Title

Bilingual language intrusions and other speech errors in Alzheimer’s disease

Authors

Gollan, Tamar H
Stasenko, Alena
Li, Chuchu
et al.

Publication Date

2017-11-01

DOI

10.1016/j.bandc.2017.07.007

Published in final edited form as:
Brain Cogn. 2017 November ; 118: 27–44. doi:10.1016/j.bandc.2017.07.007.

Bilingual Language Intrusions and Other Speech Errors in Alzheimer's Disease

Tamar H. Gollan,
University of California, San Diego

Alena Stasenko,
San Diego State University/University of California, San Diego

Chuchu Li, and
University of California, San Diego

David P. Salmon
University of California, San Diego

Abstract

The current study investigated how Alzheimer's disease (AD) affects production of speech errors in reading-aloud. Twelve Spanish-English bilinguals with AD and 19 matched controls read-aloud 8 paragraphs in four conditions (a) English-only, (b) Spanish-only, (c) English-mixed (mostly English with 6 Spanish words), and (d) Spanish-mixed (mostly Spanish with 6 English words). Reading elicited *language intrusions* (e.g., saying *la* instead of *the*), and several types of *within-language* errors (e.g., saying *their* instead of *the*). Patients produced more intrusions (and self-corrected less often) than controls, particularly when reading non-dominant language paragraphs with switches into the dominant language. Patients also produced more within-language errors than controls, but differences between groups for these were not consistently larger with dominant versus non-dominant language targets. These results illustrate the potential utility of speech errors for diagnosis of AD, suggest a variety of linguistic and executive control impairments in AD, and reveal multiple cognitive mechanisms needed to mix languages fluently. The observed pattern of deficits, and unique sensitivity of intrusions to AD in bilinguals, suggests intact ability to select a default language with contextual support, to rapidly translate and switch languages in production of connected speech, but impaired ability to monitor language membership while regulating inhibitory control.

Corresponding Author: Tamar H. Gollan, Department of Psychiatry, University of California, San Diego, 9500 Gilman Drive, 0948, La Jolla, CA 92093-0948, tgollan@ucsd.edu.

Declaration of Conflicting Interests
The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

Keywords

bilingualism; switching; speech errors; Alzheimer's disease; diagnosis; reading-aloud

---

Although there is unanimous agreement that Alzheimer's disease (AD) results in language impairment, there is some debate as to which aspects of linguistic functioning are impaired in early stages of the disease and the underlying cognitive mechanism(s). By some accounts, linguistic impairments in AD primarily reflect damage to semantic representations (Adlam et al., 2006; Butters, Granholm, Salmon, Grant, & Wolfe, 1987; Hodges & Patterson, 1995; Hodges, Salmon, & Butters, 1992) leading to the production of semantic errors, and leaving morphosyntactic aspects of speech relatively intact (here the assumption is that grammatical encoding is relatively modular and automatic; e.g., Kempler, Curtiss, & Jackson, 1987; Kavé & Levy, 2003). Others have suggested that AD results in broader linguistic impairments even in early stages of the disease (e.g., Altmann, Kempler, & Anderson, 2001; Croot, Hodges, Xuereb, & Patterson, 2000), possibly reflecting a primary deficit in working memory (e.g., Almor, Kempler, MacDonald, Anderson, & Tyler, 1999). Still others suggest that semantic errors predominate initially, but the production of morphosyntactic and phonological errors increases with disease progression (e.g., Murdoch, Chenery, Wilks, & Boyle, 1987), a pattern that fits with the propensity of the disease to first affect regions of temporal and parietal cortex followed by progression to cortical regions in the frontal lobes (Forbes-McKay, Shanks, & Venneri, 2013).

A consideration when studying language impairment in AD is that different types of tasks may be better suited than others for revealing specific aspects of linguistic impairment. For example, picture description reveals semantic impairments (as in Kavé & Levy, 2003), but free speech produced during a semi-structured interview may be better suited for exposing morphosyntactic processing deficits (Sajjadi, Patterson, Tomek, & Nestor, 2012). Although a large number of studies have attempted to characterize linguistic impairments in AD, relatively few have considered if and how speech errors might be elicited as a possible diagnostic tool.

In the present study we considered the possible effects of AD on production of bilingual speech errors in a read aloud task. Although reading comprehension is impaired in AD (e.g., Bayles, Tomoeda, & Trosset, 1992; Cummings, Houlihan, & Hill, 1986), reading aloud is a relatively spared skill. Some have argued that reading aloud is one of the skills most resistant to AD, remaining intact even in more advanced stages of disease progression (Cummings, Houligan, & Hill, 1986; Friedman Ferguson, Robinson, & Sunderland, 1992; Sasanuma, Sakuma, & Kitano, 1992; but see Glosser & Grossman, 2004). Indeed, it is only at moderate stages of cognitive impairment that patients may exhibit increased difficulty with reading low-frequency words with irregular spelling-to-sound correspondences (e.g., *caste, cough, or sew* relative to matched regular words such as *carve, couch, and sag*; Strain, Patterson, Graham & Hodges, 1998). Few studies have examined reading aloud in AD beyond the single word level (e.g., Chan, Salmon, & De La Pena, 1999; Monti, Gabrieli, Wilson, & Reminger, 1994; Monti et al., 1997) and none of these have systematically examined production of speech errors during reading aloud though these could reveal the nature of linguistic impairments in AD and the cognitive mechanisms of the speech production system (Fromkin, 1971; Garrett, 1975; 1982).

In recent work we used a paragraph reading task (Kolers, 1966) to demonstrate that reading aloud elicits connected speech in a manner that engages the language production system and leads speakers to produce speech errors that resemble errors produced in spontaneous conversation (Gollan, Schotter, Gomez, Murillo, & Rayner, 2014; Gollan & Goldrick, 2016, in press). For example, bilinguals produced significantly more errors when reading aloud in their non-dominant language than when reading in their dominant language (e.g., producing *kept* instead of *keep*, or *turn* instead of *turning*). In addition, bilinguals reading aloud mixed-language paragraphs sometimes produced *intrusion errors* in which they spontaneously produced translations of written target words in their speech (e.g., saying *pero* instead of *but*). Aging bilinguals produced more intrusions than proficiency-matched young bilinguals in both the read aloud task (Gollan & Goldrick, 2016) and in verbal fluency tasks (Gollan, Sandoval, & Salmon, 2011). Additionally, bilinguals with AD produced more intrusion errors with disease progression in both picture-naming and word translation (Costa, et al., 2012). However, it is not known if bilinguals with AD produce more intrusion errors than cognitively healthy bilingual matched controls.

In previous work we reported a counterintuitive pattern of linguistic impairment in which picture-naming deficits in bilinguals with AD were more robust in the dominant language than in the non-dominant language (Gollan, Salmon, Montoya, & da Pena, 2010). This pattern was apparent in the initial stages of disease progression in bilinguals who had one clearly more proficient language (Ivanova, Salmon, & Gollan, 2014; Kowoll, Degen, Gladdis, & Schröder, 2015), whereas balanced bilinguals exhibited parallel decline of both languages (Costa et al., 2012; Salvatierra, Rosselli, Acevedo, & Duara, 2007). This finding was counterintuitive because naming deficits are usually more pronounced for low-frequency than high-frequency words (Hodges, et al., 1992; Kirshner, Webb, & Kelly, 1984; Ober & Shenaut, 1988; Thomspon-Schill, Gabrieli, & Fleischman, 1999) and bilinguals generally speak their non-dominant language less frequently than their dominant language. Furthermore, anecdotal evidence and caregiver reports suggest that bilinguals with AD increasingly avoid the non-dominant language as dementia progresses (e.g., Mendez, Perryman, Pontón, & Cummings, 1999). Thus, the frequency effect in naming should favor the dominant language. Greater impairment in the dominant language would also be unexpected assuming deficits in executive control in AD (Baudic et al., 2006; Lafleche & Albert, 1995; Perry & Hodges, 1999); on this view, production of the non-dominant language should be more difficult because it requires bilinguals to control interference from the more accessible dominant language (Meuter & Allport, 1999; Green, 1998).

To explain the counterintuitive pattern, we initially suggested that dominant language representations might be more richly represented at the semantic level than non-dominant language representations and, therefore, more sensitive to subtle changes in the integrity of semantic representations at an early stage of disease (Gollan et al., 2010). In subsequent work, however, we found that the dominant language declines more rapidly than the non-dominant language in later stages of disease. To accommodate this finding we relied on the proposal that language impairments in AD may reflect deficits in effortful retrieval (Balota, Watson, Duchek, & Ferraro, 1999; McGlinchey-Berroth & Millberg, 1993; Nebes, Martin, & Horn, 1984; Ober, Shenaut, & Reed, 1995). Retrieval of words in the non-dominant language might generally be more effortful than retrieval of words in the dominant language. However, it is most difficult to retrieve very low frequency words, and the lowest-frequency/most difficult words bilinguals know most likely belong to their dominant language. Thus, if patients with AD have a deficit in effortful retrieval it may be most apparent compared to controls for these low frequency words that are only known in the dominant language. Neither patients nor controls are likely to know very low frequency words in the non-dominant language, so the retrieval deficit will initially not be as apparent in the non-dominant language (Ivanova et al., 2014). From this viewpoint, any task that elicits production of very low frequency words should expose greater differences between patients and controls in the dominant than in the non-dominant language. Relatively easier tasks (i.e., those with less effortful retrieval demands) should reveal the opposite pattern.

In our previous studies with the read aloud task, within-language errors (e.g., function word substitutions, omission errors, inflection errors) elicited the expected pattern of more errors in the non-dominant than the dominant language. Thus, on a difficulty based account, within-language errors in the non-dominant language should be more sensitive to AD than dominant within-language errors. This prediction assumes that the read aloud task circumvents the problem we hypothesized might arise in picture naming; i.e., that the most difficult known words belong to the dominant language. This assumption is justified because the paragraphs selected for the present study generally contained relatively simple language and were not designed to elicit production of very difficult (i.e., low frequency) words. In contrast, targets in picture naming tests become progressively more difficult (i.e., lower frequency) as the test proceeds. In addition, when reading aloud full paragraphs, language production is aided by semantic context and grammatical encoding that is not available in picture-naming tests. We also sought to determine if different types of error subtypes among within-language errors might be differentially sensitive to AD. A previous study, for example, showed that the summed duration of all hesitations produced in a 4 minute sample of spontaneous speech distinguished patients with AD from controls, whereas speech rate (phonemes per second) and grammatical errors (in syntax, or inflectional or derivational morphology) did not (Hoffmann, Nemeth, Dye, Pákáski, Irinyi, & Kálmán, 2010).

Different predictions however, would follow for intrusion errors, which produced significantly *reversed-dominance effects* in previous studies such that bilinguals replaced dominant-language targets with non-dominant-language translations more often than the reverse (i.e., more often than they replaced non-dominant-language targets with dominant-language translations; Gollan et al., 2014; Gollan & Goldrick, 2016, in press). For example, an English-dominant bilingual would be more likely to replace the English word *reason* with its Spanish equivalent, *razón*, when reading aloud a sentence written mostly in Spanish (e.g., *Es por esa reason que digo que la leyenda de La Llorona es verdad*) than he would be to replace the Spanish *razón* with *reason*, when reading the English equivalent sentence (i.e., *It is because of that razón that I say that the legend of the Weeping Woman is true*). Moreover, reversed-dominance effects though slightly smaller, were not significantly smaller in older than in young bilinguals, and neither young nor older bilinguals exhibited more intrusions in the non-dominant than the dominant language (Gollan & Goldrick, 2016).

Importantly, reversed dominance effects have also been observed in young cognitively healthy bilinguals who named pictures more quickly in their nondominant than their dominant language when tested in mixed language blocks (e.g., Christoffels, Firk, & Schiller, 2007; Costa & Santesteban, 2004; Gollan & Ferreira, 2009; Verhoef, Roelofs, & Chwilla, 2009). Full reversal of language dominance effects suggests the operation of an inhibitory control process applied to the dominant language (Gollan & Ferreira, 2009; Kroll, Bobb, Misra, & Guo, 2008), and could also imply activation of the nondominant language (to the point that its accessibility exceeds that of the typically dominant language; for reviews see Declerck & Philipps, 2015a; Runnqvist, Strijkers, Sadat, & Costa, 2011). The finding of intact reversed dominance effects in aging bilinguals with clear impairments in independent measures of nonlinguistic executive control abilities could further imply that language control mechanisms are relatively modular and specialized (Gollan & Goldrick, in press). An alternative possibility is that relatively little control is applied within the linguistic system to achieve language selection – so that even though impaired in normal aging – sufficient control abilities remained intact in aging, which therefore does not elicit impaired dominance effects (Gollan & Goldrick, 2016). Thus, bilinguals with AD will provide a stronger test of the hypothesis that general mechanisms of executive control underlie reversed dominance effects; executive control may be sufficiently impaired in bilinguals with AD to reveal weaker or absent dominance reversal (relative to cognitively intact aging bilinguals).

Finally, although we did not manipulate part of speech in language switched targets (e.g., content words versus function words), we expected to observe different outcomes for intrusions versus within-language errors in which types of words elicited more errors. In previous work, most intrusion errors involved function word targets, whereas most within-language errors involved content word targets (Gollan et al., 2014; Gollan & Goldrick, 2016, in press; these same part of speech effects on intrusion errors have also been found in bilinguals' spontaneous speech, see Poulisse, 1999; Poulisse & Bongaerts, 1994). Indeed, young and cognitively healthy older bilinguals very rarely produced intrusion errors with language switches on content-word targets. Given these results, we elicited switches only on content word targets in mixed-language paragraphs in the present study (see Appendix). The choice to elicit language switches specifically on words that *less* often elicited errors in cognitively healthy bilinguals may seem unexpected, but any intrusions elicited from bilingual patients with AD from such targets may be pathognomonic. Thus, these targets may be ideally suited for distinguishing patients from controls using the read aloud task.

Method

Participants

Twelve Spanish-English bilinguals diagnosed with probable AD (1 with Lewy Body variant), and 32 cognitively healthy controls participated in the study as part of their annual evaluation at University of California San Diego's (UCSD) Shiley-Marcos Alzheimer's Disease Research Center (ADRC). At the time of testing, 7 patients were mildly impaired (Dementia Rating Scale scores between 119 and 134; Mattis, 1988), 4 were moderately impaired (DRS scores between 95 and 112), and 1 was severely impaired (a DRS score of

93). Diagnoses were determined using criteria developed by the National Institute of Neurological and Communicative Disorders and Stroke (NINCDS) and the Alzheimer's Disease and Related Disorders Association (ADRDA; McKhann et al., 1984). Participants with a history of alcoholism, drug abuse, severe psychiatric disturbances, severe head injury, and learning disabilities are excluded from participation in the ADRC study.

To match controls to bilinguals with AD, we excluded 13 cognitively healthy controls whose proficiency score in their non-dominant language was lower than that of any patient (less than 20/68 on the picture naming test described below), controls who had the same proficiency score in both languages based on our objective proficiency measure (our intention to examine language dominance effects precluded inclusion of such balanced bilinguals), or had more than 18 years of education (the highest educated patients had 16 years of education). With these restrictions we were able to select 19 cognitively healthy controls (5 male; 19 Hispanic) who did not differ significantly from our 12 patients (5 male; 10 Hispanic) in age, education, and several language proficiency variables (see Table 1). Objective proficiency was determined by picture naming ability on the Multilingual Naming Test (MINT; Gollan, Weissberger, Runnqvist, Montoya, & Cera, 2012; Ivanova, Salmon, & Gollan, 2013) which consists of 68 black-and-white line drawings, administered in order of increasing difficulty (e.g., item #1 is *hand*, and item # 68 is *axle*). Participant demographics, performance on annual neuropsychological test battery measures (see Weissberger, Salmon, Bondi & Gollan; 2013), and self-reported language history questionnaire responses are summarized in Table 1[1].

All participants reported being exposed to Spanish from birth. Seven patients and 8 controls chose English as their dominant language, while 5 patients and 11 controls chose Spanish. In all but one case, self-selected language dominance matched classification based on the MINT. One bilingual with AD chose English as her preferred language for neuropsychological testing at the ADRC, but rated herself as a 7 on the proficiency scale in Spanish (i.e., "like a native speaker"), but a 6 for English (i.e., "very good"), and objective testing confirmed Spanish dominance as well (51 versus 57/58 correct in English versus Spanish, respectively). This participant's responses were classified based on self-reported dominance in all data analyses reported below (but we repeated critical analyses excluding this participant, and the pattern of results did not change substantially). Of self-reported English-dominant participants, 11 (4 patients, 7 controls) reported the USA as their country of origin, 3 (2 patients, 1 control) reported Mexico, and 1 (patient) Panama. Of self-reported Spanish-dominant participants, 1 (a patient) reported the USA as country of origin, 12 (2 patients, 10 controls) Mexico, 1 (a patient) Columbia, 1 (a patient) Nicaragua, and 1 (a

---

[1] With one exception noted below, the pattern of results did not change when excluding the one severely impaired patient. In addition, the results seemed to apply equally to bilinguals with relatively balanced versus less balanced knowledge of their two languages. For this purpose, we calculated a bilingual index score (Gollan, et al., 2012) for each participant by dividing the non-dominant language MINT score by the dominant language MINT score (and then multiplied by 100). Scores for patients ranged from 43–90% bilingual, and for controls from 43–94% bilingual. We repeated the analyses below including bilingual index scores as a variable and found no significant main effects or interactions with the bilingual index score ($ps \geq .105$), with one exception which was a significant interaction between target language and index score when collapsing all within-language error subtypes ($p < .001$), which was mainly driven by phonological errors ($p = .059$). Bilinguals produced more within-language errors in the non-dominant than the dominant language, and this difference was significantly reduced in bilinguals with high index scores (i.e., in relatively balanced bilinguals). All other effects reported below were not significantly modulated by the extent to which bilinguals were balanced though this should be interpreted tentatively given the small number of participants tested here.

control) Peru. All participants were right-handed except for one control who reported being ambidextrous. Informed consent was obtained from all individuals (or their caregivers when needed) prior to their participation in the research study. Study procedures were approved by the UCSD Institutional Review Board.

Materials

A native Spanish-English bilingual selected and adapted eight paragraphs from published English-Spanish translations of short stories (modified from Gollan & Goldrick, in press). A second native Spanish-English bilingual read through the paragraphs to check for errors and confirm the intended manipulations. Each paragraph was adapted so that it could be presented in each of four conditions (between participants; see *Procedure*): (a) English-only, (b) Spanish-only, (c) English-mixed, (d) Spanish-mixed. The mixed language paragraphs were written primarily in one language but had 6 language switches on content words. These *Switch-out* points were distributed evenly throughout the paragraph and were followed by immediate *Switch-back* points that were switches back into the paragraph's main language. One mixed-language paragraph had just 5 switches (an experimenter error; this paragraph was read by 4 patients and 5 controls). Switches on words with more than 2 syllables, cognates, proper names, and words sometimes classified as function and sometimes as content words were avoided as much as possible. Switches on words that had previously already been switched within the same paragraph were also avoided. On average, paragraphs had 121.8 words ($SD = 10.5$; range 103–137). An example of each type of paragraph is presented in the Appendix.

Procedure

Participants were tested individually in a quiet, well-illuminated room by a Spanish-English bilingual psychometrist. Paragraphs were presented on paper with words printed in Times New Roman font, size 20, double spaced. Each paragraph was presented on a single page. Participants were instructed to read the paragraph aloud as accurately as possible at a comfortable pace, and were audio-recorded and timed with a stop-watch. Each participant read 8 paragraphs that comprised 2 paragraphs in each of 4 conditions presented in the following order: (a) Dominant-language, (b) Non-dominant-language, (c) Dominant-language-mixed, (d) Non-dominant-language-mixed. Prior to reading the first paragraph of each condition, participants completed a shorter practice paragraph. The experimenter corrected participants if they produced any errors during these practice trials.

Counterbalancing assured that each paragraph appeared in every condition between participants. To achieve this level of counterbalancing, paragraphs were randomly paired into groups of two and then rotated across the four conditions using a Latin Square design. Paragraphs were presented in one of 8 different fixed-order lists: 4 lists for English-dominant bilinguals and 4 for Spanish-dominant bilinguals. Each bilingual was presented with just one of these lists. Errors were marked on a coding sheet during testing and were later checked against audio recordings. Errors were defined as any word produced differently from what was written on the page. Examples of error types are shown in Table 2.

Results

Following the methods of Gollan et al. (2014) and Gollan and Goldrick (2016, in press), native Spanish-English bilingual research assistants transcribed errors and classified them into error types. A small number of speech productions could have been classified as more than one type of error or as possibly correct. In these cases, preference was given to 1) classifying them as errors (rather than marking them as correct), 2) classifying them as intrusion errors (whenever these were produced) over other kinds of errors, and 3) classifying them to specific error types over classification as repetitions or other disfluencies. For example, when a participant said the correct target word but then quickly self-corrected to an intrusion error (i.e., the target was *la* and the participant said "*la… the…*"), this production was classified as a single intrusion error for the analyses that follow. In another 5 cases, intrusion errors were initially transcribed as more than one error (e.g., the speaker repeated the same intrusion twice consecutively before continuing to read, or the speaker produced both a within-language error and then an intrusion error saying "*but… porque"* when the target word was *because*) and were counted as a single intrusion error in the analyses that follow. Similarly, 2 accent errors, 3 inflection errors, 1 insertion error, 2 omission errors, 10 function word substitution errors, and 47 phonological errors were considered as a single error even though there were consecutive repetitions or multiple attempts at the target word (e.g., the target was *clearly* and the speaker said "*clearing, cleary"*). In a small number of cases (n=6) bilinguals produced a within language error and also inserted a word afterwards (e.g., producing …*every morning* as "*in the morning"*). In these cases, the insertion error was ignored and the within language error was classified (i.e., as a function-word substitution in this example). Note that many errors classified here as inflection errors are in fact analogous to "omission errors" in MacKay & James, 2004 who reported that older speakers were more likely than young speakers to omit an inflection at the end of a word, e.g., saying *rip* instead of *ripped*.

One assistant (assistant 2) transcribed errors produced by 4 patients and 1 control, another assistant (assistant 1) transcribed errors produced by 8 patients and 18 controls. During coding, one or the other coder was not fully confident about how to classify a production of a target word in 0.09% of cases; these cases were flagged and settled through discussion. Inter-rater reliability was established by having assistant 2 re-code 8 paragraphs (6 mixed language/2 one language only) produced by 8 different bilingual participants (4 patients and 4 controls) that had been initially coded by assistant 1. Participants were chosen for reliability check by randomly sorting patients and controls coded by assistant 1 and selecting the first four in each group. Paragraphs were chosen by selecting those that elicited at least 1 intrusion error. An attempt was made to include at least some mixed-language (n=6) and single-language (n=2) paragraphs, and paragraphs with variations in error rate and with multiple different error types. Error rates for selected paragraphs averaged 12.3% (*SD*=8.4%, range=2.2–25.0% errors). Agreement between raters was very high; classifications matched across raters for 95.5% of words produced (855 words produced correctly and 73 errors that included 7 intrusions, 21 disfluencies, 18 phonological errors, 18 function word substitutions, 2 inflection errors, 2 insertions, 4 omissions, and 1 unrelated word error). On 3.4% (n=33) of words produced, one assistant classified it as an error while the other classified it as correct (this included 2 intrusion errors that were produced in single-language paragraphs, 4 accent errors, 14 disfluencies, 5 phonological errors, 3 function word substitutions, 1 inflection error, 2 insertions, and 2 omissions). The two assistants classified a response as different types of errors on only 1.1% of words (n=11) and none of these involved intrusion errors. These results show that overall there was 100% agreement between raters for intrusion errors produced in mixed-language paragraphs.

In addition to intrusion errors, a number of within-language errors were coded. Within-language errors included: disfluencies, phonological errors, function-word substitution errors, inflection errors, omissions, and a small number of other errors (see Table 2). We conducted detailed analyses of intrusion errors (the error of primary interest) and within-language errors[2], and subset analyses of within-language error subtypes for all subtypes with more than 100 data points. We used logistic mixed-effects regressions (Jaeger, 2008) to analyze both intrusion and within-language errors (using the R package lme4; Bates, Maechler, Bolker, & Walker, 2015). For each error type, we first conducted the analyses with contrast-coded fixed effects of participant type (patients with AD, matched controls), language of the target word (dominant, non-dominant), paragraph type (single-language, mixed-language), and all interactions between these factors. Subjects and individual words were entered as two random intercepts (some models with random slopes failed to converge so we removed the random slopes from all the models). The significance of each fixed effect was assessed via likelihood ratio tests (Barr, Levy, Scheepers, & Tily, 2013).

Intrusion Errors

All 12 bilinguals with AD produced at least 1 intrusion error (*M*=3.0, *SD*=1.7, range=1–7). By contrast, 9/19 controls produced no intrusion errors and only 4 controls produced more than 1 intrusion error (*M*=1.2, *SD*=1.8, range= 0–7). Figure 1 shows the average percentage of words that elicited intrusion errors at switch-out, switch-back, and non-switch points in each condition and group. We began with an analysis of intrusion rates on content word targets at switch-out points in the mixed-language paragraphs. As noted above, switch-out points elicited the majority of intrusion errors in previous work though rarely with content word targets (Gollan & Goldrick, 2016, in press). In the present study, switch-out words were all content words but still elicited the majority of intrusion errors. Patients produced more intrusions than controls at switch-out points ($M = 6.34\%$ vs. 2.00%; $\beta = 1.36$; $SE\ \beta = .58$; $\chi^2\ (1) = 5.60$, $p = .018$), and overall bilinguals produced slightly, but not significantly, reversed dominance effects with more intrusions with dominant than non-dominant language targets ($M = 4.62\%$ vs. 2.72%; $\beta = .54$; $SE\ \beta = .53$; $\chi^2\ (1) = 1.07$, $p = .301$). Planned comparisons showed that bilinguals with AD produced significantly more intrusions than controls with dominant language targets at switch-out points ($M = 9.22\%$ vs. 1.76%; $\beta = 7.86$; $SE\ \beta = 3.37$; $\chi^2\ (1) = 10.81$, $p = .001$), but not with non-dominant language targets ($M = 3.50\%$ vs. 2.23%; $\beta = 1.47$; $SE\ \beta = 2.34$; $\chi^2 < 1$), although the interaction between participant group and target language-dominance was not significant ($\beta = 1.37$; $SE\ \beta = 1.08$; $\chi^2\ (1) = 1.67$, $p = .197$)[3]. Of great interest, patients exhibited significantly reversed

---

[2] In the analysis of within language errors (collapsing all subtypes), the language of an inserted word was classified as the language of whatever word was produced before the inserted word.

dominance effects, producing significantly more intrusions at switch-out points with dominant than non-dominant language targets ($M = 9.22\%$ vs. 3.50%; $\beta = 1.48$; $SE\ \beta = .78$; χ^2 (1) = 4.44, $p = .035$)[4], while controls produced intrusions equally often with dominant and non-dominant language targets ($M = 1.77\%$ vs. 2.23%; $\beta = 3.92$; $SE\ \beta = 4.52$; $\chi^2 < 1$)[5].

Because our switch-out words were all content words, we next considered if similar results would be found in mixed-language paragraphs at all points (i.e., at switch-out, switch-back, and non-switch words) which could potentially elicit intrusions with both function and content word targets. The results of this analysis revealed that intrusions were produced at a strikingly lower rate when all points (collapsing together switch-out, switch-back and no switch) were considered, but otherwise the effects were quite similar to those found at switch-out points alone. In mixed-language paragraphs, patients produced significantly more intrusions than controls ($M = 0.58\%$ vs. 0.18%; $\beta = 1.51$; $SE\ \beta = .46$; χ^2 (1) = 11.30, $p < .001$), and bilinguals produced intrusions equally often in their two languages ($M = 0.34\%$ vs. 0.33%; $\beta = -.19$; $SE\ \beta = .38$; $\chi^2 < 1$). However, the difference between patients and controls tended to be stronger with dominant than with non-dominant language targets; although this interaction between participant type and target language-dominance was only marginally significant ($\beta = 1.27$; $SE\ \beta = .76$; χ^2 (1) = 2.89, $p = .089$). In planned comparisons patients produced significantly more intrusions than controls with dominant language targets ($M = 0.71\%$ vs. 0.11%; $\beta = 2.90$; $SE\ \beta = .78$; χ^2 (1) = 20.25, $p < .001$), while the difference between patients and controls was not significant with non-dominant language targets ($M = 0.44\%$ vs. 0.26%; $\beta = 1.44$; $SE\ \beta = .96$; χ^2 (1) = 2.58, $p = .108$). Dominance effects were not significant in either participant group (χ^2s ≤ 1.43, ps ≥ 0.231).

Bilinguals produced very few intrusions (n=7) in single language paragraphs; thus effects including paragraph type should be interpreted with caution. However, to include at least one analysis that incorporated all intrusions produced, we examined condition effects on all intrusion errors. Bilinguals produced significantly more intrusions in mixed-language than in single-language paragraphs ($M = 0.34\%$ vs. 0.04%; $\beta = 6.13$; $SE\ \beta = 4.83$; χ^2 (1) = 34.53, $p < .001$). Bilinguals with AD did not produce more intrusions than controls overall ($M =$ 0.31% vs. 0.11%; $\beta = -3.19$; $SE\ \beta = 4.83$; $\chi^2 < 1$); however, patients produced significantly more errors than controls in mixed-language paragraphs ($M = 0.58\%$ vs. 0.18%; $\beta = 1.46$; $SE\ \beta = .45$; χ^2 (1) = 10.93, $p < .001$) while there were no significant group differences in single-language paragraphs ($M = 0.03\%$ vs. 0.05%; $\beta = -1.06$; $SE\ \beta = .01$; $\chi^2 < 1$). This interaction between participant type and paragraph type was significant ($\beta = 9.40$; $SE\ \beta =$ 9.64; χ^2 (1) = 4.49, $p = .034$). Surprisingly given results reported above, bilinguals produced more intrusions overall with non-dominant than dominant language targets ($M = 0.21\%$ vs. 0.18%; $\beta = 4.78$; $SE\ \beta = 4.83$; χ^2 (1) = 5.68, $p = .017$). Dominance effects trended in

---

[3] A traditional between subjects ANOVA (as reported in Gollan et al., 2010), revealed a significant interaction between participant type and language-dominance at switch-out points (F_1 (1,58)= 5.52, p= .022). But note that ANOVA is not recommended for analysis of proportional data because of the restricted range in possible means, and potential for scaling artifact in conditions that approach floor or ceiling (Dixon, 2008; Jaeger, 2008; Winer, 1971).

[4] This difference was no longer significant after excluding the one severely impaired bilingual with AD. However, the means trended in the same direction, M= 8.53% vs. 3.82%; β = 1.08; $SE\ \beta$ = 0.83; χ^2 (1) =1.92, p= 0.166.

[5] These results did not change after excluding the one bilingual patient whose self-selected language dominance did not match objective language dominance based on the MINT scores. Patients still produced significantly more intrusions than controls with dominant targets ($p = .002$) but not with non-dominant targets ($p = .610$)

opposite directions in patients vs. controls (χ^2s ≤ 2.31, *p*s ≥ .128), although the interaction between participant type and target language dominance was not significant ($\beta = -7.30$; *SE* $\beta = 9.64$; $\chi^2 < 1$). It should be noted that intrusion rates were very low overall; patients produced only very slightly more errors with dominant than non-dominant targets (*M* = 0.36% vs. 0.26%), while controls produced only very slightly more errors with non-dominant targets than dominant targets (*M* = 0.17% vs. 0.06%). Dominance effects also trended in opposite directions in single versus mixed-language paragraphs. In mixed-language paragraphs bilinguals tended to produce more intrusions with dominant than with non-dominant language targets (*M* = 0.34% vs. 0.33%), whereas in single-language paragraphs they tended to produce more errors with non-dominant than with dominant language targets (0.08% vs 0.01%). These dominance effects were not significant on their own in either single-language or mixed-language paragraphs (*p*s ≥ .250), although the interaction between target language-dominance and paragraph type was significant ($\beta = 9.31$; *SE* $\beta = 9.65$; χ^2 (1) = 4.75, $p = .029$). Finally, in mixed-language paragraphs, the majority of intrusions were produced at switch-out points and therefore only a minority involved function word targets (19/51 or 37%). In contrast, in single-language paragraphs only one intrusion involved a content word target and the rest all had function word targets (6/7 or 86%).

Self-correction rates—In recent work, we found that older bilinguals self-corrected their intrusion errors at a significantly lower rate than younger bilinguals, a result that implicates monitoring in the higher rate of intrusion errors (Gollan & Goldrick, 2016). To consider this possibility we compared self-correction rates in patients versus controls with logistic regressions with participant type as the fixed effect and self-correction as the dependent variable (we did not include language in this model because of the relatively small number of data points available for this analysis). Subjects and individual words were entered as two random intercepts (random slopes were not included due to the failure to converge). Gollan and Goldrick (2016) compared how many intrusions younger versus older bilinguals self-corrected in mid-utterance (i.e., [partial intrusions/(partial intrusions + full intrusions)]; however, in the current study there were very few partial intrusions (see Table 2), and this model revealed no effect of participant type (*M* = 18.18% vs. 15.38% for patients vs. controls, $\beta = 6.01$; *SE* $\beta = 5.03$; χ^2 <1). However, considering full intrusion errors that were subsequently self-corrected, patients self-corrected their errors at a significantly lower rate than controls (*M* = 11.11% vs. 40.91%, $\beta = 1.67$; *SE* $\beta = .67$; χ^2 (1) = 6.84, $p = .009$). A final analysis combining partial and full intrusions revealed the same pattern; bilinguals with AD self-corrected their intrusions significantly less often than controls (*M* = 22.72% vs. 46.15%, $\beta = 1.32$; *SE* $\beta = .68$; χ^2 (1) = 4.48, $p = .034$).

Within Language Errors

We began our analysis of within-language errors by collapsing together all the different error types (see Table 2). First, matching previous studies using the read aloud task (Gollan et al., 2014; Gollan & Goldrick, 2016, in press), bilinguals produced within-language errors much more often than intrusions. Foreshadowing the results briefly, shown in Figure 2 overall by condition, and in Figure 3 when broken down into error subtypes, bilinguals with AD produced significantly more errors than controls. Bilinguals also produced significantly more errors in their non-dominant than in the dominant language. That is, there were significant dominance effects on all but one within-language error subtype (i.e., omissions) in the expected direction – and in the opposite direction as found for intrusion errors, shown in Figure 4. Differences between patients and controls tended, if anything, to be more robust with dominant than with the non-dominant language targets, but not consistently for all error subtypes (e.g., function word substitution errors exhibited trends in the opposite direction).

Collapsing together all within-language error subtypes, bilinguals with AD produced significantly more errors than controls ($M = 9.05\%$ vs. 6.35%; $\beta = .62$; $SE\ \beta = .22$; $\chi^2\ (1) = 7.08$, $p = .008$), and bilinguals produced more errors in the non-dominant than in the dominant language ($M = 10.84\%$ vs. 3.94%; $\beta = -1.17$; $SE\ \beta = .05$; $\chi^2\ (1) = 539.07$, $p < .001$). Despite the lower rate of errors in the dominant language, the difference between patients and controls was more robust in the dominant ($M = 5.66\%$ vs. 2.85%; $\beta = .75$; $SE\ \beta = .22$; $\chi^2\ (1) = 9.97$, $p = .002$) than in the non-dominant language ($M = 12.45\%$ vs. 9.84%; $\beta = .47$; $SE\ \beta = .28$; $\chi^2\ (1) = 2.75$, $p = .098$), a significant interaction between participant type and language dominance ($\beta = .32$; $SE\ \beta = .11$; $\chi^2\ (1) = 8.18$, $p = .004$). Stated from a different perspective, language-dominance effects were significantly diminished in patients versus controls (as reported by Gollan et al., 2010, but see below). Dominance effects were significant in both patients ($M = 12.45\%$ vs. 5.66%; $\beta = -.97$; $SE\ \beta = .07$; $\chi^2\ (1) = 184.89$, $p < .001$) and controls ($M = 9.84\%$ vs. 2.85%; $\beta = -1.34$; $SE\ \beta = .08$; $\chi^2\ (1) = 300.66$, $p < .001$).

As reported in previous studies using the read aloud task, paragraph type did not affect production of within-language errors. Bilinguals produced within language errors equally often in single-language and mixed-language paragraphs, regardless of participant type and language-dominance of the target word (χ^2 s ≤ 1.25, ps $\geq .264$). Therefore, we collapsed across paragraph type in our analyses of within-language error subtypes. Figure 3 shows the rate of five within-language error subtypes that had sufficient data to support interpretation of subset analyses (i.e., disfluencies, phonological errors, function-word substitution errors, inflection errors, and omissions). The goal of these subset analyses was to determine if one or more error subtype could distinguish patients from controls.

Sub-types Analyses—Although disfluencies were the most frequent subtype of within-language error, patients produced only marginally more of these errors than controls ($M = 3.03\%$ vs. 2.16%; $\beta = .42$; $SE\ \beta = .22$; $\chi^2\ (1) = 3.21$, $p = .073$). Combining patients and controls, bilinguals produced more disfluencies in the non-dominant than the dominant language ($M = 3.57\%$ vs. 1.41%; $\beta = -.97$; $SE\ \beta = .08$; $\chi^2\ (1) = 150.04$, $p < .001$), and the difference between patients and controls did not vary significantly with target language dominance, a non-significant interaction between participant type and target language dominance ($\beta = .06$; $SE\ \beta = .16$; $\chi^2 < 1$). Means however trended in the opposite direction of that reported for intrusion errors, the difference between patients and controls in production of disfluency errors was slightly smaller in the dominant ($M = 1.80\%$ versus 1.18%, difference = 0.62%) than in the non-dominant language ($M = 4.27\%$ versus 3.14%, difference = 1.13%; χ^2 s ≤ 2.65, ps $\geq .105$). Dominance effects were significant in each language in both patients and controls (i.e., they produced more disfluency errors in the non-dominant than the dominant language, χ^2s ≥ 62.09, ps $< .001$).

Analysis of phonological errors, the second most common subtype of within-language errors (see Table 2), revealed that patients produced marginally more errors than controls (M = 1.83% vs. 1.59%; β = .50; $SE\ \beta$ = .26; χ^2 (1) = 3.54, p = .060), and bilinguals produced more errors in the non-dominant than in the dominant language (M = 2.85% vs. 0.52%; β = −1.88; $SE\ \beta$ = .13; χ^2 (1) = 259.08, $p < .001$). In the dominant language alone, bilinguals with AD produced significantly more phonological errors than controls (M = 0.78% vs. 0.34%; β = 1.08; $SE\ \beta$ = .43; χ^2 (1) = 2.83, p = .016), while the two groups produced these errors equally often in the non-dominant language (M = 2.87% vs. 2.84%; β = .15; $SE\ \beta$ = .25; $\chi^2 < 1$), a significant interaction between participant type and target language dominance (β = .66; $SE\ \beta$ = .28; χ^2 (1) = 5.61, p = .018). Both patients and controls produced significantly more phonological errors in the non-dominant language than in the dominant language, but this difference was slightly but significantly smaller for patients (M = 0.78% vs. 2.87%; β = −1.69; $SE\ \beta$ = .20; χ^2 (1) = 91.38, $p < .001$) than for controls (M = 0.34% vs. 2.84%; β = −2.29; $SE\ \beta$ = .23; χ^2 (1) = 140.94, $p < .001$). Thus, phonological errors exhibited the same pattern as reported above for within-language errors overall and the same as reported by Gollan et al., (2010) where language dominance effects appeared to be diminished in AD.

Analysis of function-word substitution errors revealed that patients produced significantly more of these errors than controls (M = 2.07% vs. 1.17%; β = .80; $SE\ \beta$ = .19; χ^2 (1) = 15.36, $p < .001$), and overall bilinguals produced more errors in the non-dominant than in the dominant language (M = 2.16% vs. 0.88%; β = −1.07; $SE\ \beta$ = .11; χ^2 (1) = 100.83, $p < .001$). Unlike the pattern reported for phonological errors, the difference between patients and controls was significant in both the dominant (M = 1.28% vs. 0.62%; β = 1.19; $SE\ \beta$ = .31; χ^2 (1) = 13.81, $p < .001$) and the non-dominant (M = 2.86% vs. 1.72%; β = .66; $SE\ \beta$ = .22; χ^2 (1) = 8.31, p = .004) languages, and though the interaction between participant type and dominance was just marginally significant (β = .42; $SE\ \beta$ = .23; χ^2 (1) = 3.26, p = .071), it trended in the opposite direction of that reported above (i.e., larger difference between patients and controls in the non-dominant than the dominant language, 1.14% versus 0.66%, respectively). Dominance effects (i.e., more function word substitution errors in the non-dominant than in the dominant language) were again significant in both patients (M = 2.86% vs. 1.28%; β = −.89; $SE\ \beta$ = .15; χ^2 (1) = 36.58, $p < .001$) and controls (M = 1.72% vs. 0.62%; β = −1.23; $SE\ \beta$ = .19; χ^2 (1) = 49.03, $p < .001$).

Analysis of inflection errors revealed that patients produced significantly more of these errors than controls (M = 0.90% vs. 0.78%; β = .82; $SE\ \beta$ = .35; χ^2 (1) = 5.13, p = .024), and bilinguals produced more errors in the non-dominant language than in the dominant language (M = 1.31% vs. 0.34%; β = −1.50; $SE\ \beta$ = .21; χ^2 (1) = 56.92, $p < .001$). The difference between patients and controls again did not vary significantly by target language dominance (β = .06; $SE\ \beta$ = .16; χ^2 (1) = 2.54, p = .111), but planned comparisons revealed that patients produced more inflection errors than controls in the dominant language (M = 0.66% vs 0.14%, β = 1.60; $SE\ \beta$ = 0.58; χ^2 (1) = 7.11, p = .008), but not in the non-dominant language (M = 1.14% vs. 1.42%, β = .40; $SE\ \beta$ = .42; $\chi^2 < 1$; i.e., the counter-intuitive pattern reported above for phonological errors and in Gollan et al., 2010). Dominance effects (i.e., more inflection errors in the non-dominant than in the dominant language) were again significant in both patients (M = 1.14% vs 0.66%, β = −1.22; $SE\ \beta$ =

0.28; χ^2 (1) = 20.60, $p < .001$) and controls (M = 1.42% vs 0.14%, $\beta = -2.35$; $SE\ \beta = 0.42$; χ^2 (1) = 35.72, $p < .001$).

Analysis of omission errors revealed that patients omitted words significantly more often in the read-aloud task than controls (M = 0.74% vs. 0.40%; $\beta = .70$; $SE\ \beta = .22$; χ^2 (1) = 9.78, $p = .002$), and unlike all other error subtypes which exhibited highly robust language dominance effects, bilinguals produced omissions equally often with dominant and non-dominant targets (M = 0.50% vs. 0.56%; $\beta = -.12$; $SE\ \beta = .16$; $\chi^2 < 1$). The difference between patients and controls did not vary significantly by language dominance, i.e., the interaction between participant type and target language dominance was not significant ($\beta = -.15$; $SE\ \beta = .34$; $\chi^2 < 1$). Patients omitted both dominant language targets (M = 0.70% vs. 0.38%; $\beta = .69$; $SE\ \beta = .28$; χ^2 (1) = 5.41, $p = .020$) and non-dominant language targets (M = 0.79% vs. 0.42%; $\beta = .99$; $SE\ \beta = .33$; χ^2 (1) = 9.13, $p = .003$), significantly more often than controls, and both participant groups omitted targets equally often in their two languages (χ^2 s < 1).

Self-correction rates—Unlike the analyses reported above, bilinguals with AD and controls did not differ significantly in self-correction rates of within-language errors. Patients and controls self-corrected their errors at about the same rate for disfluencies (M = 51.56% vs. 54.00%), phonological errors (M = 14.55% vs. 11.86%), function word substitution errors (M = 22.41% vs. 23.15%), inflection errors (M = 8.60% vs. 4.82%), and omissions (M = 18.39% vs. 25.67%; all ps ≥ .41). Additionally, collapsing together all the subtypes of within-language errors (including those not produced often enough for subtype analysis; see Table 2), patients and controls still self-corrected their errors about equally often (M = 28.43% vs. 28.27%, χ^2 <1).

Summary of error analyses—Summarizing the analyses of errors in the read aloud task, bilinguals produced intrusion errors when reading aloud and several within-language error subtypes, most often including disfluencies, phonological errors, function-word substitution errors, inflection errors, and omissions (in that order overall; see Table 2). Bilinguals with AD produced significantly more errors than controls for almost all error types, including intrusions in mixed-language (but not single-language) paragraphs, and most within-language error subtypes with the exception of disfluencies (and only in the dominant language for phonological errors and inflection errors). The higher rate of intrusions in bilinguals with AD appeared to reflect a monitoring deficit since patients self-corrected their intrusion errors significantly less often than controls, whereas patients and controls did not differ in self-correction rates of within-language errors.

With the exception of omission errors, which both patients and controls committed equally often in their two languages, all other within-language error subtypes exhibited highly robust language dominance effects, such that both patients and controls produced significantly more errors in their non-dominant language than in their dominant language. By contrast, dominance effects for intrusion errors were either not significant (in controls), or exhibited a fully reversed pattern (at switch-out points in bilinguals with AD). Finally, and similar to findings reported by Gollan et al., 2010, dominance effects on within-language error subtypes were diminished in patients relative to controls for some error types (significantly overall, for phonological errors, and trending in this direction for inflection errors), but not for other error types such as disfluency errors and function word substitution errors, the latter of which exhibited marginal trends in the opposite direction.

Reading Times

Although our primary measure of interest in the read aloud task was production of speech errors, it was important to also consider possible differences between groups in reading times to determine if this could be influencing any of the observed error effects. Table 3 shows reading time by participant group in each condition. Analysis of reading times was conducted using linear mixed-effects models in the R package lme4 (Bates et al., 2015). Contrast coded fixed effects included paragraph type (single-language, mixed-language), participant type (patients, controls), primary language of the paragraph (dominant, non-dominant), and all interactions between these factors. Subjects and individual paragraphs were entered as random intercepts. The significance of each fixed effect was assessed via likelihood ratio tests (Barr et al., 2013).

Bilinguals read paragraphs written primarily or exclusively in their dominant language significantly more quickly than those written primarily or exclusively in their non-dominant language (M = 53.2s vs. 72.7s; $\beta = -19.28$; $SE\,\beta = 2.77$; $\chi^2\,(1) = 25.87$, $p < .001$). Bilinguals tended to read mixed-language paragraphs more slowly than single-language paragraphs, but this difference was only marginally significant (M = 64.8s vs. 61.0s; $\beta = 2.52$; $SE\,\beta = 1.30$; $\chi^2\,(1) = 3.57$, $p = .059$). Patients tended to read paragraphs a bit more slowly than controls, but this difference was not significant (M = 65.8s vs. 61.1s; $\beta = 4.76$; $SE\,\beta = 5.41$; $\chi^2 < 1$). None of the interactions between factors were significant (χ^2s < 1).

Receiver Operating Characteristic (ROC) Curves

To consider the potential diagnostic utility of speech errors, we submitted all error types that revealed a significant difference between patients with AD and controls in the mixed effects models to ROC curve analyses. In two cases (i.e., function word substitutions and omissions) patients differed significantly from controls in both the dominant and non-dominant languages. In those cases, we conducted the following three analyses. First, we maximized power by combining all such errors produced in both languages. Second, we asked if such errors in the dominant-language could successfully discriminate patients from controls. If so, these measures might also be effective in monolinguals. Third, we asked if such errors in the non-dominant language could successfully discriminate patients from controls. If so, these measures might represent a unique tool for discriminating bilingual patients with AD from bilingual controls.

Table 4 shows the results of these analyses and Figure 5 shows plotted ROC curves for the four error types that were most effective at discriminating between patients and controls. The area under the curve (AUC), sensitivity, specificity, and ideal hypothetical cut-off scores for predicting group membership were calculated using receiver operating characteristic curve (ROC) analyses in SPSS 24. The AUC measure provides an overall indication of the diagnostic accuracy for each measure; generally classifying AUC values of 0.9 or higher as excellent and above 0.8 as good (values closer to .50 are at chance for discriminating between groups). Sensitivity (i.e., the true positive rate) and specificity (i.e., the true negative rate) were calculated and sensitivity was plotted as a function of specificity. Thus, the ROC curves illustrate diagnostic accuracy for all possible cut-off scores (from which an optimal cut-off can then be determined). In addition to the AUC, the Youden index, another commonly used ROC curve summary measure, was used to determine the rank order of the effectiveness of the predictors. Cutoff values shown in Table 4 were calculated by SPSS using a function that assumes the smallest cutoff value is the minimum observed test value minus 1, the largest cutoff value is the maximum observed test value plus 1, and all other cutoff values are the averages of two consecutive ordered observed test values. Because these values do not correspond to actual possible cutoff scores in the read-aloud task (e.g., errors were produced in whole numbers but the SPSS generated cutoffs include half numbers), for the four measures illustrated in Figure 5 we also report optimal cutoff scores where the sum of sensitivity and specificity was highest (to maximize both sensitivity and specificity; e.g., Salmon et al., 2002).

Intrusions—Intrusions at dominant language targets (collapsing all those produced at switch-out, switch-back, and non-switch points) provided excellent diagnostic accuracy (AUC = .92, $p < .001$) with maximum sensitivity and specificity of 1 and .68, respectively (cut-off value = 0.5; see Figure 5A). Actual scores ranged from 0–1 in controls, and 1–3 in patients, with an optimal cut-off at 1 intrusion or more giving 100% sensitivity and 68% specificity. The next two predictors were equally effective and provided good diagnostic accuracy: 1) intrusions at dominant language targets at switch-out points and 2) all intrusions produced (collapsing targets in both languages, and all points in the paragraphs). These two measures provided slightly better specificity than did intrusions at dominant language targets, but lower sensitivity (see Table 4). All intrusions produced ranged from 0–7 in controls, and 1–7 in patients, with an optimal cut-off at 2 or more giving 83% sensitivity and 79% specificity.

Within-language errors—Combining all within-language errors produced (including all types shown in Table 2) resulted in an AUC of .71, falling short of the criterion for a good predictor. However, dominant language inflection errors provided good discrimination (see Table 4, Figure 5B). Another good predictor was function-word substitution errors collapsed across languages (see Table 4, Figure 5C). Actual scores for this measure ranged from 1–33 in controls, and 10–37 in patients, with an optimal cut-off at 14 substitution errors or more giving 92% sensitivity and 63% specificity. Note that when using the Youden index instead of the AUC, function word substitutions separated by language dominance provided better discrimination than when the languages were combined. Omission errors overall (collapsed across language dominance) showed fair diagnostic accuracy, and similarly, omission errors in the non-dominant language showed fair diagnostic accuracy (see Table 4). However, omissions in the dominant language showed poor diagnostic accuracy and did not reach statistical significance (AUC = .67, $p = .11$). Phonological errors in the dominant language also did not reach statistical significance (AUC = .70, $p = .06$).[6]

---

[6] When excluding one patient whose self-reported language dominance did not match with objectively determined language dominance (i.e., MINT scores), the general pattern of results and rank orders of AUC values remained the same for all intrusion measures and for most other error types with a couple of minor exceptions. Dominant language inflections and function-word

Finally, we combined our best measures for discriminating patients from controls to determine if this could maximize discrimination even further. For this combined measure we summed dominant intrusions, dominant inflections, and all function word substitution errors. This measure did not improve discrimination over our top measure; it ranked second best (AUC = .86, $p = .001$), and provided sensitivity and specificity of 92% and 68%, respectively. We also attempted to find a measure that might work in monolinguals by combining top performing dominant-language measures; however, none of these performed better than dominant language inflections alone (shown in Table 4) so we do not report details of these analyses here.

Discussion

The present study was designed to investigate how AD affects production of bilingual speech errors produced in a read aloud task. Bilinguals with AD produced between-language intrusion errors and various types of within-language errors at a significantly higher rate than controls. Though bilinguals produced significantly more errors in the less frequently used, and less proficient, non-dominant language, there was no consistent evidence to suggest that patients had more difficulty than controls specifically with producing the non-dominant language. Intrusion errors exhibited trends in the opposite direction such that words in the dominant language were more often targets of error than non-dominant language targets (but see below). Function word substitution errors exhibited a marginally significant trend in the direction of larger patient versus control difference in the non-dominant language, and disfluency errors were also in this direction (though the interaction between participant type and language was not significant). By contrast, differences between patients and controls were significantly larger in the dominant than in the non-dominant language for within-language errors overall, and this interaction was also significant for phonological errors alone. The ROC curves (see Table 4) also did not exhibit a consistent pattern; sometimes implying better discrimination between patients and controls using errors bilinguals produced in the dominant, and for other error types the non-dominant language. Together these results have implications for understanding bilingual language control, linguistic impairments in AD, and paint a more mixed picture than previously reported (Gollan et al., 2010) for how to best discriminate bilingual patients from controls.

Why do Bilinguals with AD Produce More Intrusion Errors than Controls?

Because dominant-language targets produced the greatest difference between patients and controls in production of intrusion errors, it could be argued that the results we obtained with intrusion errors are generally consistent with the counterintuitive pattern of linguistic deficits in bilingual AD that we originally reported with a picture-naming task (i.e., Gollan et al., 2010; Ivanova et al., 2013; and see also Multilingual Naming Test results in Table 1). However, different accounts of the findings are likely needed to explain AD effects on picture naming versus intrusions in the read-aloud task. Though patients replaced dominant-language target words with their non-dominant language translation equivalents more often

---

substitutions all ranked 4th and 3rd respectively (instead of 3rd and 4th; see Table 4), and dominant language phonological errors and dominant language omission errors ranked 9th and 8th (instead of 8th and 9th; see Table 4).

than the reverse, these intrusion errors were produced in the context of reading aloud paragraphs that were *written primarily in the non-dominant language*. Critically, bilinguals with AD did not produce more intrusion errors than controls when reading aloud paragraphs written *exclusively* in the non-dominant (or dominant) language – indeed both patients and controls produced very few intrusions in paragraphs with no language mixing. Thus, in the read aloud task, deficits in bilinguals with AD were apparent only in conditions that placed greater demands on language control mechanisms. Specifically, when bilinguals needed to occasionally switch into the dominant language when producing sentences written primarily in the non-dominant language.

In this respect, the results we report for intrusion errors do not resemble those reported for picture-naming, although in both cases dominant language targets appeared to be most sensitive to patient versus control differences. In the read aloud task, paragraph types that elicited the highest rates of intrusion errors (non-dominant language with a small number of dominant language targets) also elicited the greatest difference between participant groups (whereas conditions with higher error rates did not elicit greater differences between groups in picture naming; Gollan et al., 2010). Importantly, though controls did not produce more intrusions with dominant than non-dominant language targets in the present study, we have previously reported this condition to be most error prone in both young and older bilinguals (Gollan et al., 2014; Gollan & Goldrick, 2016, in press). Indeed in the present study, controls simply did not produce many intrusion errors, even at switch-out points. This was likely because all switch words in the present study were content words (e.g., nouns, verbs, adjectives), whereas in previous work switches on function words (e.g., pronouns, articles, prepositions, and conjunctions) elicited the majority of intrusion errors.

In light of these considerations, it seems likely that the increased production of intrusion errors in bilinguals with AD relative to controls in the read aloud task reflects deficits in executive control (e.g., Perry & Hodges, 1999), possibly with switching languages, or more likely with a monitoring process that ensures that selected lexical items in planned speech match the intended language of production. Consistent with this view, patients self-corrected their intrusion errors at a significantly lower rate than controls. The distinction between switching and monitoring may be needed because of the contrast between switch-out points, which elicited the majority of intrusions and revealed the biggest difference between bilinguals with AD and controls, versus switch-back points which rarely elicited intrusions in patients or controls (see Figure 1; and for similar findings in healthy young bilinguals see Gollan & Goldrick, in press). Thus, it would seem that switching languages per se is not necessarily difficult, rather what is difficult is switching out of the default language.

The contrast between switches out of and switches back into the paragraph's default language could resemble processes involved when bilinguals spontaneously produce mixed-language speech (Myers-Scotton, 1993, 1997; Myers-Scotton & Jake, 2009), possibly involving inhibitory control (Myers-Scotton, 2006) of whichever language is not selected as the main force driving the utterance, i.e., the default language. In previous studies, we explained the finding of reversed language dominance effects on intrusion errors by invoking inhibitory control of the dominant language (Gollan et al., 2014; Gollan & Goldrick, 2016, in press). When reading mixed-language paragraphs, particularly paragraphs written primarily in the non-dominant language, bilinguals partially inhibit the dominant language, leading dominant language targets to be replaced with non-dominant language translation equivalents more often than the reverse. Possibly similar inhibitory control processes might explain results why the dominant language sometimes exhibits larger switch-costs than the non-dominant language, though such effects might reflect transient rather than sustained control mechanisms (Bobb & Wodniecka, 2013; Declerck & Philipp, 2015a; Green, 1998; Meuter & Allport, 1999; Kroll, et al., 2008; Philipp, Gade, & Koch, 2007; Philipp & Koch, 2009). The finding of significantly reversed language-dominance effects on intrusion errors in patients but not in controls in the present study has important implications for understanding the cognitive mechanisms underlying default language selection in bilingual language production. In particular, default language selection appears to be relatively automatic, perhaps aided by processing of context (each paragraph was written mostly in one or the other language), which could remain relatively intact in AD. In this respect, the inhibitory control mechanisms involved in default language selection appear to differ from inhibitory control as sometimes assessed in linguistic tasks that are often impaired in AD, such as the Hayling test (in which speakers try, often unsuccessfully, to avoid completing a sentence with a predictable target such as *Cats see well at* ______; Belleville, Rouleau, & Van der Linden, 2006). However, additional work is needed to explore the cognitive mechanisms underlying reversed dominance effects in the read-aloud task; default language selection, if aided by grammatical encoding should affect function more than content word retrieval (Gollan & Goldrick, in press), and intrusions in the present study primarily involved content words (as switch-out points were content words only).

An alternative possibility is that different cognitive processes control application versus release of inhibition, with release mechanisms being more impaired in AD than application mechanisms (Ivanova, Montoya, Murillo, & Gollan, 2016), or that bilinguals with AD over-applied inhibition of the dominant language. Indeed, the near total absence of intrusion errors in single-language paragraphs, contrasted with patterns found in mixed-language paragraphs, invites the conclusion that steady application of inhibition is less demanding than rapidly modulating the degree of inhibition between consecutive words (for related discussion see Gollan & Ferreira, 2009). These accounts differ considerably from those proposed previously to explain deficits in picture-naming in AD (Gollan et al., 2010; Ivanova et al, 2013), but – as noted above – different tasks might reveal different cognitive deficits in AD. The read-aloud task provides very explicit cues to the intended targets (i.e., written words). It is possible that bilinguals' relatively increased difficulty with picture-naming in the non-dominant language with disease progression (Ivanova et al, 2013) also reflects increased difficulty with controlling (inhibiting) the dominant language when facing competition for selection between languages in production without explicit retrieval cues from written words. On this view, retrieval difficulty might explain patterns observed when comparing patients to controls in picture naming (or other language selective tasks), while executive control explains deficits found in picture naming with disease progression. By contrast, the read-aloud task with a switching manipulation specifically targets language control mechanisms (in a manner that single-language paragraphs and picture-naming simply did not) thereby revealing control deficits in bilinguals with AD.

A final bit of evidence that seems consistent with the hypothesis that increased production of intrusion errors in bilinguals with AD reflects a deficit in linguistic control processes comes from correlations with neuropsychological tests. Tables 5a and 5b illustrate these correlations for cognitively healthy bilinguals and patients, respectively. Bilinguals with AD who produced more intrusion errors also tended to have relatively weaker letter fluency scores ($r = -.611$, $p = .04$), a relationship that could implicate executive control in production of intrusions since letter fluency strongly engages executive control (e.g., Baldo et al., 2006; Birn et al., 2010; Costafreda et al., 2006; Martin et al., 1994; Monsch, 1994; Mummery et al., 1996, but see Shao et al., 2014). Patients who produced many intrusion errors also produced more function word substitution errors ($r = .643$, $p = .02$) and phonological errors ($r = .647$, $p = .02$). Though these correlations were not found in cognitively healthy bilinguals, many controls produced no intrusion errors (paragraphs had switches only on content words which elicit few intrusions in cognitively healthy bilinguals; Gollan et al., 2014; Gollan & Goldrick, 2016, in press). By contrast, both patients and controls exhibited significant correlations between those subtypes of within-language speech errors that were produced in large numbers by participants in both groups (e.g., phonological errors and function word substitution errors, rs ≥ .757, ps < .01). However, the observed correlations must be viewed as strictly exploratory (given the small number of participants tested and low rates of intrusion errors); additional work is needed to test the possibility that language intrusion errors present the executive control system with a particularly difficult challenge relative to other types of speech errors.

Can Speech-Errors Discriminate Between Bilinguals with AD and Controls?

Of key interest in the present study was if the read aloud task might elicit speech errors that can distinguish bilinguals with AD from controls. Though the results must be considered with caution given the relatively small number of participants in the present study, production of intrusion errors with dominant language targets when reading paragraphs written primarily in the non-dominant language appeared to provide an excellent diagnostic measure – the best of all types of speech errors (see Table 4, and Figure 3) even though intrusion errors were one of the least frequently produced error types overall (see Table 2). It is remarkable that just four to five minutes of reading-aloud (two paragraphs of each type; see Table 3 for reading times), each with just 6 language switches in each paragraph, was sufficient to yield excellent discrimination between patients and controls. Intrusion errors might have been more sensitive to patient-control differences than other error types produced in the present study because paragraphs were designed specifically to elicit such errors (by switching languages). If paragraphs could be manipulated to specifically elicit production of other error subtypes, it is possible they too would provide excellent discrimination between patients and controls. Additional investigation is needed to discover what such manipulation(s) might entail, and whether or not within-language speech errors in the read aloud task, which could easily be measured in monolingual speakers, might generally be useful for diagnosis of AD.

Further work might also reveal if switches on function words rather than content words could improve discrimination between patients and controls. In the present study, we built in switches exclusively on content words because cognitively intact bilinguals rarely have difficulty with such switches (Gollan et al., 2014; Gollan & Goldrick, 2016, in press) and therefore high intrusion rates with such targets might be diagnostic for AD. Consistent with previous findings, when bilinguals in the present study produced intrusions outside of switch-out points, these often involved function words. Of highest priority in pursuit of determining how to maximize the potential of the read aloud task for eliciting speech errors that might discriminate between patients and controls would be to test a greater number of participants, and to examine possible interactions between participant type and language dominance using fully counterbalanced designs. We used a fixed-order in the present study (i.e., dominant-only, non-dominant-only, dominant-mixed, non-dominant-mixed) because no previous studies had examined paragraph reading in bilinguals with AD, and we wanted to bias any possible practice effects towards the non-dominant language (the more difficult task). Practice effects, if any, seemed to have little effect in the present study given that language dominance effects were highly robust in most analyses (with the dominant language less error- prone even though it was administered prior to the non-dominant language and would therefore have benefitted less from any practice effects). Nevertheless, it would increase confidence in the patterns observed here if similar results were obtained using a fully counterbalanced design. Another high priority is to test monolingual patients and controls on the read aloud task given the observed sensitivity of within-language errors to AD.

The present set of results also lead us to modify previously drawn conclusions that bilinguals should be tested primarily in the dominant language for the purpose of discriminating patients from controls. Previously we proposed that the dominant language is more sensitive to subtle differences between patients and controls that emerge in early stages of AD possibly because AD impairs semantic processing (reviewed above; see also Butters et al., 1987; Monsch et al., 1994; Salmon, Butters, & Chan, 1999) and the dominant language has richer semantic representations (Gollan et al., 2010), or because AD produces a general retrieval deficit and the most difficult-to-retrieve words that bilinguals know belong to the dominant language (Ivanova et al., 2014). The semantic account is challenged by finding that bilinguals with AD produced more intrusion errors than controls. In the read aloud task, intrusion errors illustrate a strikingly intact ability to (a) very rapidly retrieve the meanings of written target words while (b) spontaneously and rapidly translating the written words into the other language so that they are spoken as translation equivalents of the words written on the page. Like picture naming, translation is largely a semantically driven process, particularly when translating dominant language words into the non-dominant language (Kroll & Stewart, 1994; Kroll, Van Hell, Tokowicz, & Green, 2010). Translation in the other direction, non-dominant into dominant, may be less semantically driven, and more automatic involving lexical level connections between translations. If impaired semantic processing drove production of intrusion errors in the read aloud task, then bilinguals with AD should have been less likely than controls to produce reversed dominance effects that require rapidly translating dominant language words into the non-dominant language. Instead patients tended to produce more intrusions than controls in both the dominant and non-dominant languages, and if anything this difference was more robust for the semantic translation direction (dominant replaced by non-dominant). Other error subtypes are even more difficult to explain via appeal to semantic processing. Reading aloud requires many cognitive skills and different types of errors produced could reflect a variety of different processes unlikely to involve semantic processing. For example, function words are impoverished in their semantic content compared with content words, thus function-word substitution errors likely involve grammatical encoding mechanisms, and it is not clear how semantic processing deficits could elicit such errors. Similarly, in phonological errors, target words are often just slightly distorted and mispronounced, likely reflecting difficulties with phonological encoding that arises after target words have already been selected and semantic processing is relatively complete. The observed sensitivity of these error types to AD implies the presence of broader linguistic impairments in AD.

The retrieval deficit hypothesis also faces some possible challenges in explaining the pattern of deficits reported here for bilinguals with AD. As noted above, the generally more effortful task (i.e., reading in the non-dominant language with switches to the dominant language) elicited the largest patient-control differences consistent with the notion of deficits in effortful retrieval. However, as discussed above, patterns were less consistent across subtypes of within-language errors. Language dominance effects were highly robust (see Figure 4) even though the paragraphs contained excerpts of relatively simple short stories that could easily be read by bilinguals with varying proficiency levels in either language, were not graded in difficulty, and (unlike naming tests that become progressively more difficult) we made no effort to push the limits of competence in text comprehension. Similarly, errors that were produced in greater numbers overall (thereby possibly reflecting a more difficult linguistic process) did not necessarily produce greater differences between patients and controls. Instead, some of the more rarely produced error types appeared to be best for discriminating between groups, including intrusion errors, function word substitution errors, inflection errors, and omissions (e.g., compare *n*s in Table 2 which shows error types, and Table 4 which shows results of ROC analyses). In both bilinguals and monolinguals an important avenue to consider in future work will be whether conditions that elicit very few errors and little variability in performance in controls might be best for discriminating patients from controls (see Figure 2; error bars are typically smallest for controls in the dominant language).

The only error type that did not exhibit robust language dominance effects was omission errors. This was true even though omission errors were sensitive to differences between patients and controls in both languages. Of further possible interest, Table 5a reveals that cognitively healthy bilinguals who produced more omission errors also produced more function word substitution errors ($r = .713$, $p < .001$) and phonological errors ($r = .757$, $p < .001$), and also had weaker letter fluency scores ($r = -.657$, $p = .002$). The latter result could imply that sensitivity of AD to omission errors also reflects deficits in executive control mechanisms.

Neurocognitive Mechanisms Underlying Production of Speech Errors in Bilinguals with AD

Our observation of special sensitivity of intrusion errors in bilinguals with AD could imply that bilingual language control, and possibly managing competition between translation equivalents in bilingual speech production, poses a particular challenge to brain mechanisms that support executive control. This is consistent with proposals that AD disrupts functional connectivity in brain regions associated with executive control (e.g., dorsolateral prefrontal cortex; Agosta et al., 2012; Weiler et al., 2014; see review in Brier, Thomas, & Ances, 2014). Based primarily on neuroimaging studies of language switching (largely without sentence context), Abutalebi and Green (2007; 2008; Green & Abutalebi, 2013; see also Grundy, Anderson, & Bialystok, 2017; Luk, Green, Abutalebi, & Grady, 2012) proposed that inhibition is a critical cognitive mechanism needed to maintain control over bilingual language selection, and defined a Language Control Network comprising the dorsal anterior cingulate cortex (dACC), left caudate nucleus (LCN), superior marginal gyrus (SMG), dorsolateral prefrontal cortex (DLPFC), and inferior frontal gyrus (IFG). These and adjacent regions are also recruited for nonlinguistic control tasks (De Baene, Duyck, Brass & Carreiras, 2015; Garbin et al., 2010; Weissberger, Gollan, Bondi, Clark, & Wierenga, 2015), are known to support monitoring and switching (e.g., Gurd et al., 2002; Hedden & Gabrieli, 2001), and are often impaired in AD (Faust, Balota, Ducheck, & Gernsbacher, 1997; Hutchison, Balota, & Duchek, 2010; Troyer et al., 1997). Our interpretation of the error patterns observed above implies that these same brain regions can function relatively well in AD when supported by context, and fail only when taxed by the most difficult control tasks (i.e., inhibition of the dominant language was spared, but the ability to toggle inhibition on and off rapidly to allow production of a few dominant language targets embedded in connected non-dominant language speech was impaired).

A caveat that will need to be considered with further investigation is whether the special sensitivity of intrusions to AD indeed reflects a primary deficit in the functioning of the executive control network, or if the results are specific to the task we used in the present study. It is possible that different types of speech errors could be more sensitive to AD if a different method for eliciting them were discovered (as noted above, the manipulation tested here was designed specifically to elicit intrusions, not other types of speech errors). A recent study provides evidence for the hypothesis of executive function involvement in the read aloud task; bilinguals with mild traumatic brain injury (mTBI) exhibited higher rates of intrusion errors, aberrant patterns of eye-movements, and significant correlations between these measures and performance on executive function tasks implying an important role for the frontal lobes in the read aloud task (Ratiu & Azuma, 2017). Additionally, performance on a combined flanker-go/no-go task and a non-linguistic switching task was correlated with production of both cross-language intrusion errors and within-language errors (Ratiu & Azuma, 2017; see also Table 5a), which suggests a broader role for executive control in controlling speech production – which could even operate in monolinguals.

A promising avenue to explore in the attempt to understand why intrusion errors appeared uniquely sensitive to AD is the notable contrast we observed between switches back into (not impaired in AD) versus switches out of (impaired in AD) the default language (see Gollan & Goldrick, in press for discussion of the distinction between switches out and switches back). It might also be useful to investigate similarity in phonological form between translation equivalents (cognate effects; Gollan et al., 2014), part of speech effects (Gollan & Goldrick, 2016), and syntactic influences on the ease with which bilinguals can switch. It is known, for example, that switches are easier when the participating languages share word order in a sentence (Dussias; Declerck & Philipp, 2015b), and that certain combinatorial patterns are preferred over others in spontaneous code-switches (Tamargo, Valdés Kroff, & Dussias, 2016). Such factors might be differentially sensitive to language control deficits in AD. Studies examining these factors may not only reveal which types of errors are most sensitive to AD, but also reveal cognitive mechanisms underlying bilingual language processing and production of speech errors in general.

Conclusions

Returning to the questions posed earlier, what is the nature of linguistic impairments in bilinguals with AD, and can speech errors produced in the read aloud task discriminate between patients and controls? The answer to the first question is that linguistic impairments appear to be broad and varied. Patients with AD produced more between-language errors and several types of within-language errors compared to controls. We suggested that production of intrusion errors in bilinguals with AD reflects impaired control processes, though additional work is needed to verify these claims and to pinpoint more specifically which processes are impaired (application versus release of inhibition) and what types of processing underlie increased production of other error subtypes in AD. The answer to the second question also appears to be yes, but further work is needed to determine how to maximize the potential of the read aloud task for discriminating between cognitively intact individuals and those with AD. Finally, the results of the present study call for modification of our previous recommendations to test bilinguals primarily in the dominant language when the aim is to discriminate bilinguals with AD from cognitively healthy bilinguals. Intrusion errors appeared to be most useful for this purpose but were produced primarily in paragraphs that were written mostly in the non-dominant language – a finding that could be viewed as counter to our previous suggestion.

Acknowledgments

The authors thank John Wixted for helpful discussion, Rosa Montoya and Mayra Murillo for composition of the paragraphs and error coding, and Martha Muñiz for assistance with data collection.

Funding

This research was supported by grants from the National Institute on Deafness and Other Communication Disorders (011492), by a Graduate Research Fellowship Program (GRFP) and a grant (1457159) from the National Science Foundation, and by a P50 (AG05131) from the National Institute of Aging to the University of California. 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 NIH.

References

Abutalebi J, Green D. Control mechanisms in bilingual language production: Neural evidence from language switching studies. Language and Cognitive Processes. 2008; 23:557–582.

Abutalebi J, Green DW. Bilingual language production: The neurocognition of language representation and control. Journal of Neurolinguistics. 2007; 20:242–275.

Adlam AL, Patterson K, Rogers TT, Nestor PJ, Salmond CH, Acosta-Cabronero J, et al. Semantic dementia and fluent primary progressive aphasia: Two sides of the same coin? Brain. 2006; 129:3066–3080. [PubMed: 17071925]

Agosta F, Pievani M, Geroldi C, Copetti M, Frisoni GB, Filippi M. Resting state fMRI in Alzheimer's disease: beyond the default mode network. Neurobiology of Aging. 2012; 33:1564–1578. [PubMed: 21813210]

Almor A, Kempler D, MacDonald MC, Andersen ES, Tyler LK. Why do Alzheimer patients have difficulty with pronouns? Working memory, semantics, and reference in comprehension and production in Alzheimer’s disease. Brain and Language. 1999; 67:202–227. [PubMed: 10210631]

Altmann LJ, Kempler D, Andersen ES. Speech errors in Alzheimer’s disease: Reevaluating morphosyntactic preservation. Journal of Speech, Language, and Hearing Research. 2001; 44(5): 1069–1082.

Baldo JV, Schwartz S, Wilkins D, Dronkers NF. Role of frontal versus temporal cortex in verbal fluency as revealed by voxel-based lesion symptom mapping. Journal of the International Neuropsychological Society. 2006; 12:896–900. [PubMed: 17064451]

Birn RM, Kenworthy L, Case L, Caravella R, Jones TB, Bandettini PA, Martin A. Neural systems supporting lexical search guided by letter and semantic category cues: A self-paced overt response fMRI study of verbal fluency. Neuroimage. 2010; 49:1099–1107. [PubMed: 19632335]

Balota DA, Watson JM, Duchek JM, Ferraro FR. Cross-modal semantic and homograph priming in healthy young, healthy old, and in Alzheimer's disease individuals. Journal of the International Neuropsychological Society. 1999; 5:626–640. [PubMed: 10645705]

Baudic S, Dalla Barba G, Thibaudet MC, Smagghe A, Remy R, Traykov L. Executive function deficits in early Alzheimer’s disease and their relations with episodic memory. Archives of Clinical Neuropsychology. 2006; 21:15–21. [PubMed: 16125364]

Barr DJ, Levy R, Scheepers C, Tily HJ. Random effects structure for confirmatory hypothesis testing: Keep it maximal. Journal of Memory and Language. 2013; 68:255–278.

Bates D, Maechler M, Bolker B, Walker S. Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software. 2015; 67(1):1–48. DOI: 10.18637/jss.v067.i01

Bayles KA, Tomoeda CK, Trosset W. Relation of linguistic abilities of Alzheimer’s patients to stage of disease. Brain and Language. 1992; 42:454–472. [PubMed: 1377076]

Belleville S, Rouleau N, Van der Linden M. Use of the Hayling task to measure inhibition of prepotent responses in normal aging and Alzheimer's disease. Brain and Cognition. 2006; 62(2):113–119. [PubMed: 16757076]

Bobb SC, Wodniecka Z. Language switching in picture naming: What asymmetric switch costs (do not) tell us about inhibition in bilingual speech planning. Journal of Cognitive Psycholology. 2013; 25:568–585.

Brier MR, Thomas JB, Ances BM. Network dysfunction in Alzheimer’s Disease: Refining the Disconnection Hypothesis. Brain Connectivity. 2014; 4:299–311. [PubMed: 24796856]

Butters N, Granholm E, Salmon DP, Grant I, Wolfe J. Episodic and semantic memory: A comparison of amnesic and demented patients. Journal of Clinical and Experimental Neuropsychology. 1987; 9(5):479–497. [PubMed: 2959682]

Chan AS, Salmon DP, DeLaPena J. Conceptual and perceptual contributions to the acquisition and retention of text-specific reading skill in Alzheimer's disease. Journal of Clinical and Experimental Neuropsychology. 1999; 21:325–340. [PubMed: 10474172]

Christoffels I, Firk C, Schiller NO. Bilingual language control: An event related brain potential study. Brain Research. 2007; 1147:192–208. [PubMed: 17391649]

Costa A, Calabria M, Marne P, Hernández M, Juncadella M, Gascón-Bayarri J, Reñé R. On the parallel deterioration of lexico-semantic processes in the bilinguals' two languages: evidence from Alzheimer's disease. Neuropsychologia. 2012; 50:740–753. [PubMed: 22261401]

Costa A, Santesteban M. Lexical access in bilingual speech production: Evidence from language switching in highly proficient bilinguals and L2 learners. Journal of Memory and Language. 2004; 50:491–511.

Costafreda SG, Fu CHY, Lee L, Everitt B, Brammer MJ, David AS. A systematic review and quantitative appraisal of fMRI studies of verbal fluency: Role of the Left Inferior Frontal Gyrus. Human Brain Mapping. 2006; 27:799–810. [PubMed: 16511886]

Croot K, Hodges JR, Xuereb J, Patterson K. Phonological and articulatory impairment in Alzheimer’s disease: a case series. Brain and Language. 2000; 75:277–309. [PubMed: 11049669]

Cummings JL, Houlihan JP, Hill MA. The pattern of reading deterioration in dementia of the Alzheimer type: Observations and implications. Brain and Language. 1986; 29:315–323. [PubMed: 3790984]

De Baene W, Duyck W, Brass M, Carreiras M. Brain circuit for cognitive control is shared by task and language switching. Journal of Cognitive Neuroscience. 2015; 27:1752–1765. [PubMed: 25901448]

Declerck M, Philipp AM. A review of control processes and their locus in language switching. Psychonomics Bulletin & Review. 2015a; 22:1630–1645. DOI: 10.3758/s13423-015-0836-1

Declerck M, Philipp AM. A sentence to remember: Instructed language switching in sentence production. Cognition. 2015b; 137:166–173. DOI: 10.1016/j.cognition.2015.01.006 [PubMed: 25659539]

Dixon P. Models of accuracy in repeated-measures designs. Journal of Memory and Language. 2008; 59:447–456.

Faust ME, Balota DA, Duchek JM, Gernsbacher MA, Smith S. Inhibitory control during sentence comprehension in individuals with dementia of the Alzheimer type. Brain and Language. 1997; 57:225–253. [PubMed: 9126415]

Forbes-McKay K, Shanks MF, Venneri A. Profiling spontaneous speech decline in Alzheimer’s disease: a longitudinal study. Acta Neuropsychiatrica. 2013; 25:320–327. [PubMed: 25287871]

Friedman RB, Ferguson S, Robinson S, Sunderland T. Dissociation of mechanisms of reading in Alzheimer's disease. Brain and Language. 1992; 43:400–413. [PubMed: 1446210]

Fromkin VA. The non-anomalous nature of anomalous utterances. Language. 1971; 47:27–52.

Garbin G, Sanjuan A, Forn C, Bustamante JC, Rodriguez-Pujadas A, Belloch V, Hernandez M, Costa A, Avila C. Bridging language and attention: brain basis of the impact of bilingualism on cognitive control. NeuroImage. 2010; 53:1272–1278. [PubMed: 20558314]

Garrett, MF. The analysis of sentence production. In: Bower, GH., editor. The psychology of learning and motivation. Vol. 9. New York: Academic Press; 1975. p. 133-177.

Garrett, MF. Production of speech: Observations from normal and pathological language use. In: Ellis, A., editor. Normality and pathology in cognitive functions. London, England: Academic Press; 1982. p. 19-76.

Glosser, G., Grossman, M. Vision in Alzheimer's Disease. Karger Publishers; 2004. Reading and Visual Processing in Alzheimer's Disease; p. 236-247.

Gollan TH, Ferreira VS. 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, and Cognition. 2009; 35:640–665.

Gollan TH, Goldrick M. A Switch is Not a Switch: Syntactically-Driven Bilingual Language Control. Journal of Experimental Psychology: Learning, Memory, & Cognition. (in press).

Gollan TH, Goldrick M. Grammatical constraints on language switching: Language control is not just executive control. Journal of Memory and Language. 2016; 90:177–199. [PubMed: 27667899]

Gollan TH, Salmon DP, Montoya RI, Da Pena E. Accessibility of the non-dominant language in picture naming: A counterintuitive effect of dementia on bilingual language production. Neuropsychologia. 2010; 48:1356–1366. [PubMed: 20036679]

Gollan TH, Sandoval T, Salmon DP. Cross-language intrusion errors in aging bilinguals reveal the link between executive control and language selection. Psychological Science. 2011; 22:1155–1164. [PubMed: 21775653]

Gollan TH, Schotter ER, Gomez J, Murillo M, Rayner K. Multiple Levels of Bilingual Language Control: Evidence From Language Intrusions in Reading Aloud. Psychological Science. 2014; 25:585–595. [PubMed: 24367061]

Gollan TH, Weissberger G, Runnqvist E, Montoya RI, Cera CM. 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. 2012; 15:594–615.

Green DW. Mental control of the bilingual lexicosemantic system. Bilingualism: Language and Cognition. 1998; 1:67–81.

Green DW, Abutalebi J. Language control in bilinguals: the adaptive control hypothesis. Journal of Cognitive Psychology. 2013; 25:515–530. [PubMed: 25077013]

Grundy JG, Anderson JAE, Bialystok E. Neural Correlates of Cognitive Processing in Monolinguals and Bilinguals. Annals of the New York Academy of Sciences: The Year in Review. 2017; doi: 10.1111/nyas.13333

Gurd JM, Amunts K, Weiss PH, Zafiris O, Zilles K, Marshall JC, Fink GR. Posterior parietal cortex is implicated in continuous switching between verbal fluency tasks: an fMRI study with clinical implications. Brain. 2002; 125:1024–1038. [PubMed: 11960893]

Hedden T, Gabrieli JDE. Shared and selective neural correlates of inhibition, facilitation, and shifting processes during executive control. Neuroimage. 2010; 51:421–431. [PubMed: 20123030]

Hodges JR, Patterson K. Is semantic memory consistently impaired early in the course of Alzheimer’s disease? Neuroanatomical and diagnostic implications. Neuropsychologia. 1995; 33:441–459. [PubMed: 7617154]

Hodges JR, Salmon DP, Butters N. Semantic memory impairment in Alzheimer's disease: Failure of access or degraded knowledge? Neuropsychologia. 1992; 30:301–314. [PubMed: 1603295]

Hoffmann I, Nemeth D, Dye CD, Pákáski M, Irinyi T, Kálmán J. Temporal parameters of spontaneous speech in Alzheimer’s disease. International Journal of Speech-Language Pathology. 2010; 12:29–34. [PubMed: 20380247]

Hutchison KA, Balota DA, Duchek JM. The utility of Stroop task switching as a marker for early stage Alzheimer’s Disease. Psychology & Aging. 2010; 25:545–559. [PubMed: 20853964]

Ivanova I, Salmon DP, Gollan TH. The Multilingual Naming Test in Alzheimer’s disease: Clues to the origin of naming impairments. The Journal of the International Neuropsychological Society. 2013; 19:272–283. [PubMed: 23298442]

Ivanova I, Salmon DP, Gollan TH. 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. 2014; 20:534–546. [PubMed: 24725624]

Jaeger TF. Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. Journal of Memory and Language. 2008; 59:434–446. [PubMed: 19884961]

Kavé G, Levy Y. Sensitivity to gender, person, and tense inflection by persons with Alzheimer’s disease. Brain and Language. 2003; 87:267–277. [PubMed: 14585296]

Kempler D, Curtiss S, Jackson C. Syntactic preservation in Alzheimer’s disease. Journal of Speech and Hearing Research. 1987; 30:343–350. [PubMed: 3669641]

Kirshner HS, Webb WG, Kelly MP. The naming disorder of dementia. Neuropsychologia. 1984; 22:23–30. [PubMed: 6709173]

Kolers P. Reading and talking bilingually. American Journal of Psychology. 1966; 3:357–376.

Kowoll ME, Degen C, Gladis S, Schröder J. Neuropsychological profiles and verbal abilities in lifelong bilinguals with mild cognitive impairment and Alzheimer’s disease. Journal of Alzheimer’s Disease. 2015; 45:1257–1268.

Kroll JF, Bobb SC, Misra M, Guo T. Language selection in bilingual speech: Evidence for inhibitory processes. Acta Psychologica. 2008; 128:416–430. [PubMed: 18358449]

Kroll JF, Stewart E. Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. Journal of Memory and Language. 1994; 33:149–174.

Kroll JF, Van Hell JG, Tokowicz N, Green DW. The Revised Hierarchical model: A critical review and assessment. Bilingualism: Language and Cognition. 2010; 13(3):373–381.

Lafleche G, Albert MS. Executive Function Deficits in Mild Alzheimer's Disease. Neuropsychology. 1995; 9:313–320.

Luk G, Green DW, Abutalebi J, Grady C. Cognitive control of language switching in bilinguals: A quantitative meta-analysis on functional neuroimaging studies. Language and Cognitive Processes. 2012; 27:1479–1488.

MacKay DG, James LE. Sequencing, speech production, and selective effects of aging on phonological and morphological speech errors. Psychology and Aging. 2004; 19:93–107. [PubMed: 15065934]

Martin A, Wiggs CL, Lalonde F, Mack C. Word retrieval to letter and semantic cues: A double dissociation in normal subjects using interference tasks. Neuropsychologia. 1994; 32:1487–1494. [PubMed: 7885578]

Mattis, S. Dementia rating scale: Professional manual. Odessa, FL: Psychological Assessment Resources; 1988.

McGlinchey-Berroth, R., Milberg, WP. Preserved semantic memory structure in Alzheimer’s disease. In: Cerella, J.Rybash, J.Hoyer, W., Commons, ML., editors. Adult Information Processing: Limits On Loss. New York: Academic Press; 1993. p. 407-422.

McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. Clinical diagnosis of Alzheimer’s disease: Report of the NINCDS-ADRDA Work Group. Neurology. 1984; 34:939–944. [PubMed: 6610841]

Mendez MF, Perryman KM, Pontón MO, Cummings JL. Bilingualism and dementia. Journal of Neuropsychiatry and Clinical Neurosciences. 1999; 11:411–412. [PubMed: 10440021]

Meuter RFI, Allport A. Bilingual language switching in naming: Asymmetrical costs of language selection. Journal of Memory and Language. 1999; 40:25–40.

Monti LA, Gabrieli JDE, Wilson RS, Reminger SL. Intact text-specific implicit memory in patients with Alzheimer's disease. Psychology and Aging. 1994; 9:64–71. [PubMed: 8185870]

Monti LA, Gabrieli JDE, Wilson RS, Beckett LA, Grinnel E, Lange KL, Reminger SL. Sources of priming in text rereading: Intact implicit memory for new associations in older adults and in patients with Alzheimer's disease. Psychology and Aging. 1997; 12:536–547. [PubMed: 9308100]

Monsch AU, Bondi MW, Butters N, Paulsen JS, Salmon DP, Brugger P, Swenson MR. A comparison of category and letter fluency in Alzheimer's disease and Huntington's disease. Neuropsychology. 1994; 8(1):25–30.

Mummery CJ, Patterson K, Hodges JR, Wise RJ. Generating ‘tiger’ as an animal name or a word beginning with T: differences in brain activation. Proceedings of Royal Society of Biological Sciences. 1996; 263:989–995.

Murdoch BE, Chenery HJ, Wilks V, Boyle RS. Language disorders in dementia of the Alzheimer type. Brain and Language. 1987; 31:122–137. [PubMed: 2437993]

Myers-Scotton, C. Duelling languages: Grammatical structure in codeswitching. Oxford, England: Oxford University Press; 1993,1997. (Original work published 1993)

Myers-Scotton, C., Jake, J. A universal model of code-switching and bilingual language processing and production. In: Barbara, Bullock, Toribio, A Jacqueline, editors. The Cambridge Handbook of Linguistic Code-switching. Cambridge University Press; 2009.

Nebes RD, Martin DC, Horn LC. Sparing of semantic memory in Alzheimer’s disease. Journal of Abnormal Psychology. 1984; 93:321–330. [PubMed: 6470317]

Ober BA, Shenaut GK. Lexical decision and priming in Alzheimer's disease. Neuropsychologia. 1988; 26(2):273–286. [PubMed: 3399043]

Ober BA, Shenaut GK, Reed BR. Assessment of associative relations in Alzheimer’s disease: Evidence for preservation of semantic memory. Aging and Cognition. 1995; 2:254–267.

Patterson K, Graham N, Hodges JR. The impact of semantic memory loss on phonological representations. Journal of Cognitive Neuroscience. 1994; 6:57–69. [PubMed: 23962330]

Perry RJ, Hodges JR. Attention and executive deficits in Alzheimer’s disease: a critical review. Brain. 1999; 122:383–404. [PubMed: 10094249]

Philipp AM, Gade M, Koch I. Inhibitory processes in language switching: Evidence from switching language-defined response sets. European Journal of Cognitive Psychology. 2007; 19:395–416.

Philipp AM, Koch I. Inhibition in language switching: What is inhibited when switching between languages in naming tasks? Journal of Experimental Psychology: Learning, Memory, and Cognition. 2009; 35:1187–1195.

Poulisse, N. Slips of the tongue: Speech errors in first and second language production. Amsterdam/Philadelphia: John Benjamins; 1999.

Poulisse N, Bongaerts T. First language use in second language production. Applied Linguistics. 1994; 15:36–57.

Runnqvist E, Strijkers K, Sadat J, Costa A. On the temporal and functional origin of L2 disadvantages in speech production: A critical review. Frontiers in Psychology. 2011; 2:379. [PubMed: 22203812]

Sajjadi SA, Patterson K, Tomek M, Nestor PJ. Abnormalities of connected speech in semantic dementia vs Alzheimer’s disease. Aphasiology. 2012; 26:847–866.

Salmon DP, Butters N, Chan AS. The deterioration of semantic memory in Alzheimer's disease. Canadian Journal of Experimental Psychology. 1999; 53(1):108–117. [PubMed: 10389493]

Salmon DP, Thomas RG, Pay MM, Booth A, Hofstetter CR, Thal LJ, Katzman R. Alzheimer’s disease can be accurately diagnosed in very mildly impaired individuals. Neurology. 2002; 59(7):1022–1028. [PubMed: 12370456]

Salvatierra J, Rosselli M, Acevedo A, Duara R. Verbal fluency in bilingual Spanish/English Alzheimer's Disease patients. American Journal of Alzheimer's Disease and Other Dementias. 2007; 22(3):190–201.

Sasanuma S, Sakuma N, Kitano K. Reading kanji without semantics: Evidence from a longitudinal study of dementia. Cognitive Neuropsychology. 1992; 9:465–486.

Shao Z, Janse E, Visser K, Meyer AS. What do verbal fluency tasks measure? Predictors of verbal fluency performance in older adults. Frontiers in Psychology. 2014; 5:772.doi: 10.3389/fpsyg.2014.00772 [PubMed: 25101034]

Strain E, Patterson K, Graham N, Hodges JR. Word reading in Alzheimer's disease: cross-sectional and longitudinal analyses of response time and accuracy data. Neuropsychologia. 1998; 36:155–171. [PubMed: 9539236]

Thompson-Schill SL, Gabrieli JDE, Fleischman DA. Effects of structural similarity and name frequency on picture naming in Alzheimer’s disease. The Journal of the International Neuropsychological Society. 1999; 5:659–667. [PubMed: 10645707]

Troyer AK, Moscovitch M, Winocur G. Clustering and switching as two components of verbal fluency: Evidence from younger and older healthy adults. Neuropsychology. 1997; 11:138–146. [PubMed: 9055277]

Verhoef K, Roelofs A, Chwilla D. Role of inhibition in language switching: Evidence from event related brain potentials in overt picture naming. Cognition. 2009; 110:84–99. [PubMed: 19084830]

Winer, BJ. Statistical principles in experimental design. 2. New York: McGraw-Hill; 1971.

Weiler M, Fukuda A, Massabki LHP, Lopes TM, Franco AR, Damasceno BP, Cendes F, Balthazar MLF. Default Mode, Executive Function, and Language Functional Connectivity Networks are Compromised in Mild Alzheimer's Disease. Current Alzheimer Research. 2014; 11:274–282. [PubMed: 24484280]

Weissberger GH, Gollan TH, Bondi MW, Clark LR, Wierenga CE. Language and task switching in the bilingual brain: Bilinguals are staying, not switching, experts. Neuropsychologia. 2015; 66:193–203. [PubMed: 25446970]

Weissberger GH, Salmon DP, Bondi MW, Gollan TH. Which neuropsychological tests predict progression to Alzheimer’s disease in Hispanics? Neuropsychology. 2013; 27:343–355. [PubMed: 23688216]

Appendix

An example paragraph and its variants presented between subjects across different conditions.

Single language- English

Pascual did not feel fear when he saw the man for the first time. But over and over he kept thinking about what it was that he was seeing. He turned his face toward the ground. At once, he turned his sight toward where the man was dancing but from then on he no longer saw anything. The man had vanished from the place. At that instant Pascual felt a chill. It was as if someone had tossed water on his back with a bucket. He felt like his face and feet were swollen and then he could no longer walk. With great difficulty, he was able to arrive at his house.

Single Language - Spanish

Pascual no sintió miedo cuando él vio al hombre por primera vez. Pero una y otra vez se quedó pensando sobre qué era lo que estaba viendo. Volteó la cara hacía al suelo. Al momento, volvió la vista donde estaba el hombre bailando pero de allí en adelante ya no vio nada. El hombre se había desaparecido del lugar. En ese momento Pascual sintió un escalofrío. Era como si alguien le hubiera echado agua en su espalda con una cubeta. Sintió que la cara y los pies estaban hinchados y luego ya no podía caminar. Con gran dificultad pudo llegar a su casa.

Mixed Language (Mostly in English)

Pascual did not feel miedo when he saw the man for the first time. But over and over he kept pensando about what it was that he was seeing. He turned his face toward the ground. At once, he turned his sight toward where the man was bailando but from then on he no longer saw anything. The man had vanished from the place. At that momento Pascual felt a chill. It was as if someone had tossed water on his espalda with a bucket. He felt like his face and feet were hinchados and then he could no longer walk. With great difficulty, he was able to arrive at his house.

Mixed Language (Mostly in Spanish)

Pascual no sintió fear cuando él vio al hombre por primera vez. Pero una y otra vez se quedó thinking sobre qué era lo que estaba viendo. Volteó la cara hacía al suelo. Al momento, volvió la vista donde estaba el hombre dancing pero de allí en adelante ya no vio nada. El hombre se había desaparecido del lugar. En ese instant Pascual sintió un escalofrío. Era como si alguien le hubiera echado agua en su back con una cubeta. Sintió que la cara y los pies estaban swollen y luego ya no podía caminar. Con gran dificultad pudo llegar a su casa.

Highlights

Bilinguals with AD produced more intrusions in reading aloud than matched controls.

Sensitivity of intrusions to AD was greatest for dominant language targets.

Language mixing deficits in AD may reflect a monitoring impairment.

Speech errors produced in paragraph reading may be diagnostic of AD in bilinguals.

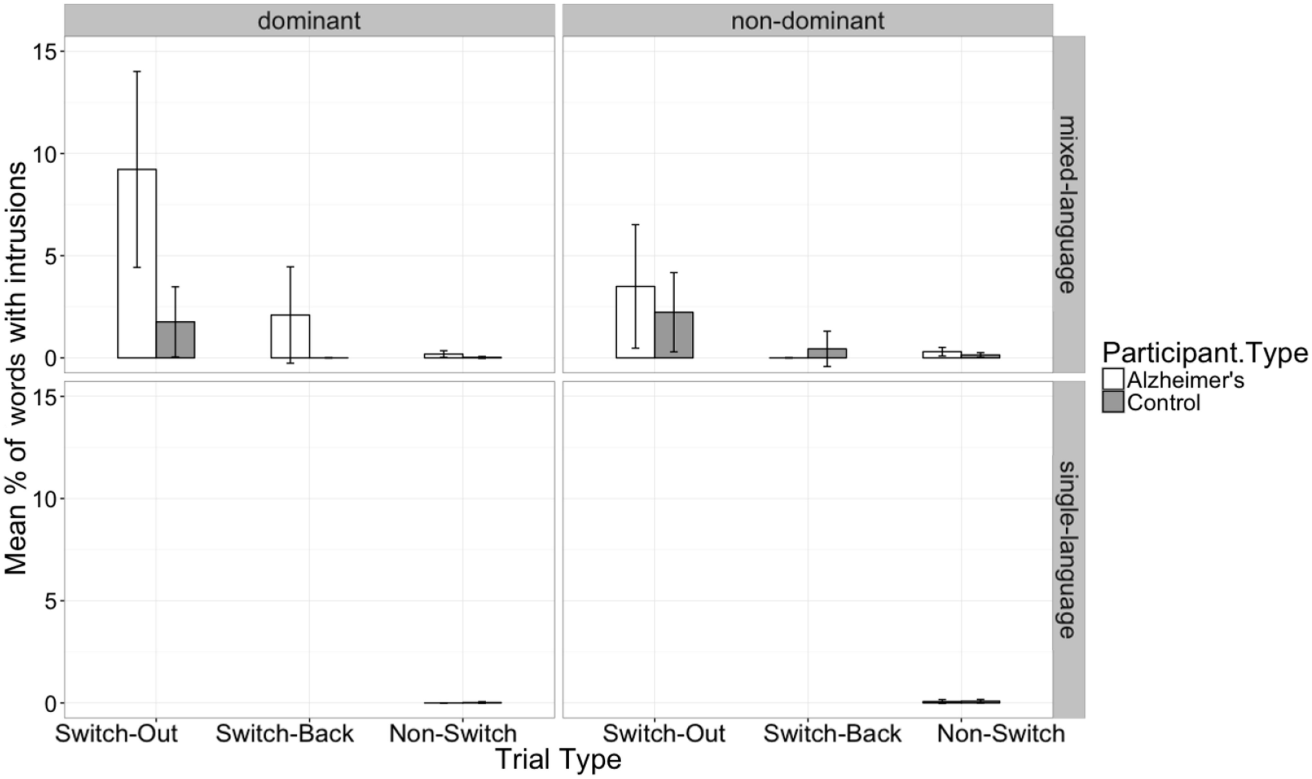

Figure 1.

Mean percent of words with intrusions for each condition and group. The error bars represent 95% Confidence Intervals. Switch-out points refer to words that switched out of the language most words in the paragraph were written in; switch-back refers to words that switched back to the paragraph main language immediately after the switch-out points; non-switch refers to all words that were not switch-out or switch-back words. Dominant vs. non-dominant refers to the language of the target word. Mixed-language vs. single-language refers to paragraph type; mixed-language paragraphs had language switches whereas single-language paragraphs were written entirely in one language.

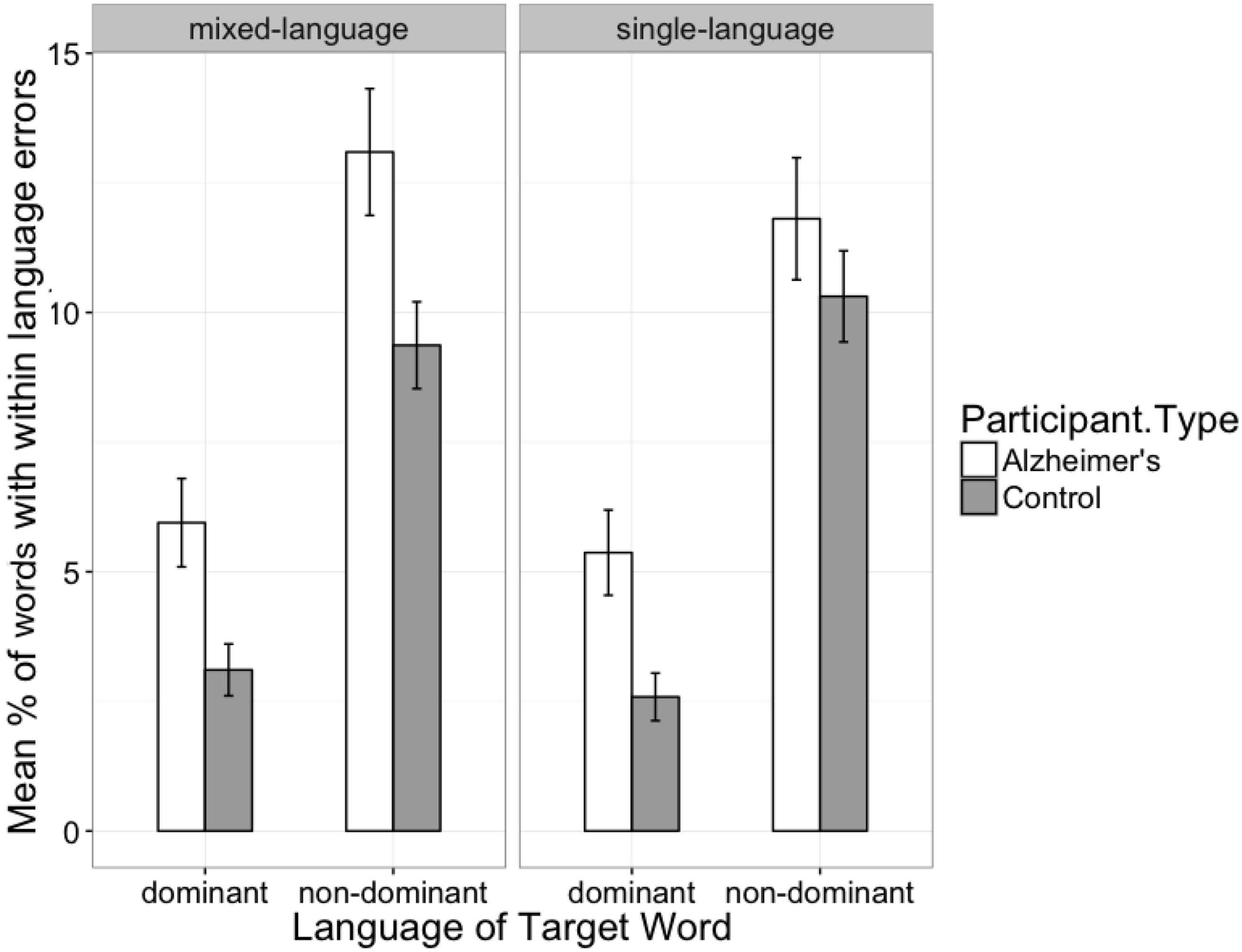

Figure 2.

Mean percent of words with within errors for each condition and group. The error bars represent 95% Confidence Intervals. Mixed-language vs. single-language refers to paragraph type; mixed-language paragraphs had language switches whereas single-language paragraphs were written entirely in one language.

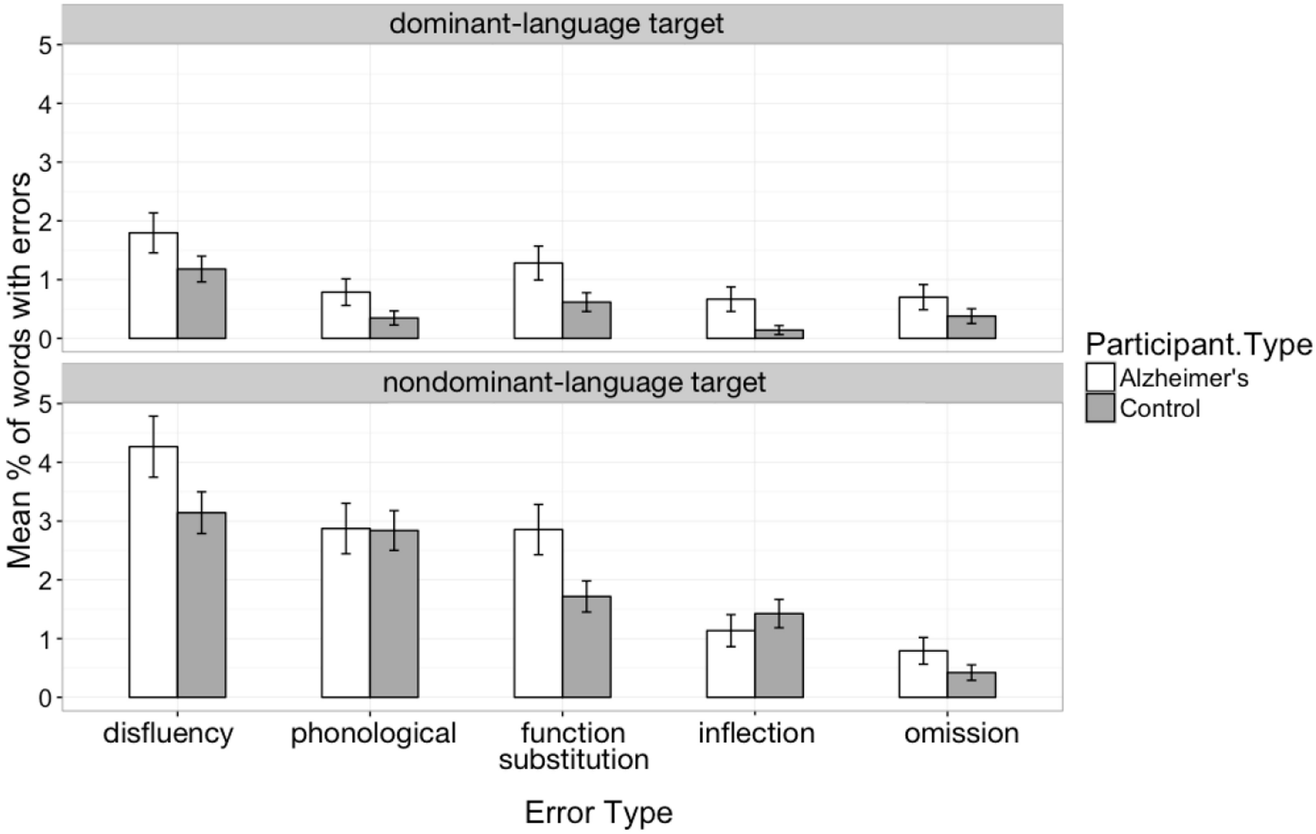

Figure 3.

Mean percent of words that elicited each error type (See Table 2). The error bars represent 95% Confidence Intervals.

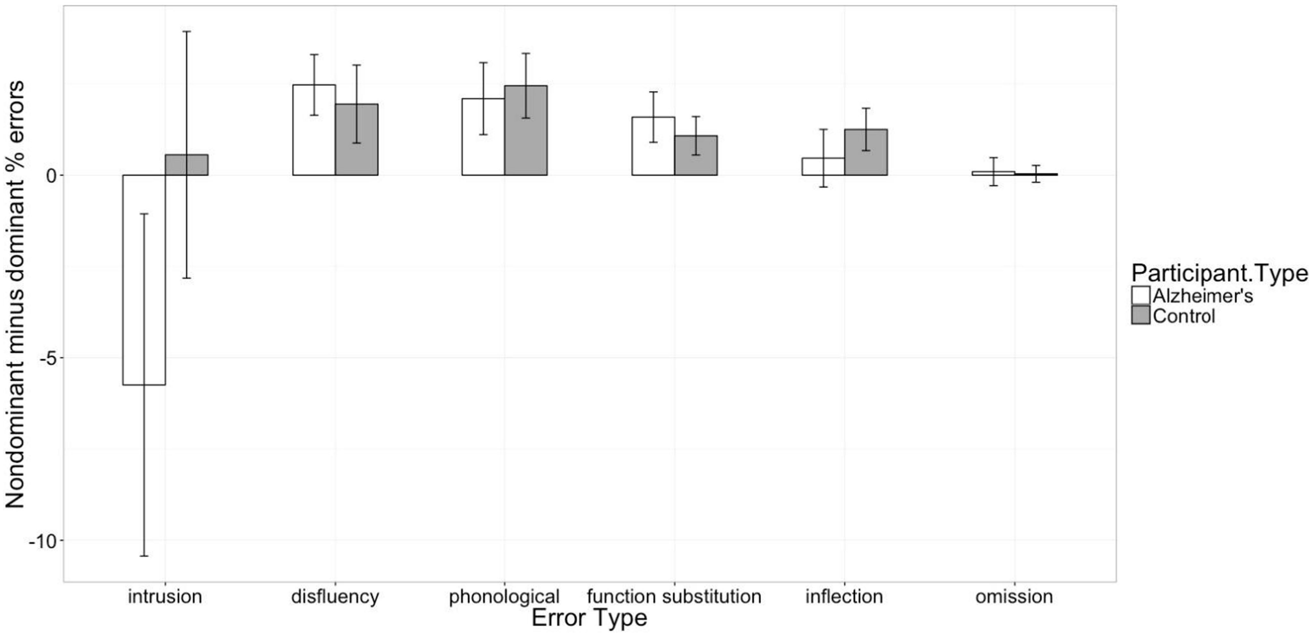

Figure 4.

Language dominance effects i.e., non-dominant minus dominant difference scores in % errors for most commonly produced error subtypes. Positive values indicate that bilinguals produced more errors with non-dominant than with dominant language targets, while negative values indicate more errors with dominant than non-dominant language targets. The error bars represent 95% confidence intervals. Note that the intrusion errors in this figure included those produced on switch-out words in mixed language paragraphs only, while all other difference scores were calculated collapsing together errors produced on all target words in all the paragraphs.

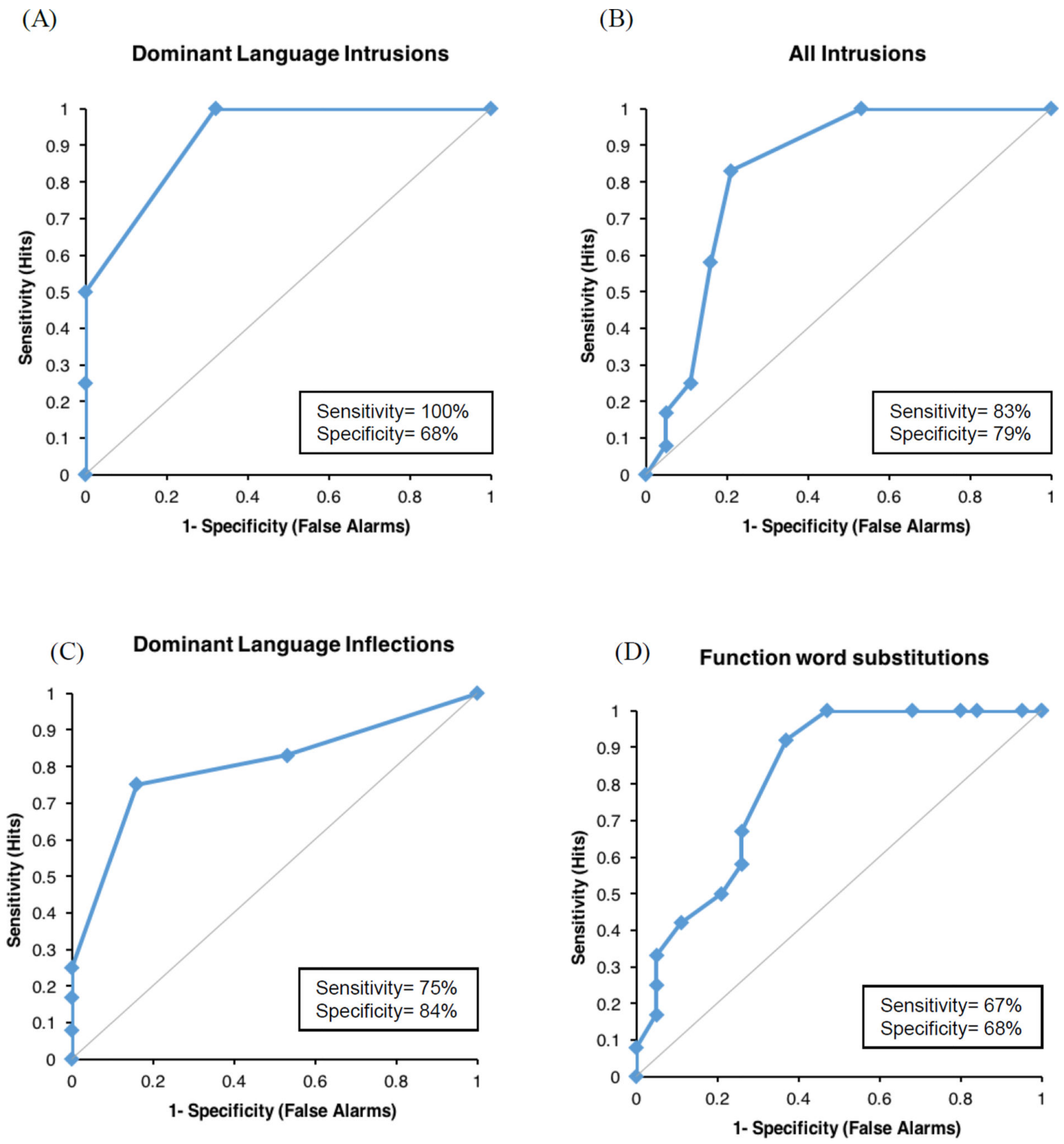

Figure 5.

Receiver Operating Characteristic (ROC) curves comparing sensitivity and specificity for discriminating cognitively healthy bilinguals from those with probable AD. Panel (A) plots intrusion errors produced with dominant language targets. Panel (B) plots all intrusions (collapsed across target language). Panel (C) plots inflection errors produced with dominant language targets. Panel (D) plots the function word substitution errors (collapsed across language dominance). Table 4 shows detailed results of the ROC analyses.

Table 1

Means, standard deviations and group comparisons for demographic characteristics and performance on neuropsychological tests.

	Probable AD (*n* = 12)		Matched Controls (*n* = 19)		Significance Test	
Language History Questionnaire	***M***	***SD***	***M***	***SD***	***t***	***p***
Age	75.8	13.0	73.5	9.1	<1	.57
% Female	58.3	--	73.7	--	--	--
Years of Education	11.8	4.0	13.4	3.9	1.1	.30
Age began using English	11.1	9.2	18.4	15.6	1.5	.15
Age began using Spanish	4.3	9.6	2.3	1.8	<1	.40
Current % English use	61.3	31.4	50.5	38.7	<1	.43
English proficiency self-rated[a]	5.5	1.5	5.1	1.8	<1	.54
Spanish proficiency self-rated[a]	6.1	1.0	6.0	1.1	<1	.76
English MINT score[b]	51.0	13.3	53.3	11.6	<1	.62
Spanish MINT score[b]	50.8	8.6	55.9	10.5	1.4	.17
Dominant-language MINT score[b]	58.8	5.4	62.6	3.5	2.4	.02
Non-dominant-language MINT score[b]	43.0	9.2	46.6	10.0	1.0	.33
Dominant-language self-rated proficiency	6.5	.7	6.8	.4	1.2	.25
Non-dominant-language self-rated proficiency	5.1	1.3	4.3	1.2	1.6	.12
Years lived in USA	55.5	17.8	49.0	25.0	<1	.44
Years lived outside USA	20.3	14.9	25.9	22.9	<1	.46
Neuropsychological Test Battery						
Dementia Rating Scale (DRS)	116.5	13.5	137.1	4.7	6.2	<.001
Trail-Making B time (sec)[d]	248.8	81.9	135.7	67.8	3.4	<.01
Logical Memory A Immediate[c]	6.5	4.9	13.3	2.9	4.8	<.001
Logical Memory A Delayed[c]	3.5	4.5	12.3	2.8	6.6	<.001
Digit Span Forward (# correct)	5.8	1.7	6.5	1.8	1.0	.34
Digit Span Backward (# correct)	3.8	1.6	4.9	1.3	2.3	<.05
Letter Fluency (total score)	28.2	11.3	41.0	9.0	3.5	<.01

	Probable AD (n = 12)		**Matched Controls (n = 19)**		**Significance Test**	
Language History Questionnaire	***M***	***SD***	***M***	***SD***	***t***	***p***
Category Fluency (total score)	27.8	8.5	46.6	6.9	6.7	<.001
Digit Symbol (raw)	22.8	10.3	42.0	10.2	5.1	<.001
Visual Reproduction A Immediate[e]	3.0	1.2	3.2	1.3	<1	.74
Visual Reproduction A Delayed[e]	0.4	.7	2.2	1.8	2.9	<.01
Visual Reproduction A Copy[e]	3.4	1.2	3.4	1.1	<1	.96
Block Design[f]	19.8	11.1	38.1	9.0	4.9	<.001
CERAD (1–3 total correct)[g]	10.8	6.3	20.2	3.3	5.3	<.001
CERAD Delayed Recall (total correct)[g]	2.1	2.7	6.8	1.8	5.7	<.001

[a] Self-rated proficiency level was averaged across ratings for speaking, comprehension of spoken speech, reading, and writing on a scale from 1 (“little to no knowledge”) to 7 (“like a native speaker”).

[b] Maximum possible score is 68. The group difference in MINT scores was not significantly modulated language dominance, (*F*s < 1).

[c] 1/12 patients did not complete this test (i.e., missing data)

[d] 6/12 patients were unable to complete this test

[e] 2/12 patients did not complete this test (i.e., missing data)

[f] 1/12 patients did not complete this test (i.e., missing data)

[g] 1/12 patients did not complete this test (i.e., missing data)

Table 2

Between-language (in first three rows) and within-language error types (all rows below the third), counts, definition, and examples for all error types that bilinguals produced during the read aloud task. *Italicized text* = error word the speaker produced; [ ] = the written target word (that the speaker should have but did not produce).

Error Type	Error Rate (%)		Definition	Example(s)	Context Example
	AD	Controls			
Intrusion (*n*=58)	3.2	1.8	Speaker produced the target in the wrong language	casa → house cheese → queso and → y	That *house* [casa] belonged to an old woman that was blind.
Partial intrusion (*n*=12)	0.7	0.3	Speaker began to produce the target in the wrong language but self-corrected before fully producing the error	woman → mu…woman [Spanish for 'woman' is 'mujer'] morning → ma…morning [Spanish for 'morning' is 'mañana']	Pasando por la orilla del cementerio viejo, cuando volvió a ver, vio a una *mu…woman* [woman] que estaba tejiendo en una esquina del cementerio.
Accent (*n*=41)	2.2	1.3	Speaker produced the correct target but with an accent that matches the other language	/pæntɹij/ → /pantri/	Ella los guardaba en la *pantry-Spanish accent* [pantry-English accent] envueltos en trapos hechos de sacos de harina.
Disfluency (*n*=753)	31.3	32.8	Speaker produced either a false start, hesitation, repetition; or speaker approached target but paused or did not finish	mysteriously → mysterious… mysteriously without → ehhhh… without	But all this came to an end from one day to the next when each one of the cabras that they had *mysterious… mysteriously* [mysteriously] died.
Phonological (*n*=508)	18.9	24.2	Speaker produced a word or nonword that has approximately 50% phonological overlap with the target	look → lock first → feast barely → rarely	Passing along the edge of the old cemetery, when he turned to *lock* [look], he saw a woman that was weaving in a corner of the cemetery.
Function word substitution (*n*=457)	21.4	17.7	Speaker substituted a function word target with a different function word	toward → to up → of what → was	He turned his face *to* [toward] the ground.
Inflection (*n*=250)	9.3	11.9	Speaker produced a syntactic error, either adding or omitting an affix	began → begin animals → animal abandoned → abandon	Very slowly I *begin* [began] to see a bit better thanks to the luz from the oil lamp in the kitchen.
Omission (*n*=161)	7.7	6.1	Speaker skipped the target word	then →	The little road [then] passed by the orilla of the old cemetery.
Insertion (*n*=66)	3.7	2.0	Speaker inserted a word not written on the page	→ not	The woman was *not* [ ] happily tejiendo but she kept her face hidden.
Other word (*n*=37)	1.3	1.8	Speaker produced a real word either semantically related or unrelated to the target	woman → man owners → home love → her	That house belonged to an old *man* [woman] that was blind.
Nonsense word (n = 4)	0.3	0.1	Speaker produced a nonword that does not share more than 50% of phonemes with the target word	casa → chente el → onde	Cuando regresó a *chente* [casa] oyó noises que venían de la cocina.

Table 3

Mean total paragraph reading times (RT) and standard deviations (SD) for patients and controls in each condition. Dominant vs. non-dominant in the RT measure refers to the language in which all (single-language) or most (mixed-language) words were written (see Appendix).

		Mixed-language paragraphs		**Single-language paragraphs**	
		RT	***SD***	***RT***	***SD***
Alzheimer's	Dominant	58.3	18.9	53.6	19.0
	Non-dominant	78.2	19.7	73.3	21.8
Controls	Dominant	53.5	12.3	49.4	12.5
	Non-dominant	71.8	18.5	69.6	19.7

Table 4

Area under curve (AUC), sensitivity, specificity and the Youden index for evaluating ROC Curves for distinguishing patients from controls by plotting the total number of speech errors of different types produced.

Measure	AUC	*p*	AUC best rank	Cut-off value	Sensitivity (%)	Specificity (%)	Youden Index	Youden Index Rank
Intrusions dominant	.92[a]	.000	1	.5	100	68	.68	1
Intrusions dominant switch-out	.84	.002	2	.5	83	79	.62	2
Intrusions all	.84	.002	2	1.5	83	79	.62	2
Inflections dominant	.81	.005	3	1.5	75	84	.59	4
Function word substitutions all	.80	.006	4	14.5	67	68	.35	9
Omissions all	.79	.007	5	4.5	92	68	.60	3
Function word substitutions non-dominant	.78	.009	6	8.5	75	83	.58	5
Omissions non-dominant	.78	.011	6	2.5	75	68	.43	7
Function word substitutions dominant	.77	.013	7	3.5	83	63	.46	6
Phonological errors dominant	.70	.059	8	3.5	50	90	.40	8
Omissions dominant	.67	.114	9	2.5	50	68	.18	10

[a] The Intrusions dominant score yielded the largest AUC value of all, though removing self-corrected intrusions from these scores improved specificity from 68% to 79% while decreasing sensitivity only slightly to from 100% to 92% (an AUC of 0.90). Removing self-corrected intrusions from the other intrusion measures decreased slightly the AUC value for Intrusions dominant switch-out from 0.84 to 0.83, and increased the AUC value for Intrusions all from 0.84 to 0.88, but no measure exceeded the AUC for Intrusions dominant (i.e., 0.92 was the highest AUC of all).

Table 5

a. Bivariate correlations between different error types produced during paragraph reading and neuropsychological measures in cognitively healthy bilinguals (*n*=19). Correlations included errors produced in both languages (collapsed together) but those that were significant in the dominant-language on its own are highlighted in bold.

Measure	**Intrusions^{a}**	**Inflections**	**Function word substitutions**	**Omissions**	**Phonological**
Intrusions^{a}	-				
Inflections	.087	-			
Function word substitutions	.220	.676**	-		
Omissions	.047	.416	.713**	-	
Phonological	.333	.605**	**.757****	.313	-
Dementia Rating Scale (DRS)	.072	−.650**	−.616**	−.232	−.488*
Logical Memory A Immediate	.207	−.118	.033	.064	.170
Logical Memory A Delayed	.157	.115	.259	.488*	.168
Digit Span Backward (# correct)	.125	−.678**	−.328	−.265	−.448
Letter Fluency (total score)	−.126	−.448	**−.681****	−.657**	−.492*
Category Fluency (total score)	.028	**−.292**	−.515*	−.524*	−.198
Digit Symbol (raw)	−.171	−.612**	**−.455**	−.259	**−.505***
Visual Reproduction A Delayed	−.076	−.438	−.424	−.209	−.281
Block Design	−.148	.203	.218	.032	.259
CERAD (1–3 total correct)	.494*	−.318	−.283	−.270	−.290
CERAD Delayed Recall (total correct)	.411	.004	−.043	−.003	−.099
Trail-Making B time (sec)	.142	.815**	**.654****	.503*	**.579****

b. Bivariate correlations between different error types produced during paragraph reading and neuropsychological measures in bilinguals with AD (*n*=12). Correlations included errors produced in both languages (collapsed together) but those that were significant in the dominant-language on its own are highlighted in bold.

Measure	**Intrusions^{a}**	**Inflections**	**Function word substitutions**	**Omissions**	**Phonological**
Intrusions^{a}	-				
Inflections	.042	-			

b. Bivariate correlations between different error types produced during paragraph reading and neuropsychological measures in bilinguals with AD (*n*=12). Correlations included errors produced in both languages (collapsed together) but those that were significant in the dominant-language on its own are highlighted in bold.

Measure	Intrusions[a]	Inflections	Function word substitutions	Omissions	Phonological
Function word substitutions	.643 *	.067	-		
Omissions	.045	−.168	.497	-	
Phonological	.647 *	.265	**.821** **	.377	-
Dementia Rating Scale (DRS)	.044	−.120	.304	.506	.140
Logical Memory A Immediate	.053	−.326	.016	.334	−.196
Logical Memory A Delayed	.070	347	−.295	−.017	**−.417**
Digit Span Backward (# correct)	−.233	−.064	.164	.367	−.179
Letter Fluency (total score)	−.611 *	.075	−.289	−.154	−.366
Category Fluency (total score)	.069	.137	.257	.386	.093
Digit Symbol (raw)	−.186	−.406	.069	.337	.108
Visual Reproduction A Delayed	.309	−.243	−.037	.113	.034
Block Design	−.313	−.458	−.050	.065	−.234
CERAD (1–3 total correct)	.071	.005	.354	.565	.149
CERAD Delayed Recall (total correct)	.327	−.264	.051	.255	.038
Trail-Making B time (sec)[b]	−.097	.505	.394	−.090	.231
Trail-Making B time (sec)[c]	−.049	.251	.099	−.305	.030

[a] Collapsing together all intrusions produced including switch-out, switch-back, and non-switch positions

** = $p < .01$;

* = $p < .05$; Bolded text highlights correlations significant at $p<.05$ level for dominant language targets only.

[a] Collapsing together all intrusions produced including switch-out, switch-back, and non-switch positions

[b] = For these correlations, there were missing data for 6/12 patients

[c] = For these correlations, data that were missing were replaced with “300” as the maximum time limit for this task

** = $p < .01$;

* = $p < .05$; Bolded text highlights correlations significant at $p<.05$ level for dominant language targets only