



Published in final edited form as:

*Neuropsychology*. 2023 July ; 37(5): 595–606. doi:10.1037/neu0000893.

## Which Language is More Affected in Bilinguals with Alzheimer's Disease?:

### Diagnostic sensitivity of the Multilingual Naming Test

Tamar H. Gollan<sup>1</sup>, Alena Stasenko<sup>1</sup>, David P. Salmon<sup>2</sup>

<sup>1</sup>Department of Psychiatry, University of California, San Diego

<sup>2</sup>Department of Neurosciences, University of California, San Diego

### Abstract

**Objective:** This study examined the joint consequences of bilingualism and Alzheimer's disease (AD) for picture naming ability to determine which language is more affected by AD and what scoring methods best distinguish patients from controls.

**Method:** Sixty-five Spanish-English bilinguals including 26 with dementia and 39 controls with equivalent age, education, and bilingual proficiency level, were tested on the Multilingual Naming Test (MINT; Gollan et al., 2012).

**Results:** Bilinguals with AD named fewer pictures than controls, and overall AD seemed to affect both languages about equally, but exploratory analyses suggested that this varied with item difficulty. In the dominant language difficult items exhibited a larger effect of AD than easy items (which were at ceiling for both patients and controls), whereas in the nondominant language items of all difficulty levels were about equally affected by AD. An "either-language" scoring procedure (that counted items as correct if produced only in one of the two languages) increased naming scores especially in balanced bilinguals, and to an equal extent in patients and controls. Receiver Operating Characteristic (ROC) analyses revealed that dominant-language and either-language naming scores classified bilinguals as patients versus controls equally well and adding nondominant-language scores did not improve diagnostic sensitivity.

**Conclusions:** Testing primarily or exclusively in the dominant language is best for detecting AD naming impairments in bilinguals. However, AD affects the ability to access names in both languages, possibly for different reasons, and simple descriptions of language decline as "parallel" or "asymmetrical" (i.e., AD affecting one language more than the other) may be misleading in terms of the theoretical implications for bilingual language processing.

### Keywords

bilingualism; language dominance; picture naming; Alzheimer's disease

---

Correspondence concerning this article should be addressed to Tamar H. Gollan 9500 Gilman Drive, La Jolla, CA 92093-0948  
tgollan@ucsd.edu.

We have no conflicts of interest to disclose.

Data are publicly available: <https://tinyurl.com/4xuys9xs>

Little is known about how Alzheimer's disease (AD) affects language abilities in bilinguals despite important implications this may have for models of bilingual language production and for maximizing diagnostic accuracy in clinical settings. What is the expected presentation of language impairment in bilinguals with AD? Assuming both languages access a shared store of meaning representations (Kroll & Stewart, 1994; Francis, 1999; 2020), and progressive deterioration of semantic knowledge in AD (Butters, et al., 1987; 1990; Monsch et al., 1994; Rascovsky et al., 2007; Salmon et al., 1999), both languages should become impaired with disease progression. However, assuming production of the nondominant language requires more executive control ability (to control competition from the dominant language; Abutalebi & Green, 2007), the nondominant language should be more affected than the dominant language, particularly with the progression of executive control impairments in more advanced stages of cognitive impairment (Ivanova et al., 2014). It is not known how diagnostic sensitivity of language tests is influenced by degree of bilingualism and by the relative dominance of the language of testing, or if the same cognitive mechanisms lead to decline in both languages.

Evidence concerning bilingual language impairment is mixed, with every possible pattern of impairment previously reported. This includes *parallel decline* where both languages are affected to the same extent (e.g., Costa et al., 2012; Gollan et al., 2010), and *asymmetrical decline* where the dominant language is more affected by AD than the nondominant language or vice versa (e.g., Gollan et al., 2010; Ivanova et al., 2014). In some studies, more than one pattern of decline was found for different types of bilinguals or within the same bilinguals on different tasks. One of the first studies compared bilinguals with AD to controls on the ability to name pictures on the Boston Naming Test (BNT; Gollan et al., 2010). For about half of the bilinguals English was the *dominant* language (and Spanish was the *nondominant* language) and for the other Spanish was the dominant language. In English-dominant bilinguals the difference between patients and controls was significantly larger in the dominant than in the nondominant language, an asymmetrical pattern of impairment. Additionally, an *either-language scoring* procedure (i.e., giving credit for correctly naming a picture in either English or Spanish), significantly reduced the difference between patients and controls. In the same study, Spanish-dominant bilinguals exhibited a similar pattern, but the difference between patients and controls was statistically equivalent in the two languages—i.e., a parallel pattern of impairment, and the either-language scoring procedure did not significantly reduce the size of the difference between patients and controls.<sup>1</sup> These results suggested that the most sensitive way to distinguish patients from controls with a picture naming test is to test only in the dominant language without allowing use of the nondominant language. Testing in the nondominant language often produces lower scores and false positive diagnoses, and can be influenced by many factors (e.g., age of acquisition, immersion experience, frequency and context of use, years of formal

---

<sup>1</sup>Importantly, in both bilingual groups, naming scores were substantially higher in the dominant than in the nondominant language. Many other differences between the two groups might have affected the results. Spanish-dominant bilinguals had lower naming scores in both languages, lower education level, had acquired English at a later age, were tested with a measure developed for their nondominant language (the Boston Naming Test; Kaplan et al., 1983), and were immersed in their nondominant language. Additionally, Spanish-dominant patients may have declined more at the time of testing than English-dominant patients. Dementia Rating Scale (DRS; Mattis, 1988) scores were about 9 points lower for Spanish-dominant vs English-dominant patients, while controls in the two language-dominance groups were matched for DRS scores.

education and similarity of the two languages). This conclusion fits with a general consensus among neuropsychologists that assessment of cognition in bilinguals is most accurate when testing is done in the dominant language (Ardila & Ramos, 2008; Gasquoine & Gonzalez, 2012; Rivera Mindt et al., 2008), but is not expected based on models of bilingualism and cognitive decline in AD (which as explained above predict either parallel decline or that the nondominant language should be more affected than the dominant language).

Another study of immigrants to Germany who spoke German and a variety of other languages found asymmetrical decline in one task but parallel decline in another (Kowoll et al., 2015). About half of the participants had immigrated early in life and were German-dominant. Overall, bilinguals with MCI or AD exhibited larger deficits (relative to controls) when they performed a semantic fluency test in their dominant versus their nondominant language. In contrast, the same patients had equivalent deficits in their two languages on a picture naming task (i.e., a 15-item short version of the BNT).

A parallel decline pattern (with both languages equally affected by AD) was reported in a few cross-sectional studies. One study of mostly Spanish-dominant Spanish-English bilinguals found equal differences between patients and controls on semantic and letter fluency tests in both languages (Salvatierra et al., 2007). Another study of relatively balanced Catalan-Spanish bilinguals found equivalent differences between patients and controls in both languages on tests of picture naming and translation (Costa et al., 2012). Another study tested individuals who immigrated to Switzerland late in life but performed similarly in French as a second language and either German, Spanish or Italian as the native language. These relatively balanced immigrant bilinguals exhibited similar differences between patients and controls in both languages on tests of picture naming, repetition, fluency, and semantic and syntactic comprehension (Manchon et al., 2015). Parallel effects of AD in both languages across so many different tests led the authors to conclude: “...it would appear that at present there is no reason to choose one language over the other for the purposes of neuropsychological assessment of DAT [Dementia of Alzheimer’s Type]” (Manchon et al., 2015, pp. 98).

Whether one sees a parallel, asymmetric or mixed pattern of language decline in bilinguals with AD may depend, in part, upon stage of disease. Ivanova et al (2014) examined picture naming data over three successive years of testing of a subset of the bilinguals studied in Gollan et al. 2010. Although the degree of naming impairment in bilinguals with AD was greater in the dominant than the non-dominant language at baseline, greater decline occurred over the next two years in the nondominant than in the dominant language naming scores. This pattern is consistent with anecdotal reports that bilinguals with AD exhibit decreasing inclination to speak in a nondominant language over time after diagnosis (Ardila & Ramos, 2008; Mendez et al., 1999; 2019; for review see Gómez-Ruiz et al., 2012). Thus, while the dominant language may be more useful for diagnosis of AD in bilinguals (in cross-sectional comparisons of patients to controls), the nondominant language may exhibit more precipitous decline with disease progression.<sup>2</sup> This may apply particularly to

---

<sup>2</sup>Ivanova et al. (2014) had 8 English-dominant and 4 Spanish-dominant bilinguals with AD. When excluding Spanish-dominant bilinguals, the pattern of decline over time was statistically equivalent for the two languages. Spanish-dominant patients in this study

bilinguals with one clearly dominant language given that a longitudinal study of relatively more balanced bilinguals with AD reported parallel decline of the two languages (Calabria et al., 2017).

The present study aimed to further examine the effects of AD on naming ability in the bilinguals' two languages with greater focus on diagnostic sensitivity. We hypothesized that pictures of varying difficulty level might be differentially sensitive to AD in the two languages, and that previous results might have been distorted by use of the BNT (Gollan et al., 2010), which was not designed for bilinguals and has items that are more difficult (or impossible) in Spanish than in English (Allegri et al., 1997; Gollan et al., 2012; Kohnert et al., 1998; thereby biasing towards English dominance). Previous work also focused exclusively on which language exhibited a larger difference between groups of patients and controls but without considering diagnostic utility on an individual patient level. To address these limitations, we used the Multilingual Naming Test (MINT; Gollan et al., 2012), which was designed for bilinguals and includes a range of item difficulty levels, and conducted analyses of sensitivity and specificity contrasting different approaches to scoring bilingual naming abilities (dominant-language, nondominant-language, either-language, both-languages) to determine which approach might be best for classifying bilinguals as patients versus controls.

## Transparency and Openness

We follow Journal Article Reporting Standards (JARS; Kazak, 2018) and report how we determined our sample size, all data exclusions (if any), all manipulations, and all measures collected in the study. All study data, analysis code, and research materials are available at [<https://tinyurl.com/4xuys9xs>]. Study design and analyses were not pre-registered.

## Method

### Participants

Participants were recruited from the Alzheimer's Disease Research Center (ADRC) at the University of California, San Diego (UCSD). Annual ADRC evaluations included detailed clinical and medical history, brief medical examination, neurological and neuropsychological assessment, screening for depression and other psychiatric symptoms, assessment of functional activities of daily living, and laboratory tests. At the completion of the evaluation at least two ADRC board-certified neurologists reviewed all information (in consultation with neuropsychologists) and classified each participant as either cognitively normal, with AD (using current NIA-AA diagnostic criteria; McKhann, et al., 2011) or with other neurodegenerative diseases (e.g., frontotemporal dementia, dementia with Lewy bodies) based on published criteria. The research protocol was approved by the UCSD Institutional Review Board in accordance with the Helsinki Declaration.

---

also had lower education level but *higher* naming scores than the English-dominant patients – a combination that is not easy to interpret.

Participants' demographic and language history characteristics are shown in Table 1. The dominant language was determined based on participants' stated preference for language of neuropsychological testing. Patients and controls did not differ in age, education, age of regular use of each language, current use of the nondominant language, or years immersed in the nondominant language (i.e., years lived outside the USA for English-dominant participants and years lived in the USA for Spanish-dominant participants). All participants had learned Spanish before English. After dividing the sample by language preference (English-dominant, Spanish-dominant) subgroups remained relatively well matched except that in the Spanish-dominant subset patients reported being slightly less proficient in English than controls (see Table 1). Importantly, within each subgroup (i.e., Spanish-dominant and English-dominant) patients and controls did not differ significantly (all  $p$ s  $> .212$ ) in degree of bilingualism as measured by an objective *bilingual index score* calculated by dividing the nondominant language naming score by the dominant language naming score (see Table 2; Gollan et al., 2012; Garcia & Gollan, 2021).

The MINT is administered annually at the ADRC. For most participants, data were taken from the first time they were tested on the MINT (all controls and 22/26 of patients). All participants self-identified as Hispanic and about 52% were born in the USA ( $n=34$ ). About 85% ( $n=55$ ) were Mexican or Mexican-American ( $n=3$  from South America,  $n=3$  from Cuba,  $n=2$  from Central America, and  $n=1$  each from Puerto Rico and Portugal).

Inclusion criteria were as follows: We included all Spanish-English bilingual ADRC participants who were classified as either a healthy control ( $n=39$ ) or who were diagnosed with dementia ( $n=26$ ;  $n=24$  with probable AD,  $n=1$  with possible AD, and  $n=1$  with frontotemporal dementia who was later autopsy confirmed to have AD), and for whom item-level data coding was available in both languages on all 68-items of the Multilingual Naming Test (MINT; Gollan et al., 2012). A number of participants were excluded for various reasons (some are counted in more than one category) as follows: (a) four participants who named only 1–4 pictures correctly in the nondominant language were considered functionally monolingual and excluded (all other participants named at least 11 pictures in the nondominant language), (b) to match patients and controls for education level we excluded 14 participants with 6 or fewer years of education, and 9 Spanish-dominant controls with high education level (patients, especially Spanish-dominant patients, tended to have lower education level), (c) to match for number of years immersed in the nondominant language we excluded 5 Spanish-dominant controls who had lived in the USA for fewer than 20 years (all Spanish-dominant patients had lived in the USA for at least 20 years). Finally, (c) we excluded two participants who were diagnosed with AD six and eight years after they were first tested on the MINT (these participants had identical naming scores in the two languages).

## Materials & Procedure

Participants were tested individually in a quiet, well-lit room by a native Spanish-English bilingual psychometrist. The 68-item version of the MINT was administered in each language according to standardized instructions, including provision of semantic and/or phonemic cues if needed (Gollan et al., 2012; Ivanova et al., 2013). An item was considered

correct if the picture was named spontaneously (i.e., without a cue) or following a semantic (but not a phonemic) cue. All 68 items were administered unless testing was stopped because 6 consecutive items had been failed (i.e., named incorrectly or not named at all; note that the MINT items become progressively more difficult with increasing item number; Gollan et al., 2012). When this stopping rule was invoked, all subsequent items were counted as incorrect. The test was completed first in the participant's stated preferred language and immediately thereafter in the other language to minimize any possible effects of testing order given that some studies showed the dominant language affected by first testing bilinguals in the nondominant language (e.g., Guo et al., 2011; Misra et al., 2012; Wodniecka et al., 2020), though note that in our studies with the MINT we found no order effects on either the dominant or the nondominant languages (i.e., naming scores in each language did not differ as a function of whether testing was done in that language first or after naming the same pictures in the other language first; Van Assche et al., 2013; Garcia & Gollan, 2021).

### Statistical Analysis

We compared groups on demographic and clinical variables using independent samples *t*-tests. Next, we conducted a  $2 \times 2$  generalized logistic mixed effects model with MINT item accuracy as the dependent variable (i.e., 0 or 1) and fixed factors of diagnosis group (AD versus control; contrasted coded as  $-.5$  and  $.5$ ) and language (dominant versus nondominant; coded as  $.5$  and  $-.5$ ), and their interaction. Random intercepts for participant and item were included, along with by-participant and by-item random slopes for the main effect of language (the maximal model did not converge). Next, we explored whether group effects were moderated by item difficulty since the MINT includes a broader range of item difficulty than the BNT (to accurately measure proficiency in the nondominant language; Gollan et al., 2012). Item difficulty was added as a continuous fixed effect (scaled and centered) to the model, removing the random effects for items. We used a model comparison (with the 'anova' function in R) to examine whether a more complicated model improved model fit using a combination of a Chi-square statistic and evaluation of AIC (Akaike information criterion). In exploratory correlations, we examined whether bilingual language characteristics and Dementia Rating Scale (DRS) scores predict a benefit from the either-language scoring procedure. Finally, we used Receiver Operating Characteristic (ROC) curves to examine the diagnostic sensitivity of the various MINT scoring methods. Since only a small number of participants were Spanish-dominant, we did not conduct separate analyses for each dominance group. However, Tables 1–2 show both overall results and separated for English-dominant versus Spanish-dominant subgroups. Below we speculate how language dominance might have affected the observed patterns of results (which would need to be tested in future studies). Data were analyzed using Statistical Package for Social Sciences (SPSS) v28 and R version 4.3 (R Core Team, 2020) and the package lme4 version 1.1.26 (Bates et al., 2015) using the 'glmer' function.

### Results

Table 2 presents means and standard deviations for the percentage of pictures correctly named by patients with dementia and healthy controls in the dominant language (whichever

language participants reported as the preferred language for testing), nondominant language, either language (i.e., counting as correct any picture named correctly in at least one language), and both languages (i.e., counting as correct only pictures that were named correctly in both languages). All but 6 bilinguals named more pictures correctly in the language they reported as dominant than in the nondominant language (2 patients and 4 controls named between 1–3 fewer pictures correctly in their preferred language, which was English in 5/6 of these cases).

### Which language is more affected by AD?

Table 3 presents detailed model output (see Table 2 for descriptive statistics). Patients named fewer pictures than controls, participants (collapsed across groups) named more pictures in the dominant- than in the nondominant-language, and unlike in Gollan et al., 2010, AD affected the two languages equally (i.e., the interaction between group and language dominance was not significant. The effect of group remained significant and of the same effect size when controlling for age, education, and sex [ $B=1.396$ ; 95%  $CI(B)=[0.727, 2.075]$ ;  $p<.001$ ].

To assess possible effects of item-difficulty, we repeated the analysis, adding item-difficulty level as a continuous factor (Table 4 presents detailed model output; see Figure 1). This revealed main effects of diagnosis group and language-dominance, as previously reported. Additionally, there was a robust main effect of item difficulty, such that correct responses decreased with each successive increase in MINT item difficulty level. There was an interaction between language dominance and item, such that language dominance effects were smaller on easier items. The interaction between participant group and item difficulty was also significant, as the difference between patients and controls was larger for difficult than for easier items. Of greatest interest, was the 3-way interaction between diagnosis, language-dominance, and item difficulty, with an odds ratio of 2.1 (which indicates at least a small effect size; Chen et al., 2010). Follow-up comparisons revealed a significant interaction between diagnosis group and item difficulty level in the dominant language [ $B=0.780$ ; 95%  $CI(B)=[.350, 1.233]$ ;  $p<.001$ ], but not in the nondominant language, [ $B=0.059$ ; 95%  $CI(B)=[-0.176, 0.293]$ ;  $p=.620$ ]. A model comparison revealed that the 3-way interaction explained significantly more variance compared to a model without this term ( $\chi^2(1)=8.392$ ;  $p=.004$ ) and achieved a lower AIC (i.e., 5,601) compared to a model without this term (AIC=5,608). Figure 1 shows that the differences between patients and controls increased with item difficulty in the dominant language, whereas in the nondominant language group differences were about equal across all items.

### Either Language Scoring Effects

Figure 2a shows the number of additional pictures named correctly when bilinguals were given credit for unique pictures named in either language (across the two MINT administrations) versus just in the dominant language as a function of their bilingual index scores. Overall, bilinguals named more pictures correctly with the either-language scoring procedure, consistent with previous reports (Gollan et al., 2007; 2010; Kohnert et al., 1998), and the benefit increased with an increasing degree of bilingualism. Importantly, patients and controls benefitted equally from the either-language scoring procedure. These

observations were confirmed in a multiple regression with diagnosis group as a fixed factor, bilingual index as a continuous predictor, and the group by bilingual index interaction, with the either-language scoring benefit (i.e., the either-language score minus the dominant language score) as the dependent variable. This analysis revealed only a main effect of bilingual index score ( $B=3.850$ ;  $SE(B)=0.867$ ;  $\beta=0.582$ ;  $p<.001$ ). Neither the main effect of group nor the group by bilingual index interaction effect approached significance ( $p$ s  $.630$ ).

Figures 2b–d show independent measures of proficiency in relation to the either-language scoring benefit. These suggested that in addition to more balanced bilingualism (including both as objectively measured by the MINT, and by self-reported proficiency level), increased current use of the nondominant language might be associated with a greater likelihood of benefitting from the either-language scoring procedure (albeit perhaps only in controls). However, degree of cognitive impairment (as measured by lower Dementia Rating Scale scores) did not predict the extent of either-language scoring benefit.

### Classifying Bilinguals as Patients versus Controls

Figure 3 presents ROC curves classifying bilinguals as patient vs. control based on dominant-language, nondominant-language, either-language (i.e., correctly named in at least one language), and both-languages (correctly named in both languages) scores. The area under the curve (AUC) values and optimal cutoff scores to maximize sensitivity and specificity for each scoring condition are shown in Table 5. The AUCs for dominant-language scores (.79) and either-language scores (.80) were good-to-excellent, whereas the AUC for nondominant-language scores (.65) and both-language scores (.66) showed poor diagnostic classification and were closer to chance (i.e., .50). The ROC curve and AUC for the either-language score were similar to dominant-language scores ( $z<1$ ;  $p=.726$ ). Thus, the either-language scoring was no better or worse at classifying bilinguals as patients versus controls than scores based only on dominant-language scores. Finally, we examined whether adding nondominant scores will aid classification accuracy by hierarchically entering dominant scores in Block 1, and nondominant scores in Block 2 in a logistic regression model. The nondominant language did not significantly improve model classification ( $p=.195$ ), nor contribute unique predictive value ( $B=-.026$ ;  $SE(B)=.020$ ;  $p=.200$ ).

### Discussion

The present study provides key insights as to how bilingualism and AD jointly affect picture naming. Critically, AD seemed to affect naming in both languages, but exploratory analyses suggested that this varied with item-difficulty. In the dominant language, only difficult items seemed to be affected by AD, whereas in the nondominant language, items of all difficulty levels were affected (see Figure 1). Language dominance effects also varied with item-difficulty, being larger for more difficult items (while bilinguals appeared to be more balanced for relatively easy items - compare means across panels in Figures 1). Overall, we found numerically larger differences between patients and controls in the nondominant than in the dominant language (but note the large variability, and lack of significance; see Tables 2–4). Even though not significant, this result is surprising given that to date, no



cross-sectional study found the nondominant language to be more affected by AD than the dominant language even though this result is predicted by models of bilingual language processing (Gollan et al., 2010) and anecdotal evidence (Mendez et al., 1999), and may best characterize changes in bilingual language production with disease progression (Ivanova et al., 2013; Mendez et al., 1999).

Second, bilinguals named more pictures when credited for names they could produce in either language relative to only in the dominant language, and this benefit increased continuously as the degree of bilingualism approached balance between the two languages, and with increased self-reported use of the nondominant language in controls<sup>3</sup>, but was not modulated by degree of cognitive impairment (see Figure 2). Importantly, the either-language scoring benefit was about equal for patients and controls. Additionally, the either-language and dominant-language naming scores were equivalent for distinguishing patients from controls in ROC analyses. Joint consideration of both the dominant and the nondominant language naming scores did not improve diagnostic sensitivity (above the dominant language score alone).

### Understanding Parallel versus Asymmetrical Impairment Patterns

Our results build on growing evidence that the answer to the question “which language declines more in bilinguals with AD?” may vary, at least in cross-sectional comparisons, as a function of the type of bilinguals tested (i.e., which language is dominant, if bilinguals are balanced or not), the type of task (e.g., fluency versus picture naming), and even difficulty of items within the same task. It is increasingly clear that both languages are affected by AD, but different measures might be needed to reveal the complete picture of the underlying neurocognitive impairments in bilinguals with AD. Although damage to semantic representations should reduce the ability to retrieve words in *both* languages given a shared conceptual representation across languages, different factors will affect which words bilinguals know in each language, what types of processes are needed to retrieve them, and whether they will be sensitive to AD. One interpretation of our results is that retrieval of all object names in the nondominant language is relatively difficult (though difficulty and diagnostic sensitivity are not one and the same as is demonstrated by ROC analyses in the present study; see also Chapman & Chapman, 1978; Melinder et al. 2005). It might be tempting to assume that there is a “diagnostic sweet spot” at about the point where controls fail to name pictures correctly about 75% of the time (this occurs at a different item difficulty range in the two languages, see Figure 1). However, we caution against this possibility because at all levels of difficulty there is simply more measurement error in the nondominant language (compare left and right sided panels of Figure 2).

---

<sup>3</sup>Elsewhere we have argued that self-report measures of language proficiency are considerably less reliable than objective measures based on studies of young Spanish-English and Chinese-English bilinguals (Tomoschuk et al., 2019). Consistent with this view, the lines in panel B of Figure 2 are perhaps slightly less steep than those in panel A; however, there is circularity of measurement in panel A (MINT scores were used to calculate scores on both the X and the Y axes), and the differences are not as striking as would be expected if self-ratings were entirely unreliable (the lines in panel B should have been flat). It is possible that older bilinguals are better than young bilinguals at rating their own proficiency level. It is also not clear why self-reported use of the nondominant language was predictive of the either-language benefit only in controls not in patients. It is possible patients had more difficulty accurately gauging their degree of use or that their use patterns changed because of their cognitive impairment.

Furthermore, items in the dominant versus in the nondominant language might be sensitive to AD for different reasons. Relatively easy items might be readily accessible in both languages leading to competition between languages, which should affect the nondominant language more than the dominant language (Gollan et al., 2008). Similar considerations apply more generally to understanding how bilingualism affects picture naming ability. For example, bilinguals had more tip-of-the-tongue states (TOT) than monolinguals for relatively easy items (e.g., Gollan & Silverberg, 2001; Gollan & Acenas, 2004), but monolinguals had more TOTs than bilinguals when the targets were very difficult words (bilinguals did not know the more difficult words; Gollan & Brown, 2006). Thus, the size and even the direction of the difference between groups can depend on idiosyncratic factors specific to the words chosen for investigation (e.g., Gollan et al., 2010 and Ivanova et al., 2013 suggested words in the dominant language have richer semantic representations leading them to be more sensitive to small changes in cognitive status, but Stilwell et al., 2015 pointed out that richer representations might just as easily be expected to be *less* vulnerable to brain damage). As such, although the majority of cross-sectional studies found parallel effects of AD on both languages (see reviews in Calabria et al., 2017; Stilwell et al., 2017) a parallel decline pattern does not necessarily implicate the same underlying cognitive mechanism.

Longitudinal studies are needed to thoroughly characterize how AD affects language processing in bilingual individuals and to test the hypothesis that the nondominant language declines faster than the dominant language as executive control impairments emerge with disease progression. As noted above, greater decline of the nondominant language with disease progression is expected given that controlling interference from the dominant language is likely the most demanding language control challenge bilinguals face when speaking. To date, only one empirical study has reported this pattern of asymmetrical decline over three years of testing (Ivanova et al., 2013), though with a naming test not ideal for use with bilinguals (the BNT) and with only twelve bilinguals with AD. A parallel pattern of decline in both languages was found in the only other longitudinal study (Calabria et al., 2017), but bilinguals in this study spoke typologically similar languages (Catalan-Spanish) and had similarly high naming scores in both languages (i.e., it could be questioned if one language was clearly dominant<sup>4</sup>). In addition, participants were only followed for one year, which might not have been sufficient to reveal an asymmetrical decline pattern. Because changes to language processing due to AD are subtle, at least initially, testing bilinguals that have one clearly dominant language and having longer test-retest intervals might be needed to observe how changes in language abilities emerge with disease progression. In addition, the effects of AD on language processing may be easier to identify in tasks other than picture naming. Language control failures increase in bilinguals with AD (Costa et al., 2012; Calabria et al., 2017) and difficulty controlling which language to speak may increase with disease progression (for review see Stilwell et al., 2015). Tasks that capture the complexity of naturalistic language production in connected speech (Kavé & Goral, 2017), or that are specifically designed to challenge control over language selection (e.g., Gollan

---

<sup>4</sup>A similar criticism applies to Manchon et al., (2015), but not to Salvatierra et al., (2007), who also observed similar sized AD effects in the two languages even though Spanish was clearly more proficient than English, on average, for those bilinguals.

et al., 2017; 2020), might be better suited for revealing how AD affects bilingual language processing. Going forward it might be more fruitful to focus on tasks that are better designed to pinpoint the underlying cognitive mechanisms that lead to differences between bilinguals with AD and controls. Furthermore longitudinal studies are needed given that abilities in the nondominant language vary considerably between individuals and are affected by factors unrelated to AD (e.g., years of immersion, frequency of use).

### Measuring Language Abilities to Diagnose AD in Bilinguals

The results of the present study help answer a critical outstanding question regarding assessment of picture naming in bilinguals with AD: “Should bilingual individuals be tested only in their dominant language or should names produced in either language be accepted as correct?”. We addressed this question using individual subject analyses that are more useful for diagnostic purposes—i.e, ROC curves with cutoffs to maximize sensitivity and specificity to classify individuals as patients with AD or controls. Based on the size of the difference between patients and controls, the only previous study that attempted to address this question surmised that the either language scoring procedure might obscure the effects of AD on naming scores (Gollan et al., 2010). However, this study did not compare diagnostic sensitivity of the two methods. The results of the present study call for revision of this conclusion; the size of the difference between groups does not necessarily indicate which measure is best for classifying participants into groups, and the either-language scoring procedure worked as well as dominant language naming scores for diagnostic classification purposes. Naming might be most sensitive to AD in bilinguals when they try to retrieve words that are part of a stable and consistently accessible lexicon. For bilinguals this might include most words in the dominant language and only a small number of words in the nondominant language. Though nondominant overall, some words in the nondominant language are actually the dominant label for some specific objects. In the present study, the either language scoring benefit began to emerge at a relatively low level of bilingual proficiency; about where bilinguals could name only about 40–50% as many pictures in the nondominant-language as in the dominant-language (see Figure 2; a bilingual with an index score of .5 can only name half as many pictures in the nondominant as in the dominant language). Previous studies reported that only balanced bilinguals benefit from the either-language scoring procedure (Gollan et al., 2007; 2010; Kohnert et al., 1998), but the present study suggests the effect might apply more broadly to different types of bilinguals (similar issues arise in assessment of bilingual children; Bedore & Peña, 2008; Bedore et al., 2012; Bialystok et al., 2010).

Importantly, we did not encourage bilinguals to use either language *during testing*. That is, we tested bilinguals in the dominant language first to minimize interference between languages since the dominant language is more affected by testing order than the nondominant language (Branzi et al., 2014;2016; Wodniecka et al., 2020)<sup>5</sup>. This is very different from inviting bilinguals to use either language during testing, which would encourage voluntary language switching, and could reduce naming scores in some bilinguals

---

<sup>5</sup>Calabria et al., (2017) and Salvatierra et al., (2007) counterbalanced testing order which introduces yet another factor that could have led to the difference in results found across studies.

(Sheppard et al., 2016). This may introduce language control requirements that could affect the dominant language more than the nondominant language (de Bruin et al., 2020; Gollan & Ferreira, 2009; Gollan et al., 2014; Kleinman & Gollan, 2016), and might be especially affected in bilinguals with AD who have impaired executive function. Investigation of a “use either language *during testing*” instruction is needed to fully characterize how AD affects bilingual naming abilities. Finally, while testing in the nondominant language seems to be less sensitive than testing in the dominant language for detecting AD, different outcomes might be found with tasks that magnify rather than minimize interference between languages (e.g., reading aloud paragraphs written primarily in the nondominant language with some switches to the dominant language; Gollan et al., 2017; 2020).

### Study Limitations

The present study adds to a growing literature on how AD affects bilinguals’ ability to speak two languages, but many important questions remain unanswered. First, the results we reported await verification when the naming test is administered in counterbalanced testing order (dominant first, nondominant first), with an invitation to mix languages freely (to measure “either-language” production in a naturalistic way), and giving participants an opportunity to name every picture (without a stopping rule after 6 failed items; we assumed that most non-administered items beyond the stop point would have been failed to the same extent across languages and across groups, but this is an empirical question). Second, better powered studies are needed to consider if the results we reported might vary with language dominance (see Table 2). We had a relatively small number of participants (especially Spanish-dominant bilinguals) and therefore could not conduct separate analyses by language dominance groups. A notable constant across all studies reviewed above and the present study is that no cross-sectional comparison has ever found a *significantly* greater AD effect in the nondominant than in the dominant language.

### Conclusions

We agree with an emerging consensus in the literature that AD affects bilinguals’ ability to speak in both languages. However, characterization of such impairments as “parallel” is likely misleading. Since it is impossible to know which words bilinguals knew prior to developing cognitive impairment (similar problems arise in bilingual aphasia; Peñalosa et al., 2019), naming scores in the nondominant language provide noisier signal, and in this respect only longitudinal studies can deliver a more definitive answer to the question “which language declines more?”. However, until more is known about how testing in the nondominant language might be useful for distinguishing bilingual patients with AD from controls, testing primarily in the dominant language continues to be the best choice for diagnostic purposes.

### Acknowledgments

This research was supported by grants from the National Institute on Deafness and Other Communication Disorders (DC011492), the National Institute on Aging (AG076415; AG077915), the National Science Foundation Grant BCS1923065, and by a P30 (AG062429) from the National Institute on Aging to the University of California. A.S. was supported by an NRSA postdoctoral fellowship from the National Institute of Neurological Disorders and Stroke (F32NS119285). D.S.S. is supported by an NRSA predoctoral MD/PhD fellowship from the National

Institute on Aging (F30-AG063440). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Institutes of Health.

## References

- Abutalebi J, & Green D (2007). Bilingual language production: The neurocognition of language representation and control. *Journal of Neurolinguistics*, 20(3), 242–275. 10.1016/j.jneuroling.2006.10.003
- Allegri RF, Mangone CA, Fernandez Villavicencio A, Rymberg S, Taragano FE, & Baumann D (1997). Spanish Boston Naming Test norms. *The Clinical Neuropsychologist*, 11, 416–420.
- Ardila A, & Ramos E (2008). Normal and abnormal aging in bilinguals. *Dementia and Neuropsychologia*, 2, 242–247. [PubMed: 29213579]
- Bedore DLM & Peña ED (2008). Assessment of Bilingual children for identification of language impairment: current findings and implications for practice. *International Journal of Bilingual Education and Bilingualism*, 11, 1–29. doi: 10.2167/beb392.0
- Bedore LM, Peña ED, Summers CL, Boerger KM, Resendiz MD, Greene K, ... Gillam RB (2012). The measure matters: language dominance profiles across measures in Spanish–English bilingual children. *Bilingualism: Language and Cognition*, 15, 616–629. doi: 10.1017/S1366728912000090 [PubMed: 23565049]
- Bialystok E, Luk G, Peets KF, & Yang S (2010). Receptive vocabulary differences in monolingual and bilingual children. *Bilingualism: Language and Cognition*, 13, 525–531. doi: 10.1017/S1366728909990423 [PubMed: 25750580]
- Branzi FM, Martin CD, Abutalebi J, & Costa A (2014). The after-effects of bilingual language production. *Neuropsychologia*, 52:102–116. [PubMed: 24144955]
- Branzi FM, Della Rosa PA, Canini M, Costa A, Abutalebi J (2016). Language Control in Bilinguals: Monitoring and Response Selection. *Cerebral Cortex*, 26, 2367–2380. [PubMed: 25838037]
- Butters N, Granholm E, Salmon DP, Grant I, & Wolfe J (1987). Episodic and semantic memory: A comparison of amnesic and demented patients. *Journal of Clinical and Experimental Neuropsychology*, 9, 479–497. [PubMed: 2959682]
- Butters N, Salmon DP, & Heindel WC (1990). Processes underlying the memory impairments of demented patients. In Goldberg E, (Ed.) *Contemporary Neuropsychology and the Legacy of Luria*. pp 99–126. Hillsdale NJ: Erlbaum.
- Calabria M, Cattaneo G, Marne P, Hernández M, Juncadella M, Gascón-Bayarri J, ... & Costa A (2017). Language deterioration in bilingual Alzheimer’s disease patients: a longitudinal study. *Journal of Neurolinguistics*, 43, 59–75.
- Chapman LJ, & Chapman JP (1978). The measurement of differential deficit. *Journal of Psychiatry*, 14, 303–311.
- Chen H, Cohen P & Chen S (2010). How Big is a Big Odds Ratio? Interpreting the Magnitudes of Odds Ratios in Epidemiological Studies, *Communications in Statistics - Simulation and Computation*, 39:4, 860–864, DOI: 10.1080/03610911003650383
- Costa A, Calabria M, Marne P, Hernandez M, Juncadella M, Gascon-Bayarri J, ... Reñé R (2012). On the parallel deterioration of lexico-semantic processes in the bilinguals’ two languages: Evidence from Alzheimer’s disease. *Neuropsychologia*, 50, 740e753. [PubMed: 22261401]
- de Bruin A, Samuel AG, & Duñabeitia JA (2020). Examining bilingual language switching across the lifespan in cued and voluntary switching contexts. *Journal of Experimental Psychology: Human Perception and Performance*, 46(8), 759–788. 10.1037/xhp0000746 [PubMed: 32324035]
- Dixon P (2008). Models of accuracy in repeated-measures designs. *Journal of Memory and Language*, 59, 447–456.
- Francis WS (1999). Cognitive integration of language and memory in bilinguals: Semantic representation. *Psychological Bulletin*, 125, 193–222. [PubMed: 10087936]
- Francis WS (2020). Shared core meanings and shared associations in bilingual semantic memory: Evidence from research on implicit memory. *International Journal of Bilingualism*, 24(3), 464–477. doi:10.1177/1367006918814375

- Garcia DL, & Gollan TH (2021). The MINT Sprint: Exploring a Fast Administration Procedure with an Expanded Multilingual Naming Test. *Journal of the International Neuropsychological Society: JINS*, 1–17. 10.1017/S1355617721001004
- Gasquoine PG & Gonzalez CD (2012). Using monolingual neuropsychological test norms with bilingual Hispanic Americans: application of an individual comparison standard. *Archives of Clinical Neuropsychology*, 27, 268–726. [PubMed: 22337934]
- Gollan TH & Acenas LA (2004). What is a TOT?: Cognate and translation effects on tip-of-the tongue states in Spanish-English and Tagalog-English bilinguals. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 30, 246–269. [PubMed: 14736310]
- Gollan TH, & Brown AS (2006). From tip-of-the-tongue (TOT) data to theoretical implications in two steps: When more TOTs means better retrieval. *Journal of Experimental Psychology: General*, 135, 462–483. [PubMed: 16846276]
- Gollan TH, Fennema-Notestine C, Montoya RI, & Jernigan TL (2007). The Bilingual Effect on Boston Naming Test performance. *The Journal of the International Neuropsychological Society*, 13, 197–208. [PubMed: 17286875]
- Gollan TH, & Ferreira VS, (2009). Should I stay or should I switch? A cost-benefit analysis of voluntary language switching in young and aging bilinguals. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 35, 640–665. [PubMed: 19379041]
- Gollan TH, Kleinman D, & Wierenga CE (2014). What's easier: Doing what you want, or being told what to do? Cued versus voluntary language and task switching. *Journal of Experimental Psychology: General*, 143, 2167–2195. [PubMed: 25313951]
- Gollan TH, Li C, Stasenko A, & Salmon DP (2020). Intact reversed language-dominance but exaggerated cognate effects in reading aloud of language switches in bilingual Alzheimer's disease. *Neuropsychology*, 34, 88–106. [PubMed: 31545627]
- Gollan TH, Montoya RI, Cera CM, & Sandoval TC, (2008). More use almost always means smaller a frequency effect: Aging, bilingualism, and the weaker links hypothesis. *Journal of Memory and Language*, 58, 787–814. [PubMed: 19343088]
- Gollan TH, Salmon DP, Montoya RI, Da Pena E (2010). Accessibility of the nondominant language in picture naming: A counterintuitive effect of dementia on bilingual language production. *Neuropsychologia*, 48, 1356–1366. [PubMed: 20036679]
- Gollan TH & Silverberg NB (2001) Tip-of-the-tongue states in Hebrew-English bilinguals. *Bilingualism: Language and Cognition*, 4, 63–83.
- Gollan TH, Stasenko A, Li C, & Salmon DP (2017). Bilingual language intrusions and other speech errors in Alzheimer's disease. *Brain & Cognition*, 118, 27–44. [PubMed: 28753438]
- Gollan TH, Weissberger G, Runnqvist E, Montoya RI, & Cera CM (2012) Self-ratings of spoken language dominance: A multi-lingual naming test (MINT) and preliminary norms for young and aging Spanish-English bilinguals. *Bilingualism: Language and Cognition*, 15, 594615.
- Gomez-Ruiz I, Aguilar-Alonso A & Espasa MA (2011). Language impairment in Catalan-Spanish bilinguals with Alzheimer's disease. *Journal of Neurolinguistics*. 10.1016/j.jneuroling.2011.06.003
- Guo T, Liu H, Misra M, & Kroll JF (2011). Local and global inhibition in bilingual word production: fMRI evidence from Chinese-English bilinguals. *NeuroImage*, 56, 2300 – 2309. doi:10.1016/j.neuroimage.2011.03.049 [PubMed: 21440072]
- Ivanova I, Salmon DP, & Gollan TH, (2013). The Multilingual Naming Test in Alzheimer's disease: Clues to the origin of naming impairments. *The Journal of the International Neuropsychological Society*, 19, 272–283. [PubMed: 23298442]
- Ivanova I, Salmon DP, Gollan TH (2014). Which language declines more? Longitudinal versus cross-sectional decline of picture naming in bilinguals with Alzheimer's disease. *The Journal of the International Neuropsychological Society*, 20, 534–546. [PubMed: 24725624]
- Jaeger TF (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, 59, 434–446. [PubMed: 19884961]
- Kaplan E, Goodglass H, & Weintraub S (1983). *The Boston Naming Test*. Philadelphia, PA: Lea & Fibiger
- Kavé G & Goral M (2017): Word retrieval in connected speech in Alzheimer's disease: a review with meta-analyses, *Aphasiology*, DOI: 10.1080/02687038.2017.1338663

- Kleinman D, & Gollan TH (2016). Speaking two languages for the price of one: Bypassing language control mechanisms via accessibility-driven switches. *Psychological Science*, 27, 700–714. [PubMed: 27016240]
- Kohnert KJ, Hernandez AE, & Bates E (1998). Bilingual performance on the Boston Naming Test: Preliminary norms in Spanish and English. *Brain and Language*, 65, 422–440. [PubMed: 9843612]
- Kowoll ME, Degen C, Gladis S, & Schröder J (2015). Neuropsychological profiles and verbal abilities in lifelong bilinguals with mild cognitive impairment and Alzheimer's disease. *Journal of Alzheimer's Disease: JAD*, 45(4), 1257–1268. 10.3233/JAD-142880 [PubMed: 25697707]
- Kroll JF & Stewart E (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language*, 33, 149–174.
- Manchon M, Buetler K, Colombo F, Spierer L, Assal F, & Annoni J (2015). Impairment of both languages in late bilinguals with dementia of the Alzheimer type. *Bilingualism: Language and Cognition*, 18(1), 90–100. doi:10.1017/S1366728914000194
- Mattis S (1988). *Dementia Rating Scale: Professional Manual*. Odessa, Florida: Psychological Assessment Resources.
- McKhann GM, Knopman DS, Chertkow H, Hyman BT, Jack CR Jr, Kawas CH, ... & Phelps CH (2011). The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & Dementia: the Journal of the Alzheimer's Association*, 7(3), 263–269.
- Melinder MRD, Barch DM, Heydebrand G, & Csernansky JG (2005). Easier Tasks Can Have Better Discriminating Power: The Case of Verbal Fluency. *Journal of Abnormal Psychology*, 114(3), 385–391. 10.1037/0021-843X.114.3.383 [PubMed: 16117575]
- Misra M, Guo T, Bobb S, & Kroll JF (2012). When bilinguals choose a single word to speak: Electrophysiological evidence for inhibition of the native language. *Journal of Memory and Language*, 67, 224–237. doi:10.1016/j.jml.2012.05.001
- Mendez MF, Perryman KM, Pontón MO, & Cummings JL (1999). Bilingualism and dementia. *The Journal of Neuropsychiatry and Clinical Neurosciences*, 11, 411–412. [PubMed: 10440021]
- Mendez MF, Chavez D, & Akhlaghipour G (2019). Bilingualism Delays Expression of Alzheimer's Clinical Syndrome. *Dementia & Geriatric Cognitive Disorders*, 48, 281–289. doi: 10.1159/000505872 [PubMed: 32045913]
- Monsch AU, Bondi MW, Paulsen JS, Brugger P, Butters N, Salmon DP, & Swenson M (1994). A comparison of category and letter fluency in Alzheimer's disease and Huntington's disease. *Neuropsychology*, 8, 25–30.
- 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. 10.1080/02687038.2019.1666082 [PubMed: 33281268]
- Rascovsky K, Salmon DP, Hansen LA, Thal LJ & Galasko D (2007). Disparate letter and semantic category fluency deficits in autopsy-confirmed frontotemporal dementia and Alzheimer's disease. *Neuropsychology*, 21, 20–30. [PubMed: 17201527]
- Rivera Mindt M, Arentoft A, Kubo Germano K, D'Aquila E, Scheiner D, Pizzirusso M, ... & Gollan TH (2008). Neuropsychological, cognitive, and theoretical considerations for evaluation of bilingual individuals. *Neuropsychology Review*, 18(3), 255–268. 10.1007/s11065-008-9069-7 [PubMed: 18841477]
- Salmon DP, Heindel WC, Lange KL (1999). Differential decline in word generation from phonemic and semantic categories during the course of Alzheimer's disease: Implications for the integrity of semantic memory. *Journal of the International Neuropsychological Society*, 5, 692–703. [PubMed: 10645711]
- Salvatierra J, Rosselli M, Acevedo A, & Duara R (2007). Verbal fluency in bilingual Spanish/English Alzheimer's disease patients. *American Journal of Alzheimer's Disease and Other Dementias*, 22, 190–201.

- Sheppard C, Kousaie S, Monetta L, & Taler V (2016). Performance on the Boston Naming Test in Bilinguals. *Journal of the International Neuropsychological Society*, 22(3), 350–363. doi:10.1017/S135561771500123X [PubMed: 26686697]
- Stilwell BL, Dow RM, Lamers C, & Woods RT (2016). Language changes in bilingual individuals with Alzheimer’s disease. *International Journal of Language & Communication Disorders*, 51(2), 113–127. 10.1111/1460-6984.12190 [PubMed: 26477451]
- Tomoschuk B, Ferreira VS, Gollan TH (2019). When a seven is not a seven: Bilingual Self Ratings Differ Between and Within Language Populations. *Bilingualism: Language and Cognition*, 22, 516–536.
- Van Assche E, Duyck W, & Gollan TH (2013). Whole-language and item-specific control in bilingual language production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 39(6), 1781–1792. 10.1037/a0032859 [PubMed: 23647380]
- Wodniecka Z, Szewczyk J, Kałamała P, Mander P, & Durlak J (2020). When a second language hits a native language. What ERPs (do and do not) tell us about language retrieval difficulty in bilingual language production. *Neuropsychologia*, 141, 107390. [PubMed: 32057934]



### Key points

**Question:**

Alzheimer's disease (AD) makes it more difficult to find words when speaking but it is not known which language is more affected by the disease in bilinguals.

**Findings:**

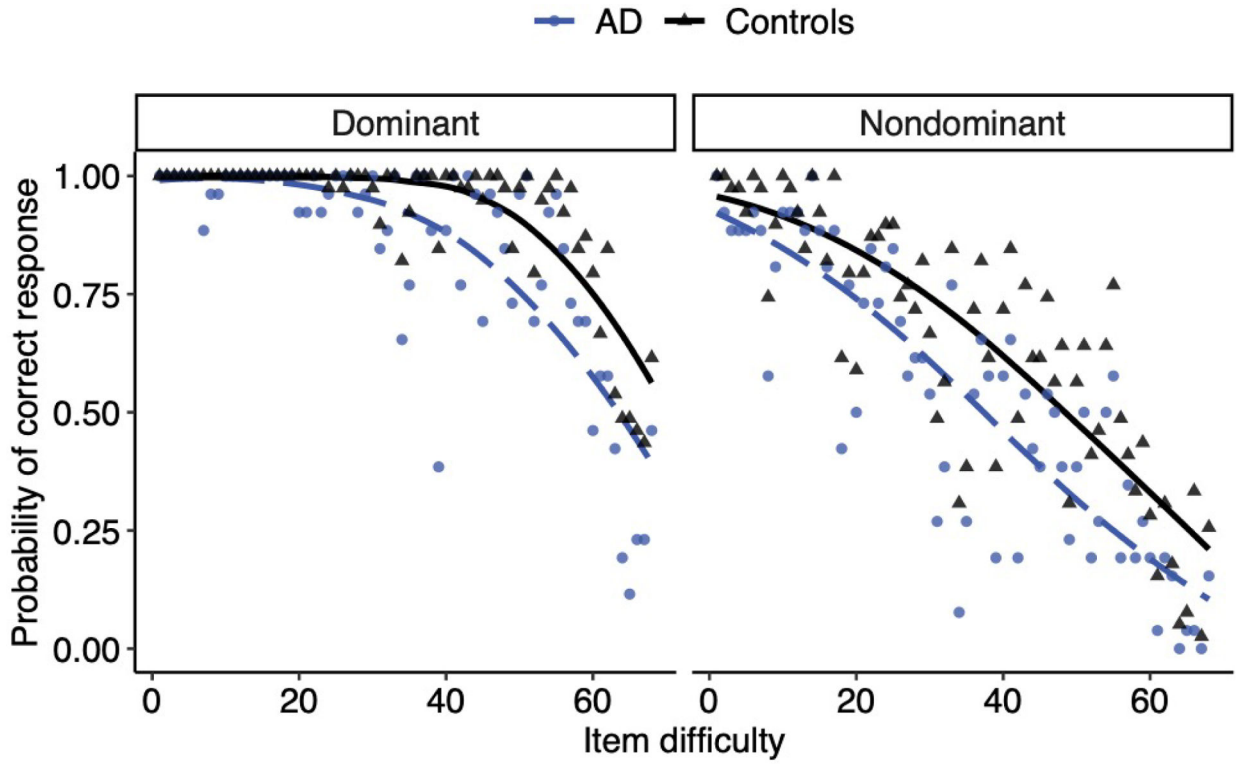
The present study showed that older Spanish-English bilinguals with AD had trouble accessing names in both languages, but in the dominant language difficult items were more affected by AD than easy items while in the nondominant language the effect of AD was similar across all item difficulty levels.

**Importance:**

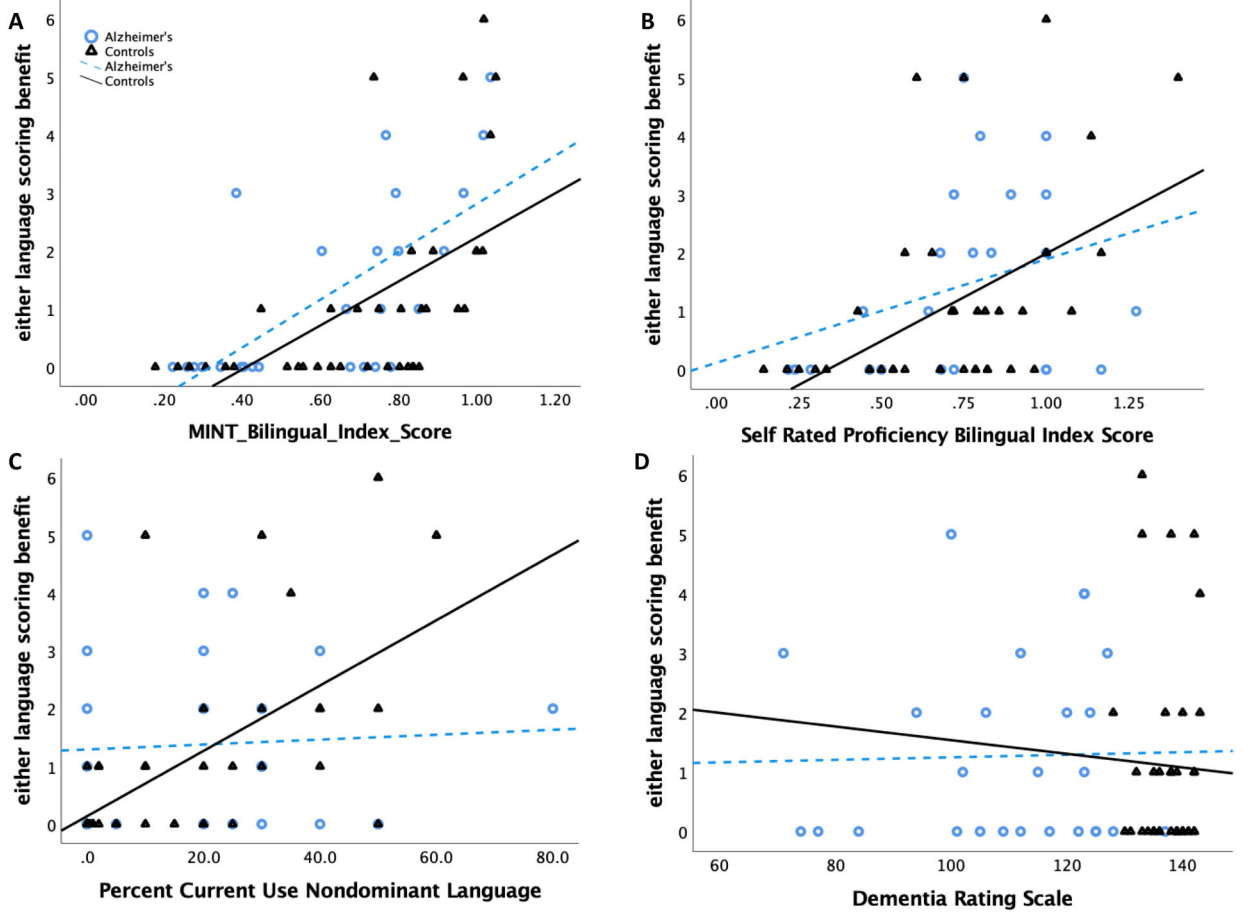
Though both languages are affected by AD a key practical finding was that testing in the dominant language, or using an "either language" scoring procedure, was best for distinguishing patients from controls.

**Next Steps:**

Follow up studies with experimental manipulations of item difficulty and longitudinal studies are needed to better identify how each language is affected by AD in bilingual speakers.

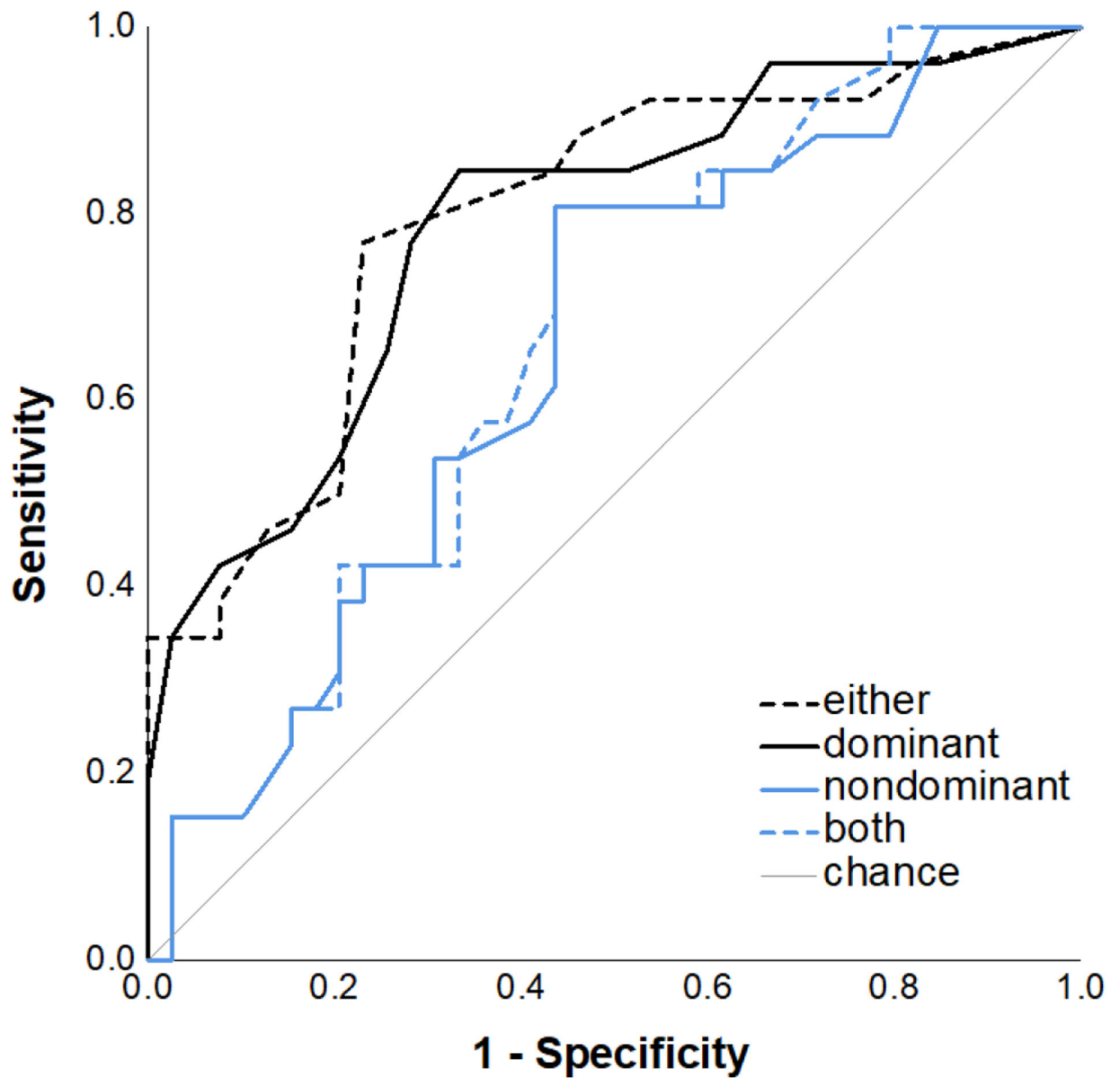


**Figure 1.** Probability of a correct naming response by difficulty (item number) for AD versus controls, plotted separately by dominant versus nondominant language. AD represented with a dashed blue line with individual means plotted as circles. Controls are represented by black solid lines, with individual means plotted as triangles.



**Figure 2A-D.**

The number of additional pictures named correctly when crediting any picture named in either language versus just in the dominant language plotted relative to (A) bilingual index calculated using MINT scores (nondominant/dominant = 1 = perfectly balanced bilingual; scores slightly above 1 indicate a score that is slightly lower in the language participants preferred for neuropsychological testing), (B) bilingual index scores calculated using self-rated proficiency level (averaged across speaking, listening, reading, writing), (C) self-reported percent of current use of the nondominant language, and (D) Dementia Rating Scale scores on the x-axis.



**Figure 3.** Receiver operating characteristic (ROC) curves classifying all bilinguals as AD vs. all cognitively healthy controls using different approaches to scoring naming responses including pictures named correctly when using the either language scoring procedure, in the dominant language only, in the nondominant language only, and in both languages.

**Table 1**

Participant characteristics.

	All Participants				English dominant				Spanish dominant			
	AD	Normal	M	SD	AD	Normal	M	SD	AD	Normal	M	SD
<i>n</i>	26	39			15	28			11	11		
Female/Male	15/11	30/9			7/8	19/9			8/3	11/0		
Right/Left/Ambidextrous	25/0/1	37/1/1			14/0/1	26/1/1			11/0/0	11/0/0		
Demographics and Tests of Cognitive Status	<i>M</i>											
	<i>SD</i>											
	Age at Testing	74.4	74.9	8.3	77.6	74.1	8.6	70.1	13.3	76.7	7.6	
	Years of Education	13.0	13.8	3.0	14.8	14.9	2.5	10.6	2.8	10.9	2.5	
Bilingual Index Scores	Dementia Rating Scale	109.8	137.6	3.8	115.8	138.4	3.5	101.7	18.7	135.4	3.9	**
	Mini Mental State Exam	21.4	28.9	1.2	22.4	29.0	1.3	20.0	4.9	28.8	1.2	**
	Multilingual Naming Test	0.63	0.71	0.25	0.62	0.72	0.27	0.63	0.29	0.67	0.21	
	Self Rated Proficiency	0.74	0.71	0.28	0.74	0.78	0.26	0.74	0.31	0.56	0.27	
Self Rated Proficiency	English Self Rated Proficiency	5.8	5.9	1.7	6.7	6.7	0.6	4.6	1.8	3.8	1.9	
	Spanish Self Rated Proficiency	5.5	5.6	1.5	5.0	5.1	1.5	6.2	0.7	6.9	0.3	**
	Dominant Language Self Rated Proficiency	--	--	--	--	--	--	--	--	--	--	--
	Nondominant Language Self Rated Proficiency	--	--	--	--	--	--	--	--	--	--	--
Other Self Report Variables	English Age Began Regularly Using	8.2	12.1	13.0	4.2	7.3	6.8	15.1	10.5	25.8	16.4	
	Spanish Age Began Regularly Using	2.5	1.4	1.7	3.2	1.2	1.8	1.5	1.5	1.6	1.3	
	%Use of Nondominant Language (Current)	19.4	17.5	17.6	15.7	18.3	18.6	24.5	26.3	15.6	15.7	
	Years Immersed in Nondominant Language	23.3	17.0	18.9	7.3	13.4	11.3	48.2	18.8	40.2	13.7	

\*\* *p* .01;

\* *p* .05;

† .10 for difference between patients and controls

Number of pictures named correctly (total possible score = 68) using different approaches to scoring bilingual naming responses.

**Table 2.**

	Alzheimer's				Controls				diff		
	M	SD	Min	Max	M	SD	Min	Max			
All Participants	either-language	58.4	5.9	42	68	63.9	3.5	56	68	5.5	**
	dominant language	57.2	6.0	39	68	62.8	4.1	54	68	5.6	**
	nondominant language	36.1	15.1	12	58	44.3	15.7	11	65	8.2	*
	both languages	34.8	14.2	12	55	43.2	15.0	11	63	8.3	*
English dominant	either-language	60.1	4.6	54	68	65.1	2.7	60	68	4.9	**
	dominant language	58.9	4.8	53	68	63.8	3.9	54	68	4.9	**
	nondominant language	36.3	13.1	12	58	45.6	16.5	11	65	9.3	
	both languages	35.1	12.1	12	54	44.4	15.8	11	63	9.3	
Spanish dominant	either-language	56.1	6.8	42	63	60.9	3.6	56	66	4.8	
	dominant language	54.7	6.8	39	61	60.1	3.3	56	66	5.4	*
	nondominant language	35.7	18.2	13	58	40.8	13.6	20	64	5.1	
	both languages	34.4	17.3	12	55	40.0	12.8	20	62	5.6	

\*\*  
p < .01;

\*  
p < .05 for difference between patients and controls

**Table 3.**

Logistic mixed effects regression output for main analysis

Fixed effect	B	95 CI (B)	Odds ratio	95 CI (odds ratio)	z-statistic	p-value
Intercept	2.618	[1.874, 3.392]	13.71	[6.508, 28.860]	6.890	<.001
Language	3.395	[2.682, 4.141]	29.81	[14.678, 60.530]	9.393	<.001
Diagnosis	1.405	[0.726, 2.093]	4.08	[2.084, 7.970]	4.105	<.001
Language × Diagnosis	0.291	[-0.989, 1.587]	1.34	[0.378, 4.740]	0.452	.652

**Table 4.**

Logistic mixed effects regression output with additional fixed effect of item difficulty

Fixed effect	B	95 CI (B)	Odds ratio	95 CI (odds ratio)	z-statistic	p-value
Intercept	2.270	[1.977, 2.572]	9.676	[7.220, 12.967]	15.196	<.001
Language	3.093	[2.532, 3.673]	22.044	[12.579, 38.630]	10.807	<.001
Diagnosis	1.522	[0.939, 2.118]	4.582	[2.564, 8.190]	5.138	<.001
Item difficulty	2.041	[1.198, 2.171]	7.692	[6.791, 8.720]	-31.947	<.001
Language × Diagnosis	1.057	[-0.060, 2.199]	2.877	[0.946, 8.750]	1.864	.062
Language × Item difficulty	0.456	[0.209, 0.715]	1.577	[1.128, 2.030]	-3.567	<.001
Diagnosis × Item difficulty	0.418	[0.172, 0.672]	1.520	[1.186, 1.940]	-3.315	.001
Language × Diagnosis × Item difficulty	0.720	[0.230, 1.220]	2.053	[1.254, 3.370]	-2.859	.004

Note: For ease of interpretation of odds ratios (i.e., to be able to interpret all fixed effects on the same scale), item difficulty was reverse-coded in this analysis (but not on x-axis in Figure 1) such that item 1 = item 68, item 2= 67, item 3 = 66 etc.).



**Table 5.**

Area under curve (AUC), 95% confidence intervals, p-values, cutoff values, sensitivity, and specificity for evaluating the ability of different types of naming scores for distinguishing patients from controls.

Naming Score	AUC	p-value	95% CI	Cutoff Value	Sensitivity	Specificity
Either	.796	<.001	[.683, .909]	61.5	.77	.77
Dominant	.788	<.001	[.674, .902]	60.5	.77	.72
Nondominant	.653	.037	[.519, .787]	41.5	.54	.67
Both	.664	.026	[.533, .796]	40.5	.58	.64