2)	Dom pre-morbid[b]	Dom post-morbid [b]
RH201	F	41	L1-Rus, L2-Heb	13	Non-fluent Severe (Agrammatism)	Non-fluent Severe (Agrammatism)	L1: 26/28 L2: 28/28	L1: 10/28 L2: 13/28	L2-Heb	L2-Heb
RH204	F	61	L1-Rus, L2-Heb	35	Non-fluent Severe (Agrammatism)	Non-fluent Severe (Agrammatism)	L1: 28/28 L2: 25/28	L1: 13/28 L2: 11/28	L1-Rus	L1-Rus
RH207	F	50	L1-Rus, L2-Heb	23	Non-fluent Moderate (Agrammatism)	Non-fluent Moderate (Agrammatism)	L1: 28/28 L2: 22/28	L1: 13/28 L2: 17/28	L1-Rus	L2-Heb
RH210	M	67	L1-Rus, L2-Heb	27	Non-fluent Severe (Agrammatism)	Non-fluent Severe (Agrammatism)	L1: 28/28 L2: 18/28	L1: 8/28 L2: 8/28	L1-Rus	L1-Rus = L2-Heb
RH213	F	70	L1-Rus, L2-Heb	23	Anomia Moderate	Fluent Moderate	L1: 28/28 L2: 28/28	L1: 21/28 L2: 18/28	L1-Rus = L2-Heb	L1-Rus
RH216	M	60	L1-Rus, L2-Heb	31	Anomia Mild	Non-fluent Moderate (Agrammatism)	L1: 28/28 L2: 15/28	L1: 17/28 L2: 10/28	L1-Rus	L1-Rus
RH219	F	64	L1-Rus, L2-Heb	43	Fluent Mild	Non-fluent Moderate (Agrammatism)	L1: 28/28 L2: 17/28	L1: 21/28 L2: 15/28	L1-Rus	L1-Rus
HE01	M	71	L1-Heb L2-Eng	11	Non-fluent (Agrammatism)	Non-fluent, (Agrammatism)	L1: 10/10 L2: 8/10	L1: 10/10[^] L2: 8/10	L1-Heb	L1-Heb
EH02	M	65	L1-Eng L2-Heb	7	Anomia Mild	Non-fluent Severe	L1: 10/10 L2: 7/10	L1: 7/10 L2: 5/10	L1-Eng	L1-Eng
EH03	M	66	L1-Eng L2-Heb	26	Non-fluent, Moderate (Agrammatism)	Non-fluent, Severe (Agrammatism)	L1:30/30 L2:27/30	L1:22/30 L2:11/30	L1-Eng	L1-Eng
EH04	M	65	L1-Eng L2-Heb	19	Non-fluent, Moderate (Agrammatism)	Non-fluent, Severe (Agrammatism)	L1:30/30 L2:30/30	L1:25/30 L2:20/30	L1-Eng = L2-Heb	L1-Eng
HE05	M	60	L1-Heb L2-Eng	7	Fluent Mild	Fluent Mild	L1: 10/10 L2: 9.5/10	L1: 7.75/10 L2: 6.75/10	L1-Heb	L1-Heb
HE06	M	73	L1-Heb L2-Eng	4	Fluent Moderate-severe	Fluent Moderate-severe	L1: 9/10 L2: 10/10	L1: 2/10 L2: 2/10	L2-Eng	L1-Heb = L2-Eng
SE02	F	51	L1-Spn L2-Eng	4	Anomia, Residual	Anomia, Residual	L1:10/10 L2:10/10	L1:7/10 L2:7/10	L2-Eng	L2-Eng
ES06	F	53	L1-Eng L2-Spn	3	Anomia, Residual	Anomia, Residual	L1: 10/10 L2:10/10	L1: 5/10 L2: 3/10	L1-Eng = L2-Spn	L1-Eng
SE07	M	74	L1-Spn L2-Eng	14	Non-fluent, Mild	Non-fluent, Mild (Agrammatism)	L1: 10/10 L2:9/10	L1: 6/10 L2:6/10	L1-Spn	L1-Spn

(continued on next page)

Table 1 (*continued*)

P no.	Sex	Age (yrs)	Languages	AoA for non-L1s (yrs)	Type and severity[a] of aphasia in L1	Type and severity[a] of aphasia in L2	Pre-morbid prof (L1:L2)	Post-morbid abilities (L1:L2)	Dom pre-morbid[b]	Dom post-morbid [b]
SE09	F	55	L1-Spn L2-Eng	13	Non-fluent Moderate	Non-fluent Moderate	L1: 10/10 L2:10/10	L1: 5/10 L2:3/10	L1-Spn = L2-Eng	L1-Spn
ESH12	F	69	L1-Eng L2-Spn L2-Heb	3, 5	Fluent Mild-Moderate	Fluent – moderate L2 Spn/Heb – moderate-severe	L1: 10/10 L2 Spn: 7/10 L2 Heb: 6/10	L1: 6/10 L2 Spn: 3/10 L2 Heb: n/a	L1-Eng	L1-Eng

Note. P = Participant; Rus = Russian; Heb = Hebrew; Eng = English; Spn = Spanish; AoA = Age of acquisition; L1 = first acquired language; L2 = second acquired language; prof = proficiency; Dom = dominance.

[a] Type and aphasia severity were determined by a clinician in each lab using the Western Aphasia Battery Revised, or a similar aphasia battery in any given language.

[b] Dominance was based on self-reports of pre- and post-morbid language abilities; post-morbid dominance was confirmed by the researchers based on language assessments.

^ HE01 rated his language skills in Hebrew as 10/10 both pre- and post-morbidly, even though he clearly had aphasia in Hebrew, as observed by a variety of language assessments.

Table 2
Categories and examples of the 4M-based analyses.

Category	Types (examples)	Specific examples from our data
Simple content words (content-s)	singular nouns, zero-marked verbs, discourse markers, pronouns, prepositions that assign theta/conceptually activated, adverbs	*'mattress'* *'argue'* *'OK'* *'really'* *'now'*
Multimorphemic content words (content-m)	e.g., plural or compound nouns, inflected verbs, verbs with particles, derivatives	*'mattresses'* *'uncomfortable'*
Early system morphemes (ESM)	e.g., determiners, conjunctions	*'ve'* in: *'moon ve star'* – 'and' in 'moon and star' *'ve'* in: *'odin brat ve sestra'* – 'and' in 'one brother and sister'
Late bridge system morphemes	e.g., possessive markers	n/a
Outsider late system morphemes	e.g., inflections, cases	n/a
Complete or incomplete islands	e.g., prepositional phrases, noun phrases, complex verb phrases	*'look lost'* – complete *'pointing in the...'* – incomplete

that the interaction between Target Language and Agrammatism was significant, B = 0.24, SE = 0.04, $\beta = 5.88$, $p < 0.001$. However, the main effects of Target (B = 0.00, SE = 0.03, $\beta = 0.07$, p = 0.942) and Agrammatism (B = −0.04, SE = 0.04, $\beta = -1.10$, p = 0.273) were not significant. For non-agrammatic participants, the pairwise contrast between L1 and L2 was not significant, B = −0.00197, SE = 0.0270, t(187) = -0.073, p = 0.9421; whereas for agrammatic participants the pairwise contrast between L1 and L2 was significant, B = −0.24112, SE = 0.0286, t(187) = -8.428, $p < 0.001$, indicating that agrammatic individuals had significantly higher rates of LM in L2 compared to L1. For group results See Fig. 1, Panel A, yet individual variation should be noted (see Fig. 1, Panel B).

Fig. 2 illustrates the distribution of utterance types across target languages (L1 vs. L2) in participants with and without agrammatism. A mixed-effects model analyzing utterance types across L1 and L2 (see Table 4) revealed no significant main effects for Agrammatism (B = 1.42, SE = 1.73, z = 0.29, p = 0.774) or Target Language (B = 0.89, SE = 1.06, z = -0.09, p = 0.925). However, a significant main effect was observed for LM utterance type, particularly for Matrix-nonTarget utterances (B = 0.01, SE = 0.02, z = −3.57, $p < 0.001$), suggesting variability in usage across utterance types. Crucially, there was a significant interaction between Agrammatism and LM type for Matrix-nonTarget utterances (B = 0.00, SE = 0.00, z = -6.87, $p < 0.001$), indicating that this utterance type was markedly more frequent in the agrammatic group. No other two- or three-way interactions reached significance. Pairwise comparisons confirmed that agrammatic participants produced significantly more Matrix-nonTarget utterances than non-agrammatic participants in both L1 ($p < 0.0001$) and L2 ($p < 0.0001$), while such utterances were virtually absent in the non-agrammatic group. A trend-level difference was also noted for Alternation utterances in L1 (p = 0.0723), with agrammatic participants tending to produce them more frequently. No significant differences emerged for Non-LM-Target, Non-LM-nonTarget, or Matrix-Target utterances across language groups, and no significant interactions were found involving the target language. These findings underscore the specificity of Matrix-nonTarget utterance production as a distinguishing feature of agrammatism, independent of the target language.

We also examined contrasts between L1 and L2 across different language-mixed utterance types in participants with and without agrammatism. The analysis revealed no significant differences between L1 and L2 in any language-mixed utterance type.

3.2. RQ2: Does the 4M model account for within-utterance LM patterns in multilingual PWA with and without agrammatism?

LM was analyzed in Matrix-target and Matrix-non-target utterances (n = 175). Fig. 3 illustrates the types of LM observed in L1 and L2 among participants with and without agrammatism, based on the 4M model. We found that in both L1 and L2, content words (content-s and content-m) were the most frequently mixed, with some early system morphemes (ESM) mixed in L2 in both participant groups. No instances of late bridge system morphemes or outsider late system morphemes were observed. Regarding islands, between 5 %-19 % of LM instances in L1 and L2 were either complete or incomplete islands, with 43 % of these being narrative comments.

Our model for RQ2 revealed significant main effects of LM type, with both ESM and incomplete island utterances showing markedly lower proportions compared to the baseline (content-s), as reflected by their near-zero estimates (both $p < 0.001$; see Table 5). In contrast, the main effects of Target Language (L1 vs. L2) and Agrammatism (Non-Agrammatism vs. Agrammatism) were not statistically significant. However, crucial three-way interactions emerged, notably Agrammatism × Target Language × Mtype for both ESM (p = 0.009) and incomplete islands ($p < 0.001$). Pairwise comparisons confirmed that agrammatic participants produced significantly more ESM and incomplete island utterances than non-agrammatic participants in L1 (both $p < 0.001$), while no significant group differences were observed in L2. Additionally, within-group

Table 3
Characteristics of the narratives per participant.

	# Total Narratives L1	# Total Narratives L2	# Total Narratives	# L1 Narratives with LM	# L2 Narratives with LM	# Total Narratives with LM	Sum TargetTVU L1	Sum TargetTVU L2	Sum LMTVU L1	Sum LMTVU L2	%LM L1-Target (out of ALL narratives)	%LM L2-Target (out of ALL narratives)
RH201	2	2	4	2	2	4	210	179	34	7	14 %	4 %
RH204	6	6	12	1	6	7	265	93	2	65	1 %	41 %
RH207	6	6	12	6	5	11	381	412	64	20	14 %	5 %
RH210	6	6	12	3	4	7	301	167	14	15	4 %	8 %
RH213	6	6	12	5	5	10	1299	967	44	47	3 %	5 %
RH216	6	6	12	0	4	4	616	709	0	12	0 %	2 %
RH219	6	6	12	0	6	6	631	582	0	85	0 %	13 %
EH02	6	6	12	0	1	1	380	138	0	5	0 %	3 %
EH03	6	6	12	1	6	7	164	47	1	81	1 %	63 %
EH04	6	6	12	1	6	7	178	22	1	28	1 %	56 %
EH05	3	3	6	2	0	2	402	438	4	0	1 %	0 %
EH06	2	3	5	1	0	1	385	478	4	0	1 %	0 %
HE01	6	5	11	2	5	7	598	235	8	108	1 %	31 %
SE02	5	6	11	4	0	4	569	654	7	0	1 %	0 %
SE06	6	5	11	0	3	3	852	817	0	7	0 %	1 %
SE07	6	6	12	1	3	4	989	1056	2	17	0 %	2 %
SE09	6	6	12	4	0	4	264	286	7	0	3 %	0 %
SE12	6	12	18	0	10	10	1402	2039	0	175	0 %	8 %
Totals	96	102	198	33	66	99	9886	9319	192	672	**2 %**	**7 %**
%s				34 % of L1 narratives	65 % of L2 narratives	**50 % of all narratives**						

Notes: Narratives with LM = narratives including at least one instance of language mixing; TVU = total verbal units; LMTVU = language-mixed total verbal units. %s = percentages.

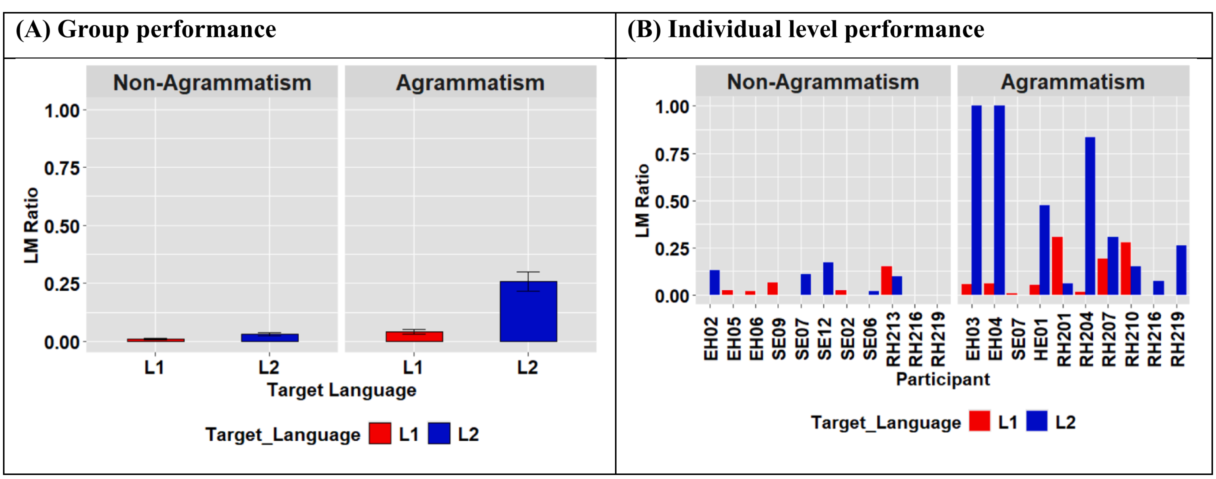

Fig. 1. Language mixing frequency in each language (L1 and L2) for participants with and without agrammatism (A) Group performance and (B) Individual level performance.

contrasts revealed that only agrammatic participants showed significant differences between L1 and L2, producing significantly more ESM and incomplete island utterances in L1 compared to L2 (both $p < 0.001$), while no such differences were observed in the non-agrammatic group. These results highlight that the disproportionate use of specific LM types – particularly ESM and incomplete islands – is a defining characteristic of agrammatic speech in the native language context, and such patterns are not mirrored in the non-agrammatic group or in L2 usage.

3.3. RQ3: Do language-specific characteristics affect within-utterance LM in multilingual PWA with and without agrammatism?

Fig. 4 shows the distribution of within-utterance LM type in each matrix language (English, Hebrew, Russian, and Spanish) across participants with and without agrammatism. The model addressing RQ3 revealed significant main effects and interactions of Matrix Language, LM type, and Agrammatism on LM percentage (see Table 6 and S1 in Supplementary materials). Content-m was nearly absent in English but frequent in Hebrew, Russian, and Spanish, with the effect especially pronounced among agrammatic participants (all $p < 0.001$). In contrast, for ESM, significant differences emerged primarily in the agrammatic group, where Hebrew had higher LM than Russian and Spanish (both $p < 0.001$), and English differed from Hebrew ($p = 0.0043$). In complete and incomplete islands, agrammatic participants exhibited robust differences across all languages, particularly with Hebrew and Spanish yielding the highest LM (all $p < 0.001$). Non-agrammatic contrasts were fewer but still revealed lower LM in English. Critically, agrammatic participants produced significantly more LM than non-agrammatic ones

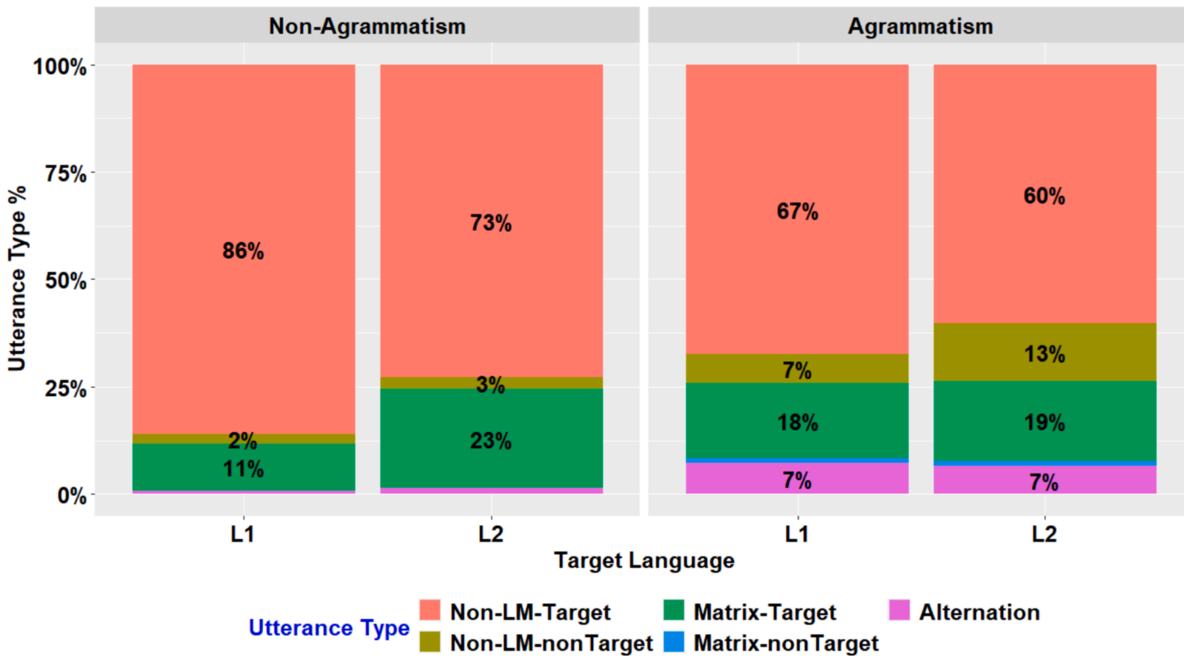

Fig. 2. Utterance type distribution for each target language: L1 and L2, in participants with and without agrammatism. (A) Group performance and (B) Individual level performance.

Table 4
Generalized Linear Mixed Model (GLMM) results for the effect of target language (L1 vs. L2), Utterance LM type and Agrammatism (YES, NO).

Predictors	*Estimates*	*SE*	*CI*	*Statistic*	*p*
(Intercept)	6.73	5.64	1.30–34.80	2.28	**0.023**
Agrammatism[Non-Agrammatism]	1.42	1.73	0.13–15.44	0.29	0.774
Target Language [L2]	0.89	1.06	0.09–9.13	−0.09	0.925
LMedUtterance[Non-LM-nonTarget]	0.10	0.12	0.01–1.04	−1.93	0.054
LMedUtterance[Matrix-Target]	0.26	0.31	0.03–2.66	−1.13	0.257
LMedUtterance[Matrix-nonTarget]	0.01	0.02	0.00–0.15	−3.57	**<0.001**
Agrammatism [Non-Agrammatism] × LMedUtterance [Matrix-nonTarget]	0.00	0.00	0.00–0.00	−6.87	**<0.001**
Random Effects					
σ^2					2.08
$\tau_{00\ Participant}$					0.00
ICC					0.00
$N_{Participant}$					18
Observations					190
Marginal R^2/Conditional R^2					0.898/0.898

in Spanish across all LM types (all $p < 0.0001$), and also in Russian and Hebrew for selected types. These results underscore how Matrix Language and LM type strongly influence LM patterns, with effects markedly heightened in agrammatic speech.

We further examined morphosyntactic accuracy of content-s language-mixed morphemes that were either nouns or verbs in participants with and without agrammatism. In the English-Hebrew PWA, we observed LM nouns or verbs in a Hebrew matrix five times – three in one PWA with agrammatism and one each in two PWA without agrammatism. For the participants without agrammatism, both instances of LM were nouns, and both were morphosyntactically correct in context (e.g., ‘*mattress*’ in the utterance ‘mattress ’אני הולך לישון על ה- = ‘[I go to sleep on the]$_{Heb}$ mattress’; ‘*business*’ in the utterance ‘business-שניהם גמרו את ה’ = ‘[The two of them finished the] $_{Heb}$ business’). For the participant with agrammatism, all three instances were zero-marked verbs; two were morphosyntactically correct in context and one was not (e.g., ‘*sad*’ in the utterance ‘sad ’ילד = ‘[A boy (is)] $_{Heb}$ sad’; ‘*argue*’ in the utterance ‘argue ’ילד וילדה = ‘[A boy and girl] $_{Heb}$ argue’; ‘*leave*’ in the utterance ‘leave ’וילדה = ‘[And the girl] $_{Heb}$ **leaves**’).

In Russian-Hebrew PWA, we observed LM of nouns or verbs 33 times in a Hebrew matrix. PWA with agrammatism produced six verbs and eight nouns in the non-target language. Two verbs were used in an incorrect infinitive form (e.g., ‘*to dance*’ in the utterance ‘דווקא מצוין танцевать ’אני = ‘[I actually] $_{Heb}$ **danced** [very well] $_{Heb}$’), whereas all nouns were morphosyntactically correct in context (e.g., ‘*homework*’ in the utterance ‘уроки ’=אבא אמר ‘[Dad said] $_{Heb}$ homework’). PWA without agrammatism produced LM of two verbs and 17 nouns, and all instances were morphosyntactically correct in context (e.g., ‘*took out*’ in the utterance ‘вылоЖила ’כל מני ספרים = ‘She took out [all kinds of books] $_{Heb}$’; ‘*cake*’ in the utterance ‘торт’אמא לקחה מגש עם ה- = ‘[Mother took a tray with the] $_{Heb}$ cake’). In the Russian matrix, 29 instances of noun and verb LM were observed, and all were morphosyntactically correct. PWA with agrammatism exhibited LM of one verb and 24 nouns, while PWA without agrammatism exhibited LM of four nouns (e.g., ‘*hat*’ in the utterance ‘посМотрел כובע и полоЖила’ = ‘Looked [hat] $_{Heb}$ and put down’; ‘*products*’ in the utterance ‘МуЖчина и Женщина ходили МеЖду מוצרים’= ‘A man and a woman walked between [products] $_{Heb}$’).

In the Spanish-English PWA, in the Spanish matrix, the person with agrammatism produced one morphosyntactically correct noun (e.g., ‘*subway*’ in the utterance ‘…que no habia tomado nunca el el el el subway’ = ‘[…that had never taken the the the] subway’) and the PWA without agrammatism produced five morphosyntactically correct nouns (e.g., ‘r*eceptionist*’ in the utterance ‘Cuando fuimo al receptionist’ = ‘[when we went to the] receptionist’). For the English matrix, the person with agrammatism produced three morphosyntactically correct nouns (e.g., ‘*cuña*’ in the utterance ‘…cuña to to to play the the the the the the baby’ = ‘crib [to to to play the the the the the the baby]’) and one morphosyntactically correct verb. The trilingual participant, who did not have agrammatism, produced two correct nouns while in a Hebrew matrix (e.g., ‘*city*’ in the utterance: ‘city’עד שנוכל לראות את כל ה- = [until we can see the whole] $_{Heb}$ city) and one correct verb. In a Spanish matrix

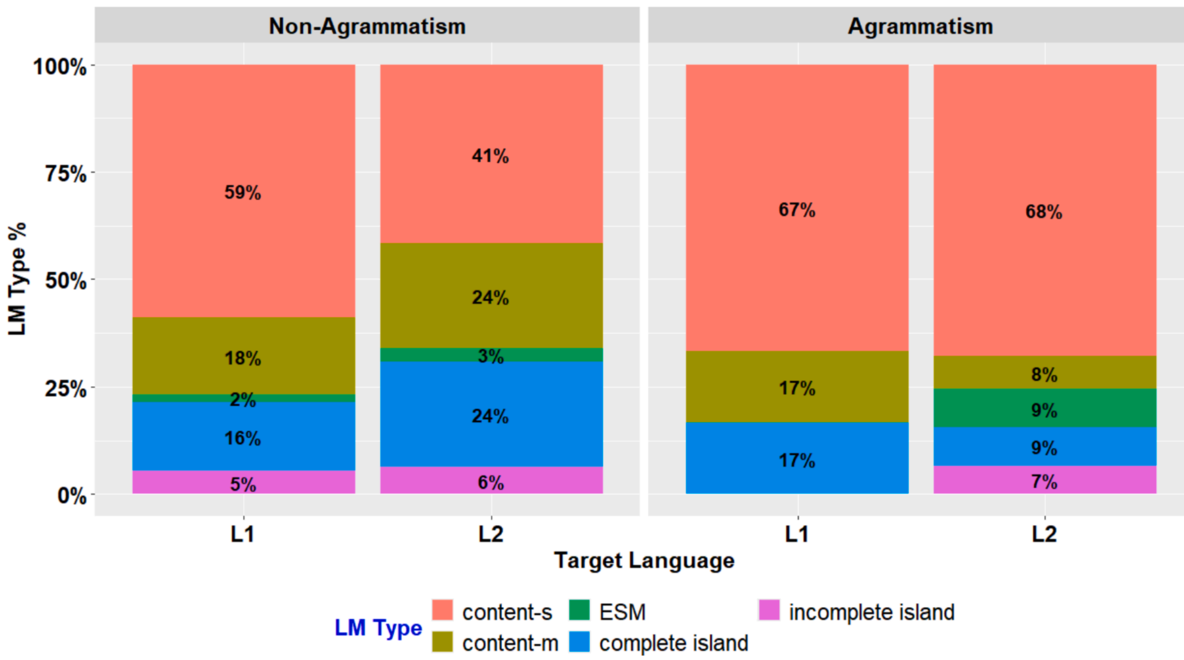

Fig. 3. Within-utterance LM type in L1 and L2 in participants with and without agrammatism (A) Group performance and (B) Individual level performance.

Table 5
Generalized Linear Mixed Model (GLMM) results for the effect of target language (L1 vs. L2) and within-utterance LM type.

	pct				
Predictors	*Estimates*	*std. Error*	*CI*	*Statistic*	*p*
(Intercept)	0.03	0.03	0.00–0.20	−3.47	**0.001**
Mtype [ESM]	0.00	0.00	0.00–0.00	−6.83	**<0.001**
Mtype [incomplete island]	0.00	0.00	0.00–0.00	−6.83	**<0.001**
Agrammatism [Non-Agrammatism] × Mtype[ESM]	2537.18	4682.25	68.16–94450.58	4.25	**<0.001**
Agrammatism [Non-Agrammatism] × Mtype[incomplete island]	8828.15	16292.08	237.14–328652.12	4.92	**<0.001**
Target Language [L2] × Mtype [ESM]	5941.50	12501.77	96.13–367236.75	4.13	**<0.001**
Target Language [L2] × Mtype [incomplete island]	36254.75	76285.41	586.56–2240887.77	4.99	**<0.001**
(Agrammatism [Non-Agrammatism] × Target Language [L2]) × Mtype [ESM]	0.00	0.00	0.00–0.17	−2.63	**0.009**
(Agrammatism [Non-Agrammatism] × Target Language [L2]) × Mtype [incomplete island]	0.00	0.00	0.00–0.01	−3.85	**<0.001**
Random Effects					
σ^2					2.17
τ_{00} Participant					0.00
ICC					0.00
N Participant					18
Observations					200
Marginal R^2/Conditional R^2					0.815/0.815

she produced one correct noun (e.g., '*building*' in the utterance 'al building' = [to the] building) and one correct verb.

4. Discussion

In this study we investigated LM in multilingual PWA with and without agrammatism, who spoke typologically different languages (English, Hebrew, Russian, and Spanish) which vary in the degree of morphological richness. We examined whether morphological aspects of within-utterance LM in PWA were similar to those reported for LM in neurotypical multilingual people, adopting the 4M model as the foundation for our analyses.

4.1. Frequency and type of language mixing in L1 and L2

Our first research question asked whether the frequency and types of LM in L1 and L2 in multilingual PWA were similar or different to each other. First, we looked at the frequency of LM overall (Fig. 1), to see whether our data sample was comparable to others in the literature and whether there were differences in the frequency of LM in the two languages. We found that on average 2 % of all words produced were mixed when speaking an L1, compared to 7 % of words mixed when speaking an L2; this difference was found to be statistically significant. These results are comparable to those observed in other studies of PWA (e.g., Bihovsky et al., 2023; Goral et al., 2019; Lerman et al., 2019), although they contradict Paplikar's (2016) finding of more LM when speaking L1 than when speaking L2. However, Paplikar (2016) explained her

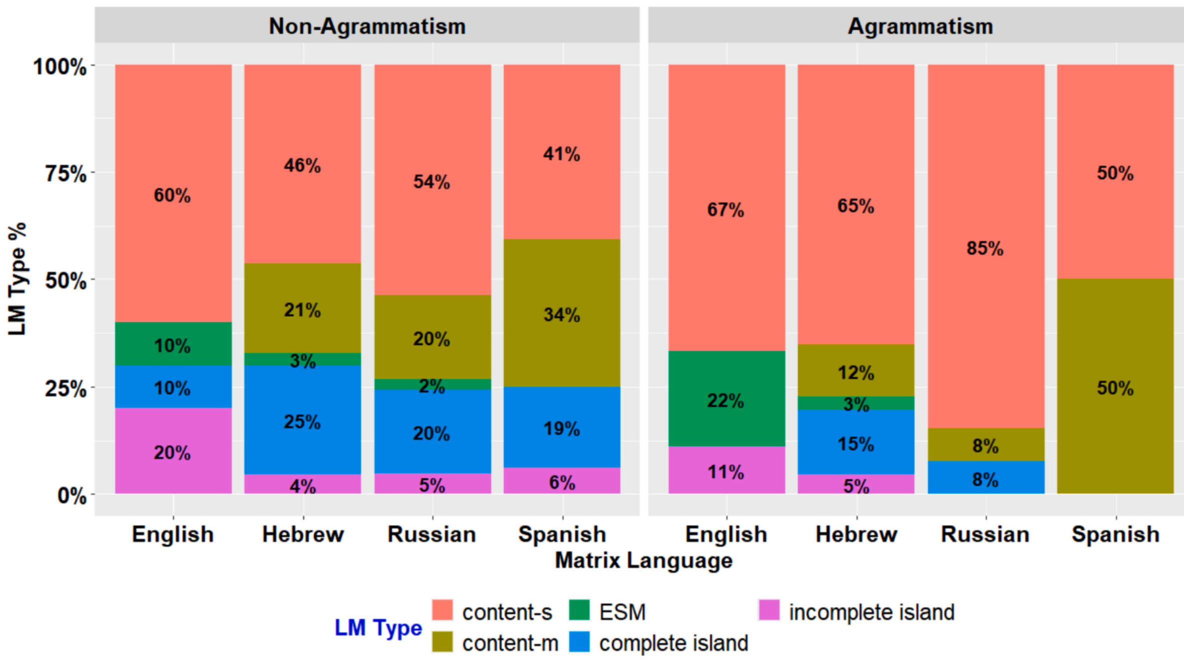

Fig. 4. Distribution of LM type in each matrix language in participants with and without agrammatism.

findings based on the pragmatic norms of the population being investigated, a population that holds L2-English to high esteem and consider it prestigious to mix English into L1-Kannada. Conversely, in our study, order of acquisition and relative language abilities appear to drive frequency patterns. Similar findings were reported in Bihovsky et al. (2023), Goral et al. (2019), and Lerman et al. (2019), for samples that contain additional data beyond the current sample. We note that even though the overall group frequency was low, some participants showed as much as 14 % LM in their narrative production in L1, and 63 % in L2.

When the LM distribution was examined for people with and without agrammatism separately, the findings demonstrated a difference between LM frequency in L1 and L2 only for people with agrammatism. This was contrary to our prediction, but could be accounted for, in part, by the low overall frequency of LM in people without agrammatism in our sample. Furthermore, by looking at the individual data of the 18 participants, we could further refine our understanding of LM frequency patterns. Out of the five participants who reported better post-stroke L2 abilities relative to L1 (RH201, RH207, and SE02) or equal L1 and L2 abilities (RH210 and HE06) we observed that LM was clearly more prevalent when speaking L1 than when speaking L2 for two of them (RH201 and RH210 – notably, both with agrammatism). Furthermore, for two more of these participants (SE02 and HE06), a very small amount of LM was observed in L1 only, in the absence of LM in L2. These results are contrary to the pattern observed at the group level. Thus, we suggest that relative post-morbid language abilities contribute to the frequency of LM across languages, together with order of acquisition. It is important to mention that although for some multilingual people the language of the environment, often the L2, becomes more dominant over time, for many PWA who have given up work before or after their stroke, communication is more limited to the home context and thus L1 may become more dominant again. While a level of analysis that includes specific information about changes in L1 and L2 use in the years leading up to the stroke and/or in the months/years since the stroke is beyond the scope of this paper, it is important to consider these factors in future studies because subtle changes in use and exposure can affect measures of dominance (Lerman et al., 2020; Peñaloza et al., 2020).

Next we examined the LM types within each language (Fig. 2). We found that the patterns of LM in target and non-target languages were similar across L1 and L2, when looking at the group data. However, there were three participants who produced most of the LM within the context of producing full utterances in the non-target language – all three when the target language was L2 (EH03, EH04 and RH204), all participants with severe non-fluent aphasia with agrammatism. Taken together, we conclude that pre-morbid proficiency – which is also related to age/order of acquisition – as well as type and severity of aphasia all impact post-morbid language abilities, and it is these abilities that are underlying differential frequency patterns of LM. Our analyses, therefore, support the hypothesis that LM is used by PWA when they encounter communication difficulties (Goral et al., 2019; Lerman et al., 2019).

4.2. *Within-utterance language mixing patterns*

Our second research question asked whether the 4M model accounts for within-utterance LM patterns in multilingual PWA. We found that, as predicted, the observed LM patterns did indeed follow the predictions of the 4M model, with most LM occurring for content morphemes, and a small percentage of LM occurring for ESM (Fig. 3). We did not observe LM at all for either type of late system morphemes. Thus, our results are similar to those of multilingual people *without* aphasia (Myers-Scotton & Jake, 2017) and also support those of Paplikar (2016) who found that LM occurred most commonly for content morphemes (in her sample, for nouns).

We also found that, at the group level, within-utterance LM produced in an L1 matrix occurred mainly for content morphemes (79 %, with only 1 % ESM), whereas when produced in an L2 matrix both content morphemes (70 %) and ESM (6 %) occurred. The differences between these distributions between L1 and L2 were driven by the participants with agrammatism. While the morpheme-type distribution in L1 and L2 was not found to be statistically significant for the non-agrammatic PWA, the differences found for PWA with agrammatism are similar to previous findings of studies on participants with varying types of aphasia, such as Paplikar's (2016) observation that more LM in bound morphemes occurred when speaking L2-English than L1-Kannada, and Bihovksy et al.'s (2023) finding that mostly content morphemes (nouns) were mixed when speaking L1-Russian, whereas both content morphemes (discourse markers) and system morphemes (function words) were mixed when speaking L2-Hebrew. Here too, we note that

Table 6
Generalized Linear Mixed Model (GLMM) results for the effect of matrix language, LM type and agrammatism.

	pct				
Predictors	*Estimates*	*std. Error*	*CI*	*Statistic*	*p*
(Intercept)	0.00	0.00	0.00–0.01	−6.44	**<0.001**
MatrixLanguage [Hebrew]	22.12	23.39	2.78–175.75	2.93	**0.003**
MatrixLanguage [Russian]	3.79	4.23	0.43–33.79	1.20	0.232
MatrixLanguage [Spanish]	20.98	21.58	2.79–157.58	2.96	**0.003**
Mtype [content-m]	0.00	0.00	0.00–0.03	−5.22	**<0.001**
Mtype [ESM]	0.02	0.02	0.00–0.14	−3.75	**<0.001**
Mtype [complete island]	0.01	0.01	0.00–0.06	−4.61	**<0.001**
Mtype [incomplete island]	0.02	0.03	0.00–0.20	−3.37	**0.001**
Agrammatism [Agrammatism]	0.02	0.02	0.00–0.20	−3.25	**0.001**
(MatrixLanguage [Hebrew] × Mtype [content-m]) × Agrammatism [Agrammatism]	0.01	0.01	0.00–0.34	−2.52	**0.012**
(MatrixLanguage [Russian] × Mtype [content-m]) × Agrammatism [Agrammatism]	0.00	0.00	0.00–0.14	−2.88	**0.004**
(MatrixLanguage [Spanish] × Mtype [content-m]) × Agrammatism [Agrammatism]	0.02	0.03	0.00–0.65	−2.18	**0.029**
(MatrixLanguage [Russian] × Mtype [ESM]) × Agrammatism [Agrammatism]	0.00	0.00	0.00–0.00	−5.12	**<0.001**
(MatrixLanguage [Russian] × Mtype [incomplete island]) × Agrammatism[Agrammatism]	0.00	0.00	0.00–0.00	−4.58	**<0.001**
(MatrixLanguage [Spanish] × Mtype [incomplete island]) × Agrammatism[Agrammatism]	0.01	0.02	0.00–0.48	−2.36	**0.018**
Random Effects					
σ^2					1.97
$\tau_{00\ \mathrm{Participant}}$					7.65
ICC					0.80
$N_{\mathrm{Participant}}$					18
Observations					720
Marginal R^2/Conditional R^2					0.558/0.910

individual differences were observed in our findings, with only five out of the 18 participants producing any LM for ESM (four in L2: SE07, RH201, RH207 and HE01 – all diagnosed with agrammatism, and one in L1: RH213 – not diagnosed with agrammatism). Further research could look at characteristics of PWA who produce a salient amount of LM for ESM, particularly when L1 is the matrix language, to try to identify features of their language background, use, and/or stroke and aphasia that may be important in determining appropriate intervention strategies for this less common type of LM.

Regarding islands, we observed that 22 % of LM instances were either complete islands (16.5 %) or incomplete islands (5.5 %), with 43 % of all islands being comments about the narrative. The percentages fall within the range of previously reported rates of islands and, as found in previous studies, demonstrate cross-language variability (Jake et al., 2002). Looking at the individual data, we observed variation across the participants – seven out of 18 participants produced no islands at all in one or both of their languages. We also observed that one participant in particular was driving the incomplete-island data – EH04, a participant with agrammatism, and severe non-fluent aphasia in his L2, who had a tendency to either give up in the middle of utterances (resulting in incomplete islands) or just switch back to his L1 (resulting in many utterances in the non-target language, see Fig. 2). Additionally, three other participants (HE01, SE02, and SE07) produced a relatively high number of incomplete islands compared to the other participants, although they had no salient similarities that might explain this finding. While these results are certainly interesting, we need more data on neurotypical multilingual people and more breakdown of islands (both complete and incomplete) in order to understand the data for multilingual PWA. This will be an important direction for future research.

4.3. Language-specific characteristics of within-utterance language mixing

Our third research question asked whether language-specific characteristics affect within-utterance LM. Aligning with our prediction, we found an effect of language. More specifically, we found that in our sample of PWA, the matrix language was driving the morpheme-related type of LM that occurred, and not the embedded language, a finding comparable to that in neurotypical multilingual people (Myers-Scotton & Jake, 2000; 2009). We were able to identify this pattern because we examined LM across different language pairs (Fig. 4). In English, the language with the least morphological richness, we observed the most within-utterance LM for ESM relative to the other languages, but no LM for multimorphemic content words, whereas in the other, morphologically richer languages, we observed LM for multimorphemic content words in 14 out of our 18 participants, consistent with our prediction that multimorphemic content words will be more likely to be mixed in these other languages than in English.

The language pairs in this study are incongruent on a number of grammatical features (i.e., phi-features in English and Hebrew; word order parameter in adjectival phrases, cf. English vs. Hebrew), and this was reflected in our results for islands: differences between the rate of complete islands were found between English vs. Spanish, English vs. Hebrew, Hebrew vs. Russian, and Hebrew vs. Spanish. Notably, only 3 % of LM was within complete islands when English was the matrix language. Similarly, differences were found between incomplete islands in English vs. Hebrew and in Hebrew vs. Russian. Thus, our findings support the hypothesis that differences across languages may be related to the morphological richness of the matrix language and congruency of the languages involved in LM (Alexiadou & Lohndal, 2018; Jake et al., 2002; Myers-Scotton, 1993; Paplikar, 2016). For example, we predicted that the morphological structure of English would allow for mixing of single words into the matrix language more easily than for other languages, and thus this may be the reason that embedding entire islands was observed to be less common in English than in the other languages.

Regarding agrammatism, differences between LM frequency in L1 vs in L2 and the distributions of morpheme types appeared to be driven by the PWA with agrammatism, because they were not observed for the PWA without agrammatism. Furthermore, we observed that content-s nouns and verbs embedded into utterances were always morphosyntactically accurate in context, in PWA without agrammatism. For PWA with agrammatism, all content-s nouns produced were morphosyntactically correct across all languages. For content-s verbs, PWA with agrammatism produced verbs in a Russian, Hebrew, or English matrix only. For Russian and English, all were morphosyntactically correct in context, whereas for Hebrew all five were either zero marked verbs or an infinitive form – with three out of five morphosyntactically incorrect in context. Thus, we conclude from our data that, as expected (e.g., Faroqi-Shah, 2023), PWA with agrammatism demonstrated partial difficulty producing verbs correctly when they were mixed into another language. PWA without agrammatism did not produce any incorrect nouns or verbs in context. Future studies could compare error analyses of nouns and verbs in PWA with agrammatism in both the target language and the non-target language, to provide a more comprehensive understanding of what is driving these effects.

4.4. Limitations, future directions, and clinical implications

This study applied the 4M model for the first time to LM patterns of multilingual PWA, providing considerable contribution to the field. However, several limitations should be acknowledged, particularly the small participant pool from each lab, relative to the large variability across aphasia severity, type, and language backgrounds. We also note that the average frequency of LM in our sample was not high (e.g., 2 % and 7 % for L1 and L2 respectively, compared to 28 % and 8 % in Paplikar, 2016), although we did observe variability across participants. Previous studies have reported higher frequency of mixing among PWA compared to neurologically healthy people (Muñoz et al., 1999; Paplikar, 2016), but they too observed variability among their participants. We did not directly compare our data to a sample of people without aphasia.

Furthermore, due to the retrospective nature of this study, participants were asked to produce narratives under slightly different conditions across the three labs, and language background information was not collected from all participants in the same manner. Our data also did not differentiate the potential effects of immigration and language immersion on the observed patterns of LM. However, in spite of these differences across participants, our results showed some clear patterns across L1 and L2, when taking into account language dominance as well, and across different language pairs, generally replicating published findings for both PWA and neurotypical adults to date.

Future studies could further examine patterns of LM in PWA in general as well as their implications for assessment and intervention. Our findings suggest that LM in PWA is typical, rather than pathological. This perspective of LM in PWA is consistent with the view that PWA could be encouraged to use their full linguistic repertoire to maximize their communication potential (Goral et al., 2024). The typicality of using one's complete linguistic repertoire is a key component of *translanguaging* (e.g., Khamis & Bortz, 2023). A translanguaging approach to assessment emphasizes the utility of acknowledging the complete linguistic repertoire of multilingual people over the arbitrary nature of externally tagging linguistic elements by language (Li Wei & García, 2022). Whereas there is still much debate regarding the concept of translanguaging both in pedagogical contexts and applied linguistics, with the concept effectively referring to a vast number of different scenarios of language processing and use (Treffers-Daller, 2024), abolishing the necessity for defining a single target language during assessment and intervention in aphasia is a promising avenue for future clinical practices.

Patterns of LM are naturally affected by linguistic and cultural norms within communities of speakers and such factors should be considered. That is, LM observed by PWA may reflect cultural and linguistic norms within their community but could also deviate from such norms (MacSwan, 2022; Paplikar, 2016). Furthermore, LM that contravenes the rules of the 4M, such as a patient who mixes more system morphemes than content morphemes, particularly late system morphemes relative to early system morphemes (not observed in our data set), could be considered atypical LM. In these cases, clinicians should consider whether intervention to reduce LM is warranted. Further research on LM in PWA who speak varying languages, and particularly in any participants with (less-common) atypical LM patterns e.g., (Bihovsky, Holzapfel, Bondi, Ben-Shachar, & Meir, 2025), should be encouraged in order to move the field forwards and provide focused evidence to support clinical practice.

5. Conclusion

Our study showed that PWA produced narratives mostly in the target language, with a clear matrix language in the event of mixing, in most cases. As a rule, PWA mixed their languages in a typical manner, a finding which is strengthened by the uniqueness of our study looking across several language pairs with varying morphological richness. This typicality was observed for LM distribution across L1 and L2, taking language dominance into account, as well as specific analyses of morpheme mixing based on the 4M model. These findings have clinical significance, such that in most cases, LM can be considered typical (although perhaps more frequent in some PWA than in neurotypical multilingual people). Thus, if and when LM occurs in the aphasia clinic, it should generally be encouraged as a communication aid, rather than discouraged. Future studies further examining patterns of LM in multilingual PWA will help determine typical and atypical multiple language use.

CRediT authorship contribution statement

Aviva Lerman: Writing – review & editing, Project administration, Investigation, Conceptualization, Writing – original draft, Methodology, Formal analysis. **Alina Bihovsky:** Methodology, Formal analysis, Writing – review & editing, Investigation, Conceptualization. **Natalia Meir:** Writing – review & editing, Methodology, Conceptualization, Writing – original draft, Formal analysis. **Michal Ben-Shachar:** Methodology, Writing – review & editing, Conceptualization. **Mira Goral:** Writing – original draft, Investigation, Conceptualization, Writing – review & editing, Methodology, Formal analysis.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://osf.io/jw8fy/.

References

Abuom, T. O., & Bastiaanse, R. (2012). Characteristics of Swahili–English bilingual agrammatic spontaneous speech and the consequences for understanding agrammatic aphasia. *Journal of Neurolinguistics, 25*(4), 276–293.
Albert, M. L., & Obler, L. K. (1978). *The Bilingual Brain: Neuropsychological and Neurolinguistic Aspects of Bilingualism*. Academic Press.
Alexiadou, A., & Lohndal, T. (2018). Units of language mixing: A cross-linguistic perspective. *Frontiers in Psychology, 9*, 1–15.
Bastiaanse, R., & Thompson, C. K. (Eds.). (2012). *Perspectives on agrammatism*. Psychology Press.
Bentahila, A., & Davies, E. E. (1992). Code-switching and language dominance. In R. J. In (Ed.), *Harris Cognitive Processing in Bilinguals* (pp. 443–458). North-Holland.
Berthier, M. L. (2005). Poststroke Aphasia. *Drugs & Aging, 22*(2), 163–182. https://doi.org/10.2165/00002512-200522020-00006
Bihovsky, A., Ben-Shachar, M., & Meir, N. (2023). Language abilities, not cognitive control, predict language mixing behavior in bilingual speakers with aphasia. *Journal of Communication Disorders, 105*, Article 106367.

Bihovsky, A., Holzapfel, S., Bondi, M., Ben-Shachar, M., & Meir, N. (2025). Unusual patterns of codeswitching in an unbalanced bilingual person with aphasia: effects of language and executive functions impairment. *Aphasiology*, *39*(7), 1006-1028..

Fabbro, F., Skrap, M., & Aglioti, S. (2000). Pathological switching between languages after frontal lesions in a bilingual patient. *Journal of Neurology, Neurosurgery, and Psychiatry, 68*, 650–652.

Faroqi-Shah, Y. (2023). A reconceptualization of sentence production in post-stroke agrammatic aphasia: The Synergistic Processing Bottleneck model. *Frontiers in Language Sciences, 2*, Article 1118739. https://doi.org/10.3389/flang.2023.1118739

Fyndanis, V., & Lehtonen, M. (2021). Cognitive control and pathological language mixing in aphasia. In R. Blackwood, & U. Røyneland (Eds.), *Multilingualism Across the Lifespan* (pp. 209–230). London: Routledge.

Goodglass, H., & Menn, L. (1985). *Is agrammatism a unitary phenomenon? In Agrammatism* (pp. 1–26). Academic Press.

Goral, M., Antolovic, K., Hejazi, Z., & Schulz, F. M. (2024). Using a translanguaging framework to examine language production in a trilingual person with aphasia. *Clinical Linguistics & Phonetics*, 1–20.

Goral, M., Norvik, M., & Jensen, B. U. (2019). Variation in code-switching in multilingual aphasia. *Journal of Clinical Linguistics and Phonetics, 33*, 915–929.

Green, D. W., & Abutalebi, J. (2008). Understanding the link between bilingual aphasia and language control. *Journal of Neurolinguistics, 21*(6), 558–576.

Grosjean, F. (1985). Polyglot aphasics and language mixing: A comment on Perecman (1984). *Brain and Language, 26*, 349–355.

Grosjean, F. (2001). The bilingual's language modes. In J. Nicol (Ed.), *One Mind, Two Languages: Bilingual Language Processing* (pp. 1–22). Blackwell.

Jake, J. L., Myers-Scotton, C., & Gross, S. (2002). Making a minimalist approach to codeswitching work: Adding the Matrix Language. *Bilingualism: Language and Cognition, 5*(1), 69–91.

Khamis, R., & Bortz, M. (2023). Translanguaging: A theoretical shift necessary for culturally and linguistically responsive practice in Speech Language Hearing Sciences. In M. Goral, & A. Lerman (Eds.), *Advances in the Neurolinguistic Study of Multilingual and Monolingual Adults* (pp. 36–50). Routledge.

Kiran, S., & Gray, T. (2018). Understanding the nature of bilingual aphasia: Diagnosis, assessment and rehabilitation. In D. Miller, F. Bayram, J. Rothman, & L. Serratrice (Eds.), *Bilingual cognition and language: The state of the science across its subfields* (pp. 371–400). John Benjamins.

Lerman, A. (2020). *Direct and indirect treatment effects in multilingual people with aphasia. Doctoral dissertation.* City University of New York.

Lerman, A., Pazuelo, L., Kizner, L., Borodkin, K., & Goral, M. (2019). Language mixing patterns in a bilingual with non-fluent aphasia. *Aphasiology, 33*(9), 1137–1153.

Lerman, A., Goral, M., & Obler, L. K. (2020). The complex relationship between pre-stroke and post-stroke language abilities in multilingual aphasia. *Aphasiology, 34* (11), 1319–1340. https://doi.org/10.1080/02687038.2019.1673303

Wei, L.i., & García, O. (2022). Not a first language but one repertoire: Translanguaging as a decolonizing project. *RELC Journal, 53*(2), 313–324. https://doi-org.jerome.stjohns.edu/10.1177/00336882221092841.

MacSwan, J. (2022). *Multilingual Perspectives on Translanguaging.* Multilingual Matters.

Menn, L., Obler, L. K., & Miceli, G. (1990). *Agrammatic aphasia: A cross-language narrative sourcebook* (Vol. 2). John Benjamins Publishing.

Muñoz, M. L., Marquardt, T. P., & Copeland, G. (1999). A comparison of the codeswitching patterns of aphasic and neurologically normal bilingual speakers of English and Spanish. *Brain and Language, 66*(2), 249–274. https://doi.org/10.1006/brln.1998.2021

Muysken, P. (2000). *Bilingual speech: A typology of code-mixing.* Cambridge University Press.

Myers-Scotton, C. (1993). *Duelling languages: Grammatical structure in code switching.* Oxford University Press.

Myers-Scotton, C. (2002). *Contact Linguistics: Bilingual Encounters and Grammatical Outcomes*. Oxford University Press.

Myers-Scotton, C., & Jake, J. (2000). Four types of morpheme: Evidence from aphasia, code switching, and second-language acquisition. *Linguistics, 38*(6), 1053–1100. https://doi.org/10.1515/ling.2000.021

Myers-Scotton, C. M., & Jake, J. L. (2017). Revisiting the 4-M model: Codeswitching and morpheme election at the abstract level. *International Journal of Bilingualism, 21*(3), 340–366.

Nguyen, L. (2024). Rethinking the matrix language: Vietnamese–English code-switching in Canberra. *International Journal of Bilingualism, 1–27*. https://doi.org/10.1177/13670069241254454

Papathanasiou, I., & Coppens, P. (2021). *Aphasia and Related Neurogenic Communication Disorders* (3rd ed.). Jones & Bartlett Learning.

Paplikar, A. (2016). *Language-mixing in discourse in bilingual individuals with non-fluent aphasia*. CUNY Graduate Center. Unpublished doctoral dissertation.

Paradis, M. (2004). *A neurolinguistic theory of bilingualism*. John Benjamins Publishing Company.

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

Perecman, E. (1984). Spontaneous translation and language mixing in a polyglot aphasic. *Brain and Language, 23*(1), 43–63.

R Core Team. (2023). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing.

Raymer, A. M., Beeson, P., Holland, A., Kendall, D., Maher, L. M., Martin, N., Murray, L., Rose, M., Thompson, C. K., Turkstra, L., Altmann, L., Boyle, M., Conway, T., Hula, W., Kearns, K., Rapp, B., Simmons-Mackie, N., & Gonzalez Rothi, L. J. (2008). Translational research in aphasia: From neuroscience to neurorehabilitation. *Journal of Speech, Language, and Hearing Research: JSLHR, 51*(1), 259–275. https://doi.org/10.1044/1092-4388(2008/020)

Treffers-Daller, J. (2024). Translanguaging: What is it besides smoke and mirrors? *Linguistic Approaches to Bilingualism, 14*(5), 1–26. https://doi.org/10.1075/lab

Walters, J. (2005). *Bilingualism: The sociopragmatic-psycholinguistic interface.* Lawrence Erlbaum.