INTRODUCTION

BACKGROUND:

Broadly speaking, frontotemporal dementia (FTD) is the result of neuronal cell damage to the frontal and temporal regions of the brain, resulting in the shrinkage of those lobes [1]. The frontal region of the brain is critical for expressive language, voluntary movement, and maintaining more complex cognitive skills such as the capacity to plan and organize one's goals [2]. The temporal region is mainly responsible for processing what we hear and encoding memory, hence playing an important role in understanding language and retaining verbal and non-verbal information [2]. Individuals diagnosed with FTD experience devastating symptoms, such as inability to control their behavior, lack of emotional regulation, trouble with communicating, and incapacity to make decisions [1]. There are three distinct types of FTD that present unique clinical symptoms: behavioral variant frontotemporal dementia (bvFTD), semantic variant primary progressive aphasia (svPPA), and nonfluent variant primary progressive aphasia (nfvPPA). As implied in the name, individuals with bvFTD typically tend to have difficulty controlling their behavior and may exhibit signs of apathy, lack of empathy, impulsivity, etc. [3]. Primary progressive aphasia (PPA) groups individuals that experience damage to the regions of the brain responsible for speech and language; those with svPPA may have trouble with understanding the meaning of words and being able to find the right words to name people/objects while those with nfvPPA have difficulty with their speech and clarity when speaking [3]. There is an additional form of PPA, called logopenic variant primary progressive aphasia (lvPPA), which does not fall under the FTD umbrella as autopsies have indicated that those with FTD exhibit pathological signs more consistent with Alzheimer's diseases instead [3]. This form of PPA impacts one's ability to think of the words they are trying to express and as a result, being to speak slower [3].

Each of these forms of FTD vary in terms of the general symptoms produced. However, individuals across the board experience a decline in their ability to think and carry out functions [3]. Note that the severity of disease progression and symptoms varies by person.

Cognitive reserve refers to one's ability to preserve function in the face of cognitive decline, such as in patients with FTD [4]. Individuals that speak more than one language may show greater brain volume, brain connectivity, and enhanced executive functioning, which could improve one's cognitive reserve later in life [5]. Prior research studies report conflicting evidence regarding the role of bilingualism as a protective factor in cognitive reserve. While it is known that bvFTD and PPA disease variants impact distinct underlying neural networks, the effect of being bilingual on these networks is yet to be determined.

RESEARCH QUESTION/GOAL:

In this study, we investigated whether the impact of bilingualism on cognitive performance varied in language-related PPAs versus non-language-related bvFTD. Due to the availability of longitudinal data, this question was evaluated both cross-sectionally and over time.

METHODS

DESCRIBING THE DATA:

Participants

Participants for this study were recruited through efforts of a longitudinal study from the UCSF Memory and Aging Center (MAC). Demographic variables, such as sex, race, education, and age were collected. A comprehensive chart review was performed to identify bilingual speakers for inclusion in this study Patients were classified as bilingual speakers if their chart

indicated that they could communicate in two or more languages in everyday interaction with other speakers. Cognitive performance was assessed through a series of neuropsychological tasks that were administered to patients at the time of data collection.

Neuropsychological Measures

There was a total of twenty-six neuropsychological tasks that were grouped into five domains: general cognition, language, visuospatial, memory, and frontal executive function. General cognition tasks consisted of the Mini Mental State Exam (MMSE) and the Gordon Diagnostic System (GDS). Language measures were Sentence Repetition, Verbal Agility, Wide Range Achievement Test (WRAT4), Irregular Word Reading, Sentence Comprehension, Peabody Picture Vocabulary Test (PPVT), Animal Fluency, and the Boston Naming Test. The three tasks that assessed visuospatial function were Rey-Osterrieth Complex Figure (Rey Copy), Calculations, and Visual Object Space and Perception (VOSP) battery. The frontal executive tasks were trail lines per second, Design Fluency, Digits Forward, Digits Backward, D words, Abstraction, Stroop Color Naming, and Stroop Inhibition. The final two tasks, the California Verbal Learning Test (CVLT) and Rey Recall fall under the memory domain.

STATISTICAL ANALYSIS:

All statistical analyses were carried out using R Version 4.3.1.

Baseline demographics:

Demographic variables (age, sex, education, and race) and the neuropsychological measures were compared by monolingual and bilingual speakers within each disorder at baseline. Continuous variables (age, education, neuropsychological measures) were compared using Welch's t-Test due to the unequal variance of monolingual and bilingual speakers in each disorder. Categorical variables (race and sex) were evaluated using Fisher's Exact Tests.

Baseline linear regression:

Multivariable linear regression models were used to evaluate the difference in cognitive performance between bilingual and monolingual speakers. Considering that the outcome variables, the neuropsychological tasks, were continuous in nature and there would be a mix of categorical and continuous predictor variables, multivariable linear regression models were the best choice to answer the research objective. To evaluate the patient cohort at baseline, the data was filtered to only contain patients evaluated at their first timepoint. Initially, the data was then further split by diagnosis so that each of the PPAs (lvPPA, svPPA, nfvPPA) and bvFTD groups had a separate baseline dataset. Age, education, lingual status, and gender were included as covariates in the linear regression model. Age and education are continuous variables, represented by years. Gender (Male/Female) and lingual status (Bilingual/Monolingual) were included in the models as categorical variables. Age, education, and gender were included in the models as covariates due to being established as significant predictors of cognitive performance in individuals with FTD and PPA in prior literature.

After running these four separate models, the lingual status covariate was corrected for multiple comparisons by using the False Discovery Rate (FDR) method. Upon further examination of the sample sizes when separating the regression models by variant, it was apparent that all the data needed to be grouped together to support analyses by diagnosis. Hence, a separate variable classifying diagnosis as either language-related FTD (nfvPPA, lfvPPA, nfvPPA) vs non-language related FTD (bvFTD) was created to include in the model. An interaction term between lingual status and diagnosis, along with the other predictors mentioned above, was also added to the model to capture instances in where the combined effect of both of those covariates significantly impacted the neuropsychological tasks. A separate linear regression model with diagnosis represented as a categorical variable with four levels (lvPPA,

svPPA, nfvPPA, and bvFTD), an interaction term between diagnosis and lingual status, and the same additional predictors was used to evaluate how the different PPA groups performed in reference to the bvFTD group.

To visualize baseline results, four separate forest plots were created for each diagnosis to interpret the coefficient for the interaction term. To check model assumptions, a Normal Quantile-Quantile (QQ) plot was used to check for the normality of residuals and a Fitted vs. Residuals plot assessed the linearity and homoscedasticity assumptions.

Longitudinal linear regression:

To determine which observations to include for longitudinal analyses, we looked at the total count of patients included in each timepoint. Timepoint 2 had roughly half the initial patients at baseline, and the numbers dwindled by half for each timepoint after. In order to run a longitudinal model with the highest sample size possible to validate results, patients with both a timepoint 1 and timepoint 2 visit were included for this analysis.

Like the baseline regression, the data was first split by diagnosis and four separate models were run with lingual status, gender, age, and education as predictors. The time between data collection was also included as a predictor and was represented as a unit of years. Given the longitudinal nature of the data, a linear mixed effects model was utilized with the random effect accounting for distinct slopes for each patient.

Additionally, a multivariable linear regression model, similar to baseline analyses, was used to evaluate the longitudinal data. The outcome for this model was the difference in neuropsychological measures from timepoint 2 to timepoint 1. However, due to the complexity in interpreting the results of the models based on the difference in cognitive tasks and the inability to account for a random effect of patients, we went forward with the linear mixed effect

model. The final version of the linear mixed effect model had age, gender, education, diagnosis, and lingual status as fixed effects. The random effect in the model were individual intercepts for each patient, modeling each patient's unique baseline score for the neuropsychological measures. One linear mixed effect model contained a binary variable of diagnosis (bvFTD vs all PPAs) while the other model contained a categorical variable of diagnosis (bvFTD, svPPA, lvPPA, and nfvPPA), both with interaction terms between diagnosis and lingual status.

Confidence intervals and p-values were generated with Kenward-Roger's degrees of freedom approximation for the linear mixed effect models. To help interpret the results visually, marginal effects were plotted over time by diagnosis for monolingual and bilingual patients over time. Model assumptions were examined through Normal Q-Q plots and Fitted vs Residuals graph, just like they were for baseline results.

RESULTS

Demographic characteristics of patients:

This study analyzed a cohort of 316 patients at baseline (224 monolingual and 92 bilingual speakers). The cohort was 53% female (n=167). The average age of the cohort was 64.7 years (SD 8.3). Patients in this cohort completed an average of 16.3 years of education (SD 2.7). These patients were categorized into four disease variants: Logopenic Variant Primary Progressive Aphasia (n=72), Semantic Variant Primary Progressive Aphasia (n=85), Nonfluent Variant Primary Progressive Aphasia (n=79), and Behavioral Variant Frontotemporal Dementia (n=80). Table 1 shows the breakdown of demographic characteristics and neuropsychological measures between monolingual and bilingual speakers by disease variant. Across each disease variant, the monolingual and bilingual speakers did not differ in terms of age, education, sex, and Clinical Dementia Rating (CDR), which is a measure of disease severity. The bvFTD, svPPA,

and lvPPA, groups did significantly differ in the distribution of race between monolinguals and bilinguals. In terms of the neuropsychological measures, the bvFTD group differed in scores for the GDS (Monolingual Mean = 8.4 ± 7.0, Bilingual Mean = 4.6 ± 4.5, p = 0.01), Rey Recall (Monolingual Mean = 7.8 ± 4.6, Bilingual Mean = 4.8 ± 4.2, p = 0.01), BNT (Monolingual Mean = 12.0 ± 3.4, Bilingual Mean = 9.1 ± 3.9, p = 0.01), and VOSP (Monolingual Mean = 8.5 ± 2.1, Bilingual Mean = 6.8 ± 2.8, p = 0.01) tasks by bilingual status. The lvPPA group differed in the Verbal Agility (Monolingual Mean = 3.4 ± 2.0, Bilingual Mean = 5.0 ± 1.2, p < 0.001) task between monolingual and bilingual individuals. The monolinguals and bilinguals in the svPPA and nfvPPA group were not significantly different in any of the neuropsychological measures.

Baseline linear regression results:

After adjusting for education, age, and gender, bilingual speakers in the bvFTD group scored significantly lower on certain tasks relating to memory **(*Rey Recall:*** p = 0.009) **,** language (***Sentence Repetition***: p = <0.001; ***Verbal Agility:*** p = 0.021*;* ***Animal Fluency:*** p = 0.010; ***BNT:*** p = 0.026), visuospatial (***VOSP****:* p = 0.004) and frontal executive tasks (***Trails (lines/second)***: p = 0.015; ***Design Fluency:*** p = 0.043) in reference to their monolingual, PPA counterparts. Figure 1A shows the forest plot for the interaction term in the baseline linear regression model where diagnosis was treated as a binary variable (bvFTD vs svPPA, lvPPA, and nfvPPA).

When looking at the model where diagnosis was treated as a categorical variable, we can see the results for each of the PPAs in reference to the bvFTD group. Figure 1B-D models the results of the interaction between lingual status and disease variant at baseline for these groups. Adjusting for the other predictors in the model, svPPA, nfvPPA, and lvPPA bilingual speakers showed an increase in scores on the neuropsychological measures. Specifically, the svPPA

bilinguals scored better on memory **(*Rey Recall:*** p = 0.023), language (***Sentence Repetition:*** p = 0.015; ***Animal Fluency:*** p = 0.015; ***BNT:*** p = 0.009) ***v***isuospatial ***(VOSP:*** p = 0.006)**,** and frontal executive tasks (***Trails (lines/second):*** p = 0.007). The nvfPPA patients who were bilingual scored better on memory (***Rey Recall:*** p = 0.032) and language tasks ***(Sentence Repetition:*** p = 0.019; ***BNT:*** p = 0.041). Bilingual speakers with lvPPA scored higher on tasks in the memory (***CVLT:*** p = 0.039 ; ***Rey Recall:*** p = 0.022) **,** language (***Sentence Repetition:*** p = <0.001; ***Verbal Agility:*** p = <0.001; ***Irregular Word Reading:*** p = 0.049; ***Sentence Comprehension:*** 0.034; ***BNT:*** 0.010), and visuospatial realm ***(VOSP:*** p = 0.031).

Longitudinal regression results:

Over time, we see similar patterns observed cross-sectionally. Table 3A and 3B showcase the regression outputs for the linear mixed effects models. Bilingual speakers with bvFTD scored lower on tasks across all five domains after adjusting for the predictors in the model: general ***(MMSE:*** p = 0.034)**,** memory ***(Rey Recall:*** p = <0.001; ***CVLT:*** p = 0.016)**,** frontal executive ***(Trails (lines/second):*** p = 0.003; ***Design Fluency:*** 0.009; ***Digits Forward:*** 0.027; ***D Words:*** 0.048; ***Stroop Color Naming:*** p = 0.015; ***Stroop Inhibition:*** p = 0.016)**,** language ***(Sentence Repetition:*** p = <0.001***; Irregular Word Reading:*** p = 0.013; ***Sentence Comprehension:*** p = 0.016 ; ***Animal Fluency:*** p = 0.038; ***BNT:*** p = 0.026)**,** and visuospatial ***(VOSP:*** p = 0.01)**.** The svPPA patients that are bilingual scored better on memory ***(Rey Recall:*** p = <0.001)**,** visuospatial ***(VOSP:*** p = 0.006)**,** language ***(Sentence Repetition:*** p = 0.028; ***Irregular Word Reading:*** p = 0.001; ***Animal Fluency:*** p = 0.002; ***BNT:*** p = <0.001; ***PPVT:*** p = 0.022), and frontal executive (***Trails (Lines/Second):*** p = 0.001; ***Design Fluency:*** p = 0.001; ***D words:*** p = 0.005; ***Stroop Color Naming:*** p = 0.017; ***Stroop Inhibition:*** p = 0.016) tasks over time. Bilingual speakers with nfvPPA scored better on memory (***Rey Recall:*** p = 0.005) and language (***Sentence Repetition:*** p = <0.001; ***Sentence Comprehension:*** p = 0.016; ***BNT:*** p = 0.044; ***Verbal Agility:*** p = 0.014)

neuropsychological measures. Those that were bilingual and had lvPPA scored higher on tasks relating to memory (***Rey Recall:*** p = 0.012), language (***Sentence Repetition:*** p = 0.012; ***Sentence Comprehension:*** p = 0.018), and frontal executive functioning (***Digits Forward:*** p = 0.012). Figure 2A and 2B show the marginal effects of the Rey Recall task plotted over time for the binary variant and categorical variant linear mixed effects model respectively. Visually, it is evident that the bvFTD bilingual speakers consistently score lower over time on Rey Recall compared to the monolingual speakers with PPA.

DISCUSSION

These findings suggest bilingualism's effects on cognitive performance differ by the FTD variants. Both cross-sectionally and over time, bvFTD bilingual speakers significantly scored lower on tasks across the cognitive domains while PPA bilingual speakers significantly scored higher. This highlights the complexity of bilingualism's role in cognitive reserve and stresses the need to compare within the variants.

Figure 3A and 3B shows a representation of the Normal Q-Q plot and Fitted vs Residuals graph for the model at baseline and over time for the Rey Recall measure. This pattern is similar to Normal Q-Q plots and Fitted vs Residuals graphs across the neuropsychological measures. Overall, the linear models fulfill the requirement of normality as the majority of the data points fall close to the line. We do see a series of points deviate from the line especially at the tail end regions of the Normal Q-Q plot, indicative of outliers. Yet, this was expected given the nature of our dataset. The dataset contains scores of neuropsychological measures strictly from individuals with neurodegenerative diseases, so the data will be skewed to a degree. The Fitted vs Residuals plot shows that the residuals are randomly scattered around zero. The residuals roughly form a

horizontal band around the zero line, deviating from any distinct pattern. This tells us that our models are not violating homoscedasticity and linearity assumptions.

A major limitation of this study was the missingness of data. The "Mono/Bi" column in Table 1 shows that not all patients in this dataset completed every neuropsychological task at baseline. This issue persists over time and reduces the sample size even further. In fact, the lvPPA group in particular only had two bilingual patients that were included in the longitudinal analyses. Multiple imputation methods were considered initially but ultimately were not used. The rationale behind this was that each patient's pathological progression is highly variable and therefore using methods such as mean imputation would not be able to realistically capture the true scoring for the neuropsychological measures. The sample size issue was the main reason that all the datapoints were combined instead of running separate linear regression models on each of the disorders. To further validate the results of this study, it is imperative that future directions focus on increasing the size of the patient cohort and ensuring a thorough evaluation, especially for the bilingual speakers, in order to get a representative sample.

REFERENCES

1) *What are frontotemporal disorders? Causes, symptoms, and treatment*. (2021, July 30). National Institute on Aging. https://www.nia.nih.gov/health/frontotemporal-disorders/what-are-frontotemporal-disorders-causes-symptoms-and-treatment#FTDcenters
2) B*rain Map Frontal Lobes*. (2022, July 12). Queensland Government. https://www.health.qld.gov.au/abios/asp/bfrontal#:~:text=The%20frontal%20lobes%20are%20important,order%20to%20achieve%20a%20goal.
3) *Frontotemporal dementia*. (n.d.). Memory and Aging Center. https://memory.ucsf.edu/dementia/ftd
4) Joshi, M. S., & Galvin, J. E. (2022). Cognitive Resilience in Brain Health and Dementia Research. *Journal of Alzheimer's disease : JAD*, *90*(2), 461–473. https://doi.org/10.3233/JAD-220755.
5) de Leon, J., Grasso, S. M., Allen, I. E., Escueta, D. P., Vega, Y., Eshghavi, M., … Henry, M. L. (2024). Examining the relation between bilingualism and age of symptom onset in

frontotemporal dementia. *Bilingualism: Language and Cognition*, *27*(2), 274–286. doi:10.1017/S1366728923000226