CAMBRIDGE
UNIVERSITY PRESS

What is proficiency? Characterizing spoken language proficiency in older Spanish-English bilinguals

Dalia L. Garcia[1] and Tamar H. Gollan[2]

[1]Joint Doctoral Program in Language and Communicative Disorders, San Diego State University/University of California, San Diego, CA, USA and [2]Department of Psychiatry, University of California, San Diego, CA, USA

Abstract

We conducted a detailed linguistic analysis of Oral Proficiency Interviews (OPIs) from older Spanish-English bilinguals ($n = 28$) to determine which cognitive, linguistic, and demographic factors predict proficiency. In the dominant language, older age was associated with lower proficiency scores, but aging effects were not significant after accounting for cognitive functioning scores. In the nondominant language, bilinguals with larger vocabulary scores, fewer speech errors, and higher education levels obtained higher proficiency scores. Multiple linguistic submeasures from the OPIs were highly correlated across languages (e.g., fast speakers spoke fast in both languages), but these same measures exhibited significant language dominance effects (e.g., bilinguals spoke faster in the dominant than in the nondominant language). These results suggest it is critical to control for cognitive functioning when examining aging effects on language production, reveal powerful individual differences that affect how people talk regardless of language, and validate the use of the OPI to measure bilingual proficiency.

Highlights

- Individual differences affect how people speak despite proficiency limitations.
- Cognitive functioning, not age, predicts proficiency in the dominant language.
- Aging effects may be masked by cognitive functioning in older adults.
- Language knowledge and education predict proficiency in the nondominant language.
- Linguistic submeasures correlated across languages and showed dominance effects.

1. Introduction

Speaking requires accessing knowledge from various linguistic domains, including vocabulary, syntax, morphology, and phonology, and coordinating their activation to enable fluent language production, which also requires organizing and planning ideas and thoughts. Speaking in a less proficient or *nondominant* language likely imposes greater cognitive demands than speaking in a stronger or *dominant* language. Although different cognitive mechanisms may enable production in each language, it is also likely that many shared mechanisms affect production in both languages. Comparisons across languages produced by the same bilingual speakers can offer unique insights into the cognitive mechanisms that underlie language proficiency in bilinguals and monolinguals alike.

Although many studies have documented language dominance effects, surprisingly little is known regarding which aspects of spontaneous language are critical for obtaining higher versus lower proficiency ratings. For example, it is well documented that bilinguals can name pictures faster and with fewer errors (Hanulová et al., 2011; Runnqvist et al., 2011) and also produce fewer grammatical errors in their dominant than their nondominant language (Pirvulescu et al., 2014). However, it is not known if more proficient speakers also talk faster or to what extent the production of errors is critical (i.e., not all proficient speakers speak quickly, and even the most proficient speakers occasionally produce some types of speech errors). This lack of knowledge is likely due to the difficulty of studying spontaneous language production. Utterances must be elicited, transcribed, and then coded for relevant linguistic properties, and for bilinguals, this must be done in a comparable manner for two languages. Recent years have brought a surge of interest in automated analyses of language, but these have yet to be applied to bilinguals, which present many challenges. Many automatic transcribers cannot accurately transcribe language samples produced by speakers with nonnative accents, and deriving equivalent measures across languages is also challenging (e.g., measures of grammatical complexity in particular might be difficult to equate across languages).

An additional consideration is that many individual differences might affect how a person talks, and proficiency in both languages might be affected by demographic, developmental,

cognitive, personality, and cultural factors that are not necessarily related to language proficiency per se. Consistent with this possibility, one study examined narrative language samples in bilingual kindergarteners and found that kindergartners who produced longer utterances in one language also did so in the other (Bedore et al., 2010). Furthermore, children who produced longer utterances tended to also produce more utterances and use more different words, and this was true in both languages the children spoke.

Similar considerations must apply in the analysis of how aging affects bilingual proficiency. In conversational language samples, older adults spoke for longer periods of time (Pereira et al., 2019), with longer utterances, produced more fillers (Horton et al., 2010), and more semantic paraphasias (Marini et al., 2005) than younger adults (see Mortensen et al., 2006 for a review). Longitudinal analyses of spontaneous language samples showed that as people aged, they spoke more slowly and repeated more words (Beier et al., 2023). Age effects can also vary with task demands, topics, or goals. In one study, older adults spoke more than younger adults about personal topics but not when describing a picture (James et al., 1998). Older-old adults also talked more than younger-old adults, but aging-related changes do not necessarily reflect cognitive decline, e.g., older adults may talk more because they have different communicative priorities (Pushkar et al., 2000). Furthermore, some aging effects, especially on speaking rate, may be confounded by overall health (Ramig, 1983) or may be due to changes in control over movement timing of articulators (Tremblay et al., 2017).

Education level can also affect how people talk. Researchers examined the effects of age and education on spontaneous language samples in three groups of adults, including young (ages 18–35), middle-aged (ages 36–50), and older adults (ages 51–76; Mulder & Hulstijn, 2011). In this study, participants produced monologues in a role-play based on a picture prompt. Participants with postsecondary education produced more words, more utterances, and fewer grammatical errors than those without postsecondary education, but aging effects were not significant. Education effects have also been found in a simple picture description task, where speakers with lower education levels produced shorter and less complete picture descriptions than speakers with higher education levels (MacKenzie, 2000).

Cognitive functioning is another individual difference that can affect different aspects of language production, including planning and fluency. For example, people with lower verbal IQ (WAIS-R Comprehension, Similarities, and Vocabulary) produced more inadvertent repetitions in a sentence repetition task (Engelhardt et al., 2019). Greater interference in the Stroop task was associated with the production of more repair disfluencies (stopping and restarting an utterance with a different word or phrase; Engelhardt et al., 2013). Individuals with cognitive impairment due to Alzheimer's disease spoke with reduced lexical diversity compared to controls (Ahmed et al., 2013; Berisha et al., 2015; Kavé & Goral, 2016), as did individuals with mild cognitive impairment who also produced fewer and shorter words and shorter video clip descriptions (Reeves et al., 2023).

The studies reviewed above used a variety of methods to elicit speech, including picture description, role playing, and a picture-verb sentence production task. Bilingual proficiency level is often assessed with Oral Proficiency Interviews (OPIs), which may also be especially useful for identifying individual differences in speaking ability and for revealing the cognitive mechanisms underlying language production. OPIs require speakers to answer a variety of questions that increase in complexity as the interview progresses in order to test the limits of communicative competence. Bilinguals

can be interviewed in both languages, providing a unique opportunity to examine speaker-dependent and language dominance-dependent language markers.

This study examined OPIs of older Spanish-English bilinguals to determine which linguistic and nonlinguistic factors best predict proficiency in the dominant versus the nondominant languages. Although OPIs have excellent construct validity, they do not provide specific information about which aspects of language production lead to higher proficiency scores in each language, which can be revealed by examining relationships between linguistic submeasures and OPI scores in each language. Specifically, we examined measures of vocabulary (number of different words), syntax (mean length utterance in words), talking speed (words per minute), verbosity (total utterances), language knowledge or monitoring failures (errors), and fluency (revisions, repetitions, and filled pauses or *RRFP*s). To provide further evidence for the validity of the OPI as a test of language proficiency, we asked which of these linguistic submeasures exhibited significant language dominance effects. We also asked whether age would survive as a predictor of OPI scores after accounting for cognitive functioning (the Dementia Rating Scale [DRS]; Mattis, 1988). This question would help to address a critical limitation in existing studies, which often included only one of these two variables (Breier et al., 2023; Horton et al., 2010; James et al., 1998). Finally, given the unique opportunity to examine both bilinguals' languages, we asked which linguistic submeasures are correlated across languages, which would suggest shared underlying cognitive abilities and individual differences that affect language production in general.

2. Methods

2.1. Participants

Twenty-eight Spanish-English bilinguals aged 64–90 (*M* = 77.0, *SD* = 8.5) who participated in the study by Weissberger et al. (2012) were included in this study. All were cognitively healthy at the time of testing based on their performance on the DRS (Mattis, 1988). Nineteen participants also completed more extensive cognitive, neurological, and neuropsychological exams as part of their participation in the longitudinal study at the UCSD Alzheimer's Disease Research Center (ADRC). Eleven of the nineteen remained healthy for as long as they were followed by the ADRC (on average 6.4 years after their participation in the study), while 8 were diagnosed with Alzheimer's Disease (AD; about 4 years later; Garcia & Gollan, 2024). All 28 participants scored between 131 and 142 (maximum score is 144) on the Dementia Rating Scale (see Table 1 and Figure S1 in the Supplementary Materials; a score of ≤130 is within normal limits for this demographic; Weissberger et al., 2013).

Seventeen of the 28 bilinguals were English-dominant, reported first being exposed to English at age 4.1 years on average (*SD* = 3.7), and reported using English about 81.4% of the time (*SD* = 16.2%). On average, these English-dominant bilinguals had lived in a Spanish-speaking country for 2.0 years (*SD* = 2.4). The remaining 11 bilinguals were Spanish-dominant, reported first being exposed to English at age 9.9 (*SD* = 7.4) on average, reported currently using English only 36.8% of the time (*SD* = 26.8%), and had lived in a Spanish-speaking country for an average of 25.5 years (*SD* = 20.6).

2.2. Materials and procedures

The research protocol was approved by the UCSD Institutional Review Board in accordance with the Helsinki Declaration. Participants

Table 1. Participant characteristics

Gender (male, female)	English dominant (*n* = 17)		Spanish dominant (*n* = 11)	
	M	*SD*	*M*	*SD*
Age	78.0	9.5	75.5	6.8
Education	14.1	2.4	11.6	3.4
Dementia rating scale	135.3	3.8	136.8	4.1
Mini mental state examination	28.7	1.3	28.7	1.8
First exposure to English	4.2	3.9	11.0	7.6
Years immersed in Spanish	2.0	2.4	30.7	30.7
Dominant language use growing up (%)	56.5	26.0	91.3	14.5
Nondominant language use growing up (%)	43.5	26.0	8.7	14.5
Dominant language use currently (%)	81.4	16.2	65.4	24.8
Nondominant language use currently (%)	18.6	16.2	34.6	24.8
Dominant language OPI score	8.7	0.7	9.3	0.6
Nondominant language OPI score	7.5	1.1	7.3	1.6
Dominant language MINT score	62.2	3.7	61.1	2.6
Nondominant language MINT score	48.8	6.8	53.9	9.6

Note: MINT = Multilingual Naming Test (maximum score = 68); OPI = Oral Proficiency Interview (maximum score = 10).

first completed a Language History Questionnaire. Those who were recruited from the community then completed the DRS in their self-reported dominant language. Bilinguals enrolled in the ADRC completed the DRS during their annual neuropsychological evaluation (and we used the score from whichever administration date was closest to the time of participation in this study). Bilinguals then completed the OPI and the Multilingual Naming Test (MINT; Gollan et al., 2012) in each language in counterbalanced order (between participants with both tests given in English first and then in Spanish, or vice versa[1]).

2.2.1. Oral proficiency interview

The OPI included five questions and a picture description in each language. The questions increased in difficulty to elicit a variety of language skills, with the last two questions eliciting higher-level language skills (e.g., complex sentence structure, defending an opinion). Participants were interviewed by a proficient native Spanish-English bilingual. A score of 1–10 was assigned to each interview by two raters (Gollan et al., 2012). A third rater assigned scores only when there was a discrepancy of two or more points between the initial two raters. An average of all the raters' scores was taken to assign each interview a final score. The OPI scores were then used to determine language dominance.

The OPIs were transcribed by 6 proficient Spanish-English bilinguals. Each interview was transcribed by one research assistant and checked by another to ensure transcription accuracy. If the two transcribers could not reach a consensus on a part of the transcription, a third transcriber was consulted. Once a transcript

---

[1]Linguistic and nonlinguistic switching tasks were administered last (Weissberger et al., 2012); performance on these tasks is not discussed herein.

was deemed accurate, one of three trained bilinguals formatted the transcript for the analysis software. We used the Systematic Analysis of Language Transcripts (SALT; Miller & Iglesias, 2020) software to extract linguistic measures from the language samples. The SALT reports provide measures in both English and Spanish of lexical diversity (e.g., number of different words; NDW), speaking speed (e.g., words per minute; WPM), syntactic complexity (e.g., mean length utterance in words; MLU, calculated by adding together all words a person produced divided by the number of utterances), errors, and a measure that combines revisions, repetition, and filled pauses, abbreviated as RRFPs (or "*mazes*" according to the SALT terminology to measure speaking fluency and planning ability; Bedore et al., 2006; Fagan, 1982; Loban, 1976). The RRFP measure was calculated by dividing the total number of RRFPs by the total number of words plus RRFPs. Utterances included independent clauses with modifiers if produced (i.e., subordinate clauses). Finally, the percentage of utterances with errors measures linguistic knowledge and the ability to monitor speech for errors while also planning and producing speech. This included word-level errors (e.g., *I *is in my house*), extra words (e.g., *After I got glasses, I saw *more better*), and utterance-level errors (e.g., *They *came store went*), and omissions (e.g., *I went *(to) the store*).

2.2.2. Multilingual naming test (MINT)

The MINT (Gollan et al., 2012) has 68 black-and-white line drawings. The test starts with easy items and difficulty increases as the test progresses. Responses were counted as correct if produced spontaneously or with a semantic cue (i.e., a prompt that provides information about the meaning of the target word).

2.2.3. Dementia rating scale (DRS; Mattis, 1988)

The DRS is a standardized 144-point mental status test with subscales for Attention (37 points; including forward and backward digit span, ability to follow commands, and visual search tasks), Initiation and Perseveration (37 points; includes semantic fluency, copying and maintaining movements modeled by experimenter, and copying designs), Construction (6 points; includes copying shapes and writing your name), Conceptualization (39 points; requires participants to identify abstract and concrete similarities and differences between words and shapes), and Memory (25 points; includes orientation questions [i.e., date, current president, etc.], sentence recall, word recognition, and visual memory). The DRS takes about 10–15 minutes to administer in healthy older adults.

2.2.4. Mini-mental state examination (Folstein et al., 1975)

The Mini-Mental State Examination (MMSE) is a brief screener that assesses orientation to time and place, attention and concentration, recall, language, and visual construction. This test takes about 8 minutes to administer. A cutoff score of 24/30 was recommended by Folstein et al. (1975) to differentiate impaired and healthy older adults.

2.3. Preprocessing and analyses

Analysis 1 consisted of running exploratory correlations between overall OPI scores, demographic variables, cognitive tests, and linguistic submeasures to identify factors that may explain unique variance in OPI scores. Variables that were correlated with OPI scores were then entered into a linear backward regression model with OPI scores as the dependent variable. OPI scores were treated

as a continuous (rather than ordinal) variable because the resulting score was an average taken from 2 to 3 raters. These models examined which variables explained unique variance in the overall OPI scores in each language. Analysis 2 consisted of repeated measures ANOVAs with each language as a repeated factor for each submeasure to determine which linguistic submeasures were sensitive to the effects of language dominance. Finally, Analysis 3 consisted of between-language correlations for each linguistic submeasure while controlling for age to examine which linguistic submeasures were subject to individual differences.

3. Results

3.1. Measures of overall cognitive ability (MMSE versus DRS)

DRS scores were correlated with four measures including overall proficiency (OPI) scores in both languages (dominant $r = .571$, $p < .001$; nondominant $r = .396$, $p < .05$), and speaking speed (WPM) in both languages (dominant $r = .508$, $p < .05$; nondominant $r = .466$, $p < .05$). However, the MMSE (which takes less time to administer than the DRS) was not significantly correlated with OPI scores (or any other measure) (all $ps \geq .288$; see Table 2); thus, we focused our analyses below on the more thorough DRS measure.

3.2. Analysis 1: What predicts proficiency?

3.2.1. Dominant language
In the dominant language, DRS scores and age were significantly correlated with OPI scores, and education effects were marginal. Looking at linguistic submeasures, only one correlation was significant, and not in an expected direction; speakers who produced more RRFPs were rated as more proficient overall ($r = .571$, $p = .002$; see Figure 1) than those who produced fewer RRFPs in the dominant language. Importantly, people who produced more RRFPs were not doing so because they were speaking faster, i.e., RRFPs and speaking speed were not significantly correlated ($r = .009$, $p = .964$). A backward regression analysis with a criterion set at a probability of F-to-remove of greater than or equal to .10 including DRS, age, education, and RRFPs showed that DRS scores, $B = .110$, $t(25) = 4.15$, $p < .001$ and RRFPs, $B = .102$, $t(25) = 3.278$, $p = .003$ together explained 50% of the variance in OPI scores in the dominant language. Age and education effects were removed from the model. See Table 3 for full model results.

3.2.2. Nondominant language
In the nondominant language, age was not significantly correlated with OPI scores, but DRS scores and vocabulary/MINT scores were, and education was marginal. Several linguistic submeasures were correlated with overall OPI scores in the nondominant language, including (in order of significance, see Figure 2) errors, lexical diversity (NDW), and total utterances. However, in the backwards regression analysis only MINT scores, $B = .10$, $t(25) = 4.21$, $p < .001$, years of education, $B = .23$, $t(25) = 3.46$, $p = .002$, and errors, $B = -.06$, $t(25) = -2.24$, $p = .035$ were significant, and together explained 51% of the variance, while DRS scores, total utterances, and NDW were removed from the model. See Table 4 for full model results.

3.2.3. Speaking rate
Since several papers reported age effects on speaking rate (Beier et al., 2023; Duchin & Mysak, 1987; Horton et al., 2010), and in the analyses above, age did not explain unique variance when entered into a model with cognitive functioning (DRS) scores. We asked a follow-up question, which was if age effects on speaking rate might be more robust even after controlling for cognitive functioning. Speaking rate (WPM) slowed with increased age in both the dominant ($r = -.427$, $p = .023$) and nondominant language ($r = -.423$, $p = .025$), and DRS scores were positively correlated with speaking rates in both the dominant ($r = .508$, $p = .008$) and nondominant languages ($r = .466$, $p = .017$). However, a backward regression analysis showed that only DRS scores significantly predicted speaking rate in the dominant language, $B = 3.21$, $t(25) = 2.89$, $p = .008$, while the effect of age was removed. Similarly, DRS predicted speaking rate in the nondominant language, $B = 3.02$, $t(25) = 2.58$, $p = .017$, but age was removed from the model. See Table 5a,b for full model results.

3.3. Analysis 2: Validating the OPI with linguistic submeasures

Language dominance effects are shown in Figure 3. Bilinguals named more pictures in the dominant than in the nondominant language, $F(1.27) = 48.89$, $\eta_p^2 = .644$, $p < .001$, and many linguistic submeasures produced significant language dominance effects. Bilinguals spoke faster, $F(1.27) = 27.98$, $\eta_p^2 = .509$, $p < .001$, produced more different words, $F(1.27) = 15.49$, $\eta_p^2 = .365$, $p = < .001$, longer utterances, $F(1.27) = 7.45$, $\eta_p^2 = .217$, $p = .011$, and marginally more utterances, $F(1.27) = 3.56$, $\eta_p^2 = .116$, $p = .07$, in the dominant than in the nondominant language. Conversely, bilinguals produced more RRFPs, $F(1.27) = 27.83$, $\eta_p^2 = .508$, $p < .001$, and more utterances with errors, $F(1.27) = 33.00$, $\eta_p^2 = .550$, $p < .001$ in the nondominant than the dominant language.

Table 2. Pearson correlations between overall oral proficiency interview scores in each language, demographic variables, and cognitive test scores

	Dominant language OPI	Nondominant language OPI	Age	Education	Dementia rating scale	Mini mental state examination	Dominant language MINT
Nondominant language OPI	.558**	–					
Age	−.455*	−.200	–				
Education	.317†	.354†	−.126	–			
Dementia rating scale	.571**	.396*	−.381†	.276	–		
Mini mental state examination	−.015	−.049	.020	.181	.314	–	
Dominant language MINT	.285	.260	−.372†	.203	.471*	.106	–
Nondominant language MINT	.245	.511**	.175	−.128	.262	.080	.181

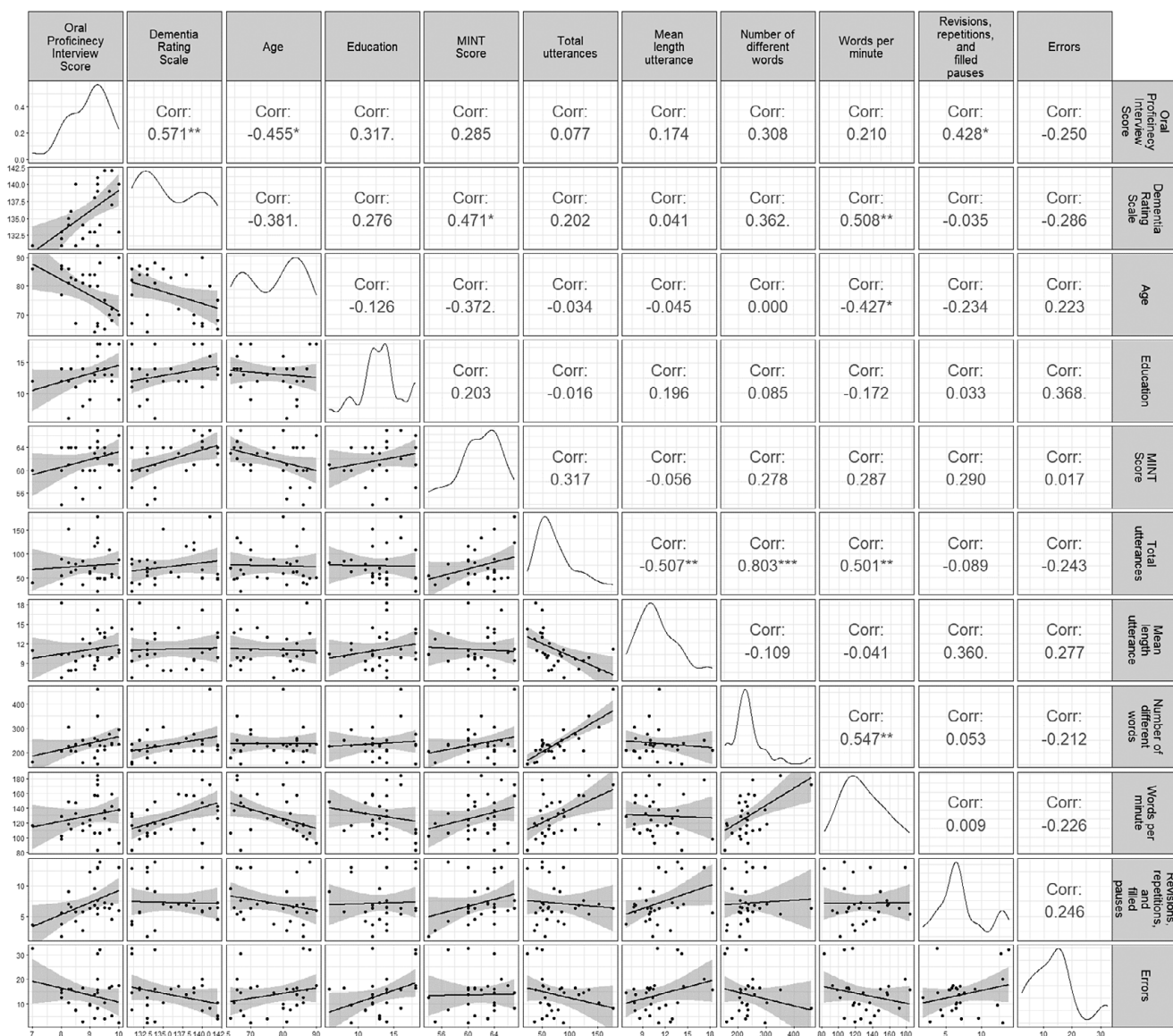†$p \leq .10$, *$p \leq .05$, **$p \leq .01$.

Figure 1. Dominant language correlation table and scatter plots.
Note: ˙*p* < .10, *p < .05, **p < .01, ***p < .001.

Table 3. Backwards regression results using demographic and linguistic variables to predict Oral Proficiency Interview scores in the dominant language

Independent variables	B	SE	Beta(β)	t	p
(Constant)	−6.78	3.62		−1.87	0.74
Dementia rating scale	0.11	0.03	0.59	4.15	<.001
RRFPs in the dominant language	0.10	0.03	0.46	3.28	0.03

Note: The table below shows the final model after removing age and education. R^2 = .540, $F(2.23)$ = 13.52, p < .001.

3.4. Analysis 3: Do individual differences affect how people talk in both languages?

Figure 4 shows partial correlations between OPI scores and linguistic submeasures from the OPI in each language while controlling for age, and for comparison, we also included MINT scores. The correlations between overall OPI scores in the dominant and nondominant language remained significant after controlling for age (r = .535, p = .004), and all other linguistic submeasures were also significantly correlated across languages. In order of correlation strength, RRFPs, (r = .768, p < .001), total utterances, (r = .725, p < .001), NDW (r = .702, p < .001), WPM (r = .628, p < .001), and MLU(r = .546, p = .003) were all correlated across languages. By contrast, picture naming scores (i.e., the MINT; r = .269, p = .175) and the percentage of utterances with errors (r = .258, p = .194) were not correlated across languages.

3.4.1. Lexical diversity and verbosity

Though we found little overlap in which linguistic submeasures predicted OPI scores in the dominant versus in the nondominant languages, several similarities between linguistic submeasures emerged within each language. Specifically, total utterances, speaking speed (WPM), and lexical diversity (NDW) were significantly correlated in all possible combinations (see Figures 2 and 3). That is,
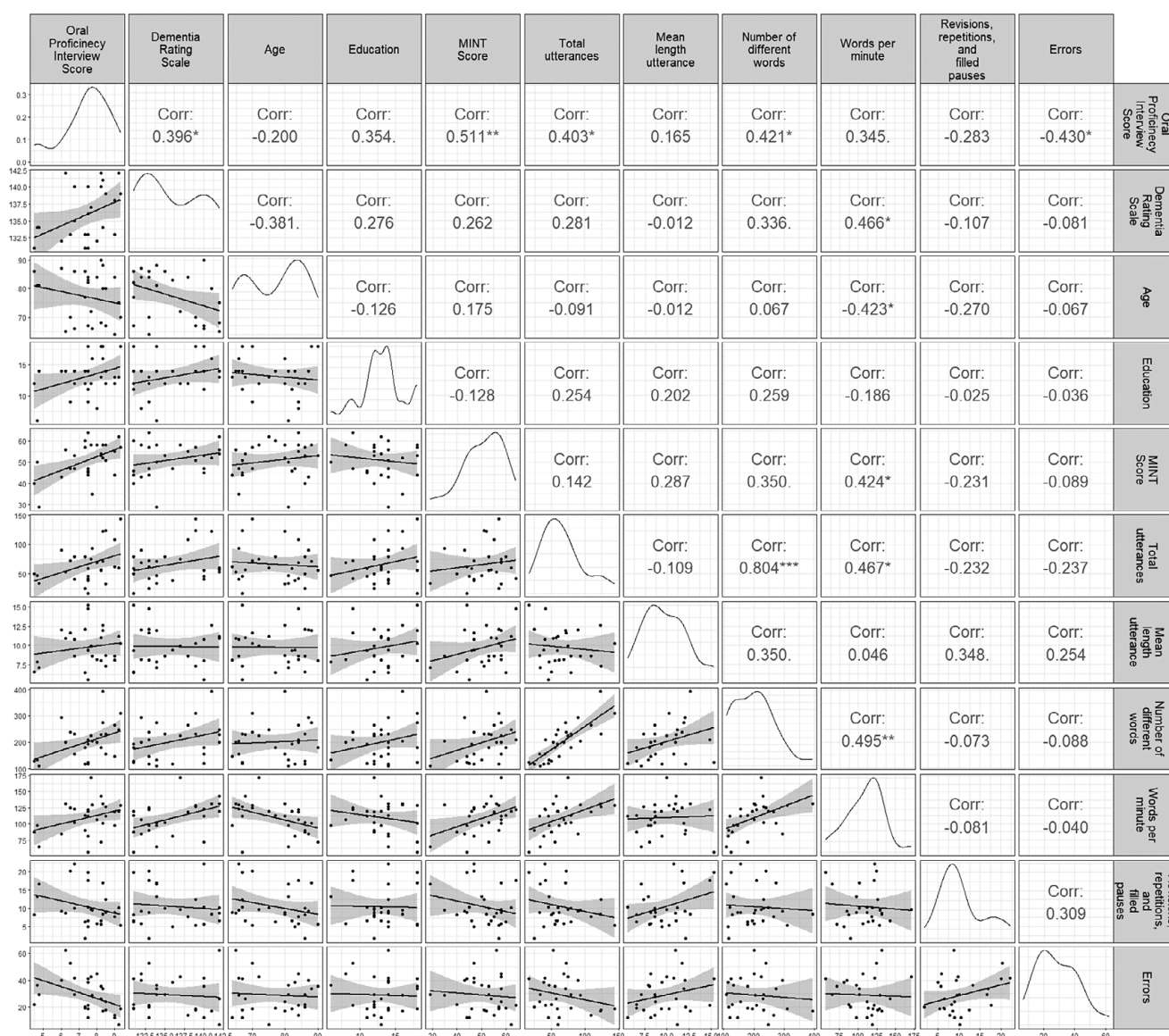
Figure 2. Nondominant language correlation table and scatter plots.
Note: ˙*p* < .10, *p < .05, **p < .01, ***p < .001.

Table 4. Backwards regression results using demographic and linguistic variables to predict Oral Proficiency Interview scores in the nondominant language

Independent variables	B	SE	Beta(β)	t	p
(Constant)	−.002	1.52		−0.01	0.989
Education	0.23	0.07	0.52	3.46	0.002
MINT score	0.10	0.02	0.59	4.21	<.001
Errors	−0.06	0.02	−0.34	−2.24	0.035

Note: The table below shows the final model after removing total utterances and number of different words. R^2 = .569, $F(3.22)$ = 9.668, $p < .001$.

Table 5. Backwards regression results include variables that were significantly correlated with speaking speed in the (a.) dominant language and (b.) nondominant language

a. Dominant language	B	SE	Beta(β)	t	p
(Constant)	−309.04	151.01		−2.05	.052
Dementia rating scale	3.21	1.11	.51	2.89	.008
b. Nondominant language					
(Constant)	−302.26	159.45		−1.90	.07
Dementia rating scale	3.02	1.17	.47	2.58	.017

Note: Only DRS remained as a significant predictor in both models, while age was removed. R^2 = .258, $F(1.24)$ = 8.36, $p = .008$ and R^2 = .217, $F(1.24)$ = 6.64, $p = .017$, respectively.

(a) people who produced a greater number of utterances, also produced a greater amount of different words in both the dominant ($r = .803$, $p < .001$) and nondominant languages ($r = .804$, $p < .001$), (b) people who produced a greater number of utterances also spoke faster in both the dominant ($r = .501$, $p = .007$) and the nondominant languages ($r = .467$, $p = .012$), and c) people who talked fast also used many different words in both the dominant ($r = .547$, $p = .003$) and nondominant language ($r = .495$ $p = .007$).
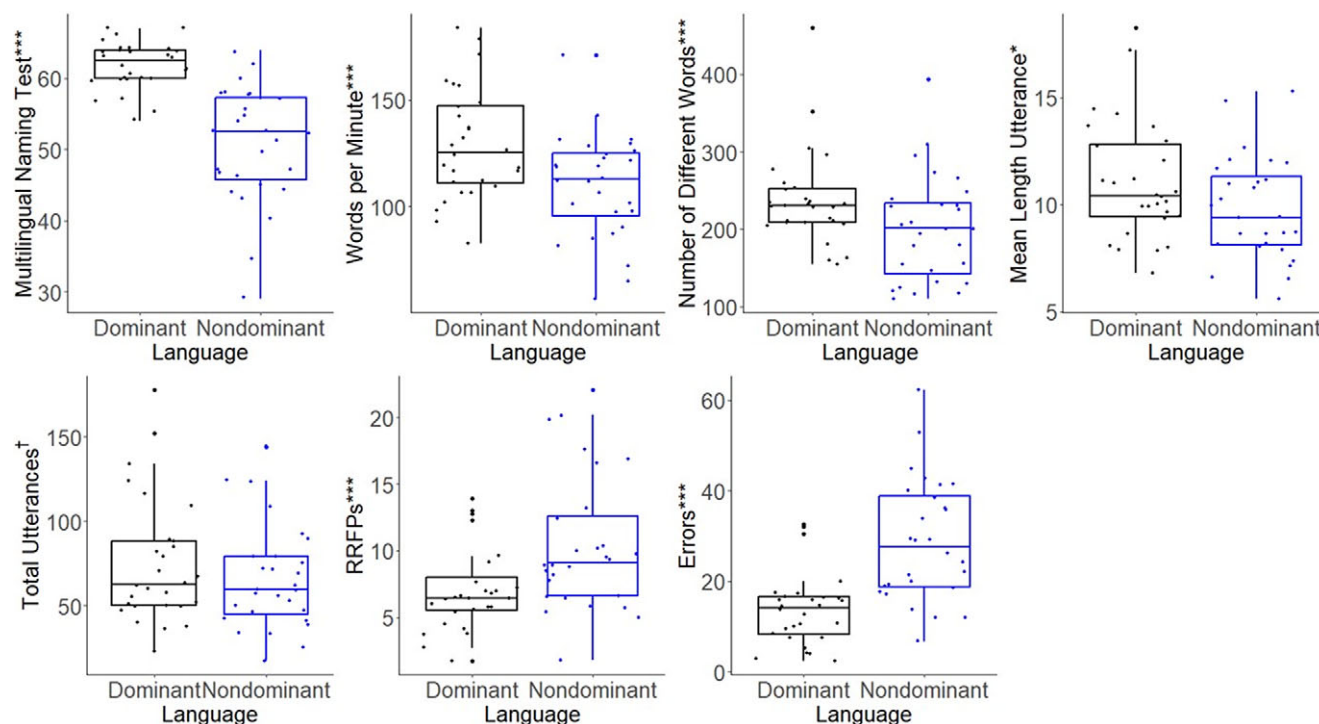
Figure 3. Mean and interquartile range for each linguistic submeasure by language.
Note: $^†p < .10$, $^*p < .05$, $^{**}p < .01$, $^{***}p < .001$.

4. Discussion

In this study, we asked which factors predict proficiency in both languages of older Spanish-English bilinguals. In the dominant language, older-old speakers were rated as less proficient and spoke more slowly than younger-old speakers, but age effects were not significant after accounting for cognitive functioning scores. No other measure was significantly correlated with OPI scores in the dominant language with just one exception that ran opposite to the expected direction (speakers who produced more RRFPs were rated as *more* proficient, though this was primarily driven by revisions). In the nondominant language, age was not significantly correlated with overall proficiency scores, and speakers with higher cognitive functioning scores were rated as more proficient and spoke marginally faster but not after accounting for productive vocabulary (MINT) scores, which were strong predictors of proficiency (see also De Jong et al., 2012). The ability to produce error-free speech and higher education level also explained unique variance in predicting overall proficiency scores in the nondominant language.

Additionally, we observed both persistence of individual traits within bilinguals' two languages and significant language dominance effects on those same traits. Most linguistic submeasures were significantly correlated across languages (see Figure 4) except vocabulary scores and production of errors, but virtually all measures exhibited significant and highly robust language dominance effects (see Figure 3). Finally, although different linguistic submeasures predicted proficiency in each language, in both languages, fast talkers also tended to speak more and with greater linguistic diversity (see similar within language correlations in Bedore et al., 2010).

4.1. Aging versus cognition effects

Previous studies reported significant aging-related changes in how people talk but few asked if age effects might have reflected other cognitive changes associated with aging. Of the six aging studies reviewed above, half did not measure cognitive ability (Breier et al., 2023; Horton et al., 2010; James et al., 1998), and two used just a brief cognitive screener (Pereria et al., 2019; Marini et al., 2005). One used a more thorough test (used by the Canadian armed forces) and like this study found that age effects were not significant after controlling for cognitive ability (Gold et al., 1988). The present study used two measures to examine cognitive functioning. The MMSE has 11 questions with a maximum score of 30, while the DRS has 36 questions that measure functioning across 5 cognitive domains with a maximum score of 144. In this study, only the DRS significantly predicted proficiency scores. Our results suggest that it is critical to use more thorough cognitive measures to distinguish the effects of age from subtle changes in cognitive functioning that may be present in some but not in other older participants.

A unique aspect of this study was the fact that some participants were followed longitudinally (see also Beier et al., 2023), and subsequently, several participants were diagnosed with dementia. Thus, it might seem that aging effects would have been more robust if our sample had a lower rate of participants with subsequent cognitive decline. To consider this possibility, we repeated the analyses predicting OPI scores in the dominant language (see Table 3) excluding the participants ($n = 8$) who were later diagnosed with AD. Analysis of this subset ($n = 20$) provided no indication that aging effects might be more robust without participants who declined in subsequent years; DRS scores ($r = .666$, $p = .003$) were still highly correlated, and age was marginally correlated
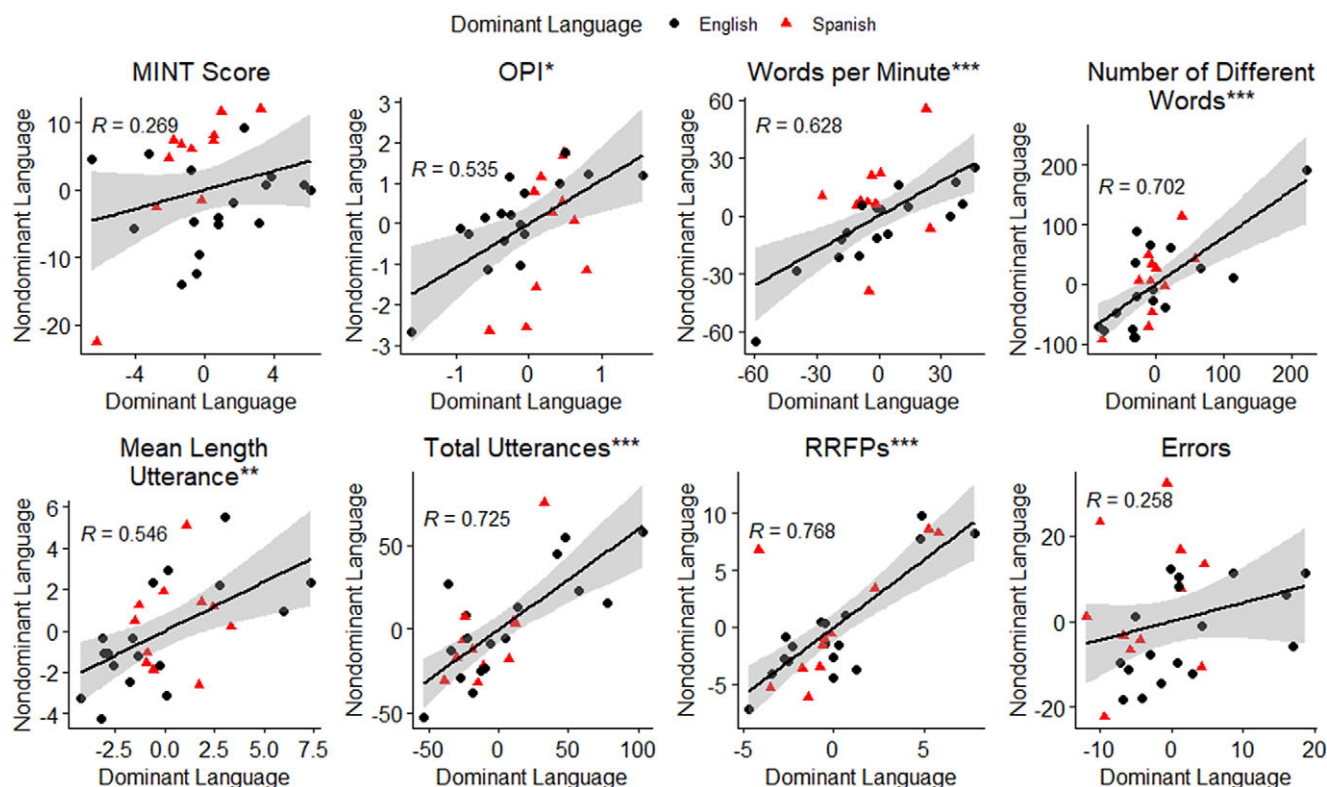
Figure 4. Partial correlations between the linguistic submeasures in the dominant and nondominant languages controlling for age. To depict the strength of the correlation, the residuals of a univariate ANOVA for each variable are plotted below.
Note: *$p < .05$, **$p < .01$, ***$p < .001$.

($r = .407$, $p = .075$) with OPI scores in the dominant language, and DRS scores ($B = .103$, $t(17) = 2.898$, $p = .012$) but not age effects ($B = −.019$, $t(17) = −1.147$, $p = .271$) significantly predicted dominant language OPI scores when entered together in a single model.[2] Given the high base rates of AD pathology (46%) in adults over 70 years old (Kern et al., 2018), and the rapid increase in pathology prevalence after age 70 (prevalence doubles every 5 years from ages 65–90; Corrada et al., 2010), it seems possible that some previously reported aging effects may have reflected subtle cognitive decline associated with preclinical AD and should not have been attributed to healthy aging (McDonough et al., 2016; Sliwinski et al., 1996).

4.2. What is proficiency in the dominant versus in the nondominant language?

A major question of interest in this study was what cognitive and linguistic abilities underlie proficiency in the bilinguals' two languages. Although cognitive ability was significantly correlated with OPI scores in both languages, backward regression models revealed no overlap in which predictors survived to best predict proficiency in each language. An important consideration is the more restricted range of scores in the dominant versus the nondominant languages. In the dominant language, OPI scores varied between 7.0 and 10.0 (with all but one score between 8 and 10, with *SD*s of 0.6–0.7). By contrast, in the nondominant language, OPI scores varied between 4.5 and 9.3 (and *SD*s were larger; see Table 1). Linguistic submeasures also showed less variability in the dominant than in

the nondominant language. While the differences in proficiency ranges between languages may have contributed to different measures predicting proficiency in each language, additional possible reasons why we obtained differences between languages with respect to which measures best predicted overall OPI scores can be considered.

The DRS was administered in the bilinguals' dominant language. As such, it might seem that the robust correlation between the DRS and overall OPI scores in the dominant language could be an artifact of linguistic submeasures in the DRS. In particular, the DRS includes one semantic fluency trial (supermarket items), which is known to be correlated with OPI scores (Neveu et al., submitted). To test if the correlation between DRS and OPI scores might have been driven by semantic fluency scores (rather than by cognitive abilities in general), we examined if the number of supermarket items produced in the semantic fluency trial in the DRS was itself correlated with OPI scores. This analysis was restricted to a subset of 22 participants for whom we had individual DRS subtest scores. In this subset of participants, DRS scores were still correlated with OPI scores in the dominant language ($r = .492$, $p = .02$).[3] Critically, the number of supermarket items produced was not significantly correlated with OPI scores in the dominant language ($r = −.116$, $p = .607$). Thus, we suggest that the explanatory power of the DRS for explaining individual differences in OPI scores was likely not driven by the use of the same language to administer the DRS. Instead, other aspects of cognitive functioning or the combination of several domains of cognition (see description of DRS

---

[2]The effect of education remained nonsignificant ($B = .069$, $t(25) = 1.507$, $p = .154$, total $R^2 = .569$, $F = 6.171$).

[3]This is important because the fluency trial in the DRS is capped at a maximum number of 20 points, which might have reduced its sensitivity.

subtests above) may be more strongly predictive of the ability to proficiently formulate and produce spoken language as needed to express thoughts in the dominant language.

Another unexpected predictor of overall OPI scores in the dominant language was the production of revisions, repetitions, and filled pauses; more proficient speakers produced more RRFPs. This is a counterintuitive result, given that RRFPs are types of disfluencies. For example, Beier et al. (2023) suggested that increased repetitions reflect processing difficulties associated with increased age. Production of more RRFPs in the nondominant language was also associated with increased likelihood of later being diagnosed with Alzheimer's disease (Garcia & Gollan, 2024). To better understand why RRFPs predicted proficiency in the dominant language, we examined revisions, repetitions, and filled pauses separately. Bilinguals who produced more revisions in the dominant language were rated as more proficient (r = .482, p = .009). Repetitions and filled pauses were not significantly correlated with OPI scores (ps ≥ .16). Revisions may have been used to add new or more complex information to utterances, thus leading to higher ratings. Revising one's speech requires the speaker to detect a problem and formulate a replacement to correct it, and producing revisions may be used to maintain speech continuity for the ease of the listener (Lake et al., 2011). A last possibility is that increased disfluencies in highly proficient speakers could reflect the use of more grammatically complex structures (but note this correlation only approached significance, r = .360, p = .06; see Figure 2).

Proficiency in the nondominant language was predicted by measures of productive vocabulary (MINT), errors, and education level. It is counterintuitive that education predicted proficiency only in the nondominant language, given that bilinguals were educated primarily in their dominant language. However, this finding is consistent with previous work in which older adults with higher education levels were more likely to learn and maintain a novel language (Blumenfeld et al., 2017). The correlation between MINT scores and OPI scores in the nondominant language also replicates previous findings (Gollan et al., 2012; Garcia & Gollan, 2021) and demonstrates that vocabulary is central to proficiency. Without adequate vocabulary knowledge in the target language, speakers may fail to communicate or make errors in order to do so (De Jong et al., 2012). Finally, the production of errors explained unique variance in predicting OPI scores only in the nondominant language. Errors produced in the dominant language might have primarily reflected accidental slips (that all speakers produce occasionally regardless of proficiency level), while in the nondominant language, the production of errors was more likely to reflect a lack of knowledge/proficiency. Thus, speaking in the nondominant language requires adequate knowledge in both lexical and grammatical domains. While this is of course, also true in the dominant language, these abilities are more likely to be at or closer to ceiling levels in the dominant language, which opens the possibility for other predictors (e.g., cognitive ability) to explain more variance.

4.3. Evidence for individual differences: Why do people talk how they talk?

Apart from speech errors and MINT scores, which constitute relatively pure measures of language-specific knowledge, all linguistic submeasures of the OPI were highly correlated across languages. Even OPI scores themselves were significantly correlated across languages, and this correlation was significant even when controlling for DRS scores (r = .444, p = .026; as were all the between

language correlations shown in Figure 2 with rs ranging from .540 to .784, and ps from .01 to < .001). Thus, it seems that many aspects of language production are subject to individual differences that will be displayed in both languages for a given speaker. The robust relationship between these measures, even after controlling for DRS scores, could suggest that between-language correlations in OPI submeasures are driven by factors unrelated to cognitive functioning per se (e.g., aspects of personality; Feldstein & Sloan, 1984), or that the DRS does not measure all relevant aspects of cognitive functioning (the DRS was designed to differentiate cognitively healthy from impaired older adults; Monsch et al., 1995, and to stage dementia progression; Vitaliano et al., 1984).

However, previous studies with young adult talkers also reported significant between-language correlations for speech rate (Bradlow et al., 2017, 2018, 2019). Reading speed is also correlated between languages for bilinguals. Kuperman et al. (2023) examined eye movements and the ability to answer comprehension questions in bilingual readers of English as a second language with 11 different native languages and found that reading rate was highly correlated across languages (r = .75, p < .001), but the ability to answer comprehension questions was much more language specific. This resembles our finding of nonsignificant correlations between languages for MINT scores and the production of errors. Kuperman and colleagues proposed that individual bilinguals use the same reading strategy in both their languages (e.g., frequency of word skipping), while reading comprehension relies on a different set of skills (i.e., language-specific skills). Bradlow and colleagues recorded Mandarin-English and Korean-English bilinguals producing sentences in each of their languages. Native speakers then listened to the productions in a sentence-in-noise recognition task, which allowed researchers to calculate an intelligibility score for each talker. Intelligibility scores were significantly lower in the second language than in the native language, but intelligibility scores were also correlated across languages; people who produced more intelligible speech in their native language also produced more intelligible speech in their second language (Bradlow et al., 2018). The generalization of cross-linguistic correlations across such different tasks (from silent reading to linguistic submeasures in an open-ended interview) and across multiple age groups, invites further investigation and supports Bradlow et al.'s (2018) proposal that such correlations may reflect central cognitive processes and deeper levels of planning that are not language specific (Bedore et al., 2010).

Importantly, while we observed high correlations between languages (see Figure 4), the linguistic submeasures were also highly sensitive to language dominance (see Figure 3). In previous work, we assumed the OPI is the gold standard because it is more comprehensive than most objective measures (e.g., naming tests only measure expressive vocabulary and lexical decision tests only measure the ability to accurately distinguish between real words and nonwords, but neither of these provide information about grammatical knowledge). To earn the highest score (10), speakers must function "like a highly educated native speaker," (Gollan et al., 2012) raising the possible criticism that the OPI measures education level as much as it does proficiency level. Although education was only marginally correlated with OPI scores in both languages, it was a robust predictor of proficiency in the nondominant language in the backward regression analysis. But this is not necessarily a problem given that both the ability to communicate in everyday scenarios and the ability to communicate complex ideas must be evaluated to obtain a complete assessment of a speaker's communicative competence. Thus, although the OPI assesses the ability to function in

everyday scenarios and also pushes bilinguals to use higher-order language skills that may not be necessary in day-to-day life (Hulstijn, 2024), this is necessary to determine the range of topics in which a speaker can function. High levels of education may typically be needed to communicate about topics that are not needed in daily life. Note, however, that if so, education effects should have been significant in both languages in this study. Thus, an alternative possibility is that more educated individuals are also more likely to have or actively seek opportunities to gain proficiency in a nondominant language (Neveu & Gollan, 2024). In addition to providing a comprehensive assessment of proficiency level, another advantage of the OPI is that it is relatively straightforward to match across languages for difficulty, which is far from trivial (Hulstijn, 2012; Peña, 2007). Similar to previous studies, we found that more proficient bilinguals spoke faster (Baese-Berk & Morrill, 2015; Guion et al., 2000), with more different words (Treffers-Daller, 2011), with longer utterances, and with fewer errors, repetitions, revisions, and filled pauses. The results of this study thus validate the use of the OPI as a comprehensive measure of language proficiency.

5. Limitations

Though the OPI provides a rich language sample, and each speaker provided well above the amount of speech considered necessary to conduct the types of analyses reported herein (i.e., 50 utterances or more; see Pezold et al., 2020) we had a relatively small number of participants. A larger study might detect both significant age effects and effects of cognitive functioning. Similarly, our relatively small sample size might explain why so few variables explained unique variance in predicting overall OPI scores in the dominant language (although note that our sample was quite heterogeneous with respect to age and education level). Though detailed analysis of spontaneous speech is labor intensive, more automated methods of analysis are emerging and may shed additional light (but see Hitczenko et al., 2021) on relationships between aging and speech production, particularly if robust measures of cognitive functioning can be included in such studies.

6. Conclusion

Together with previous studies, the results of this study suggest that proficient language production reflects the combined forces of language-specific and talker-specific characteristics, including higher-level cognitive processing and at all processing levels, from formulation to articulation across the lifespan. Similar language-independent individual differences in cognitive processing may underlie differences between individual talkers, regardless of idiosyncratic properties of the specific languages, whether for producing intelligible speech or coherent, highly rated, spoken utterances during an interview. Additionally, rather than reflecting age-related changes per se, some changes in how people talk that were previously associated with aging (e.g., slowed speech) might instead reflect changes in cognitive functioning, which can be detected with more thorough cognitive assessment (but is seldom included in studies of cognitive aging). Finally, while some aspects of speaking in the nondominant language might be sensitive to future cognitive decline in bilinguals (Garcia & Gollan, 2024), overall production of the *dominant* language is more closely tied to cognitive functioning, whereas production of the nondominant language reflects language-specific knowledge.

References

Ahmed, S., Haigh, A.-M. F., de Jager, C. A., & Garrard, P. (2013). Connected speech as a marker of disease progression in autopsy-proven Alzheimer's disease. *Brain*, **136**(12), 3727–3737. https://doi.org/10.1093/brain/awt269

Baese-Berk, M. M., & Morrill, T. H. (2015). Speaking rate consistency in native and non-native speakers of English. *The Journal of the Acoustical Society of America*, **138**(3), EL223–EL228. https://doi.org/10.1121/1.4929622

Bedore, L. M., Fiestas, C. E., Peña, E. D., & Nagy, V. J. (2006). Cross-language comparisons of maze use in Spanish and English in functionally monolingual and bilingual children. *Bilingualism: Language and Cognition*, **9**(3), 233–247. https://doi.org/10.1017/S1366728906002604

Bedore, L. M., Peña, E. D., Gillam, R. B., & Ho, T.-H. (2010). Language sample measures and language ability in Spanish-English bilingual kindergarteners. *Journal of Communication Disorders*, **43**(6), 498–510. https://doi.org/10.1016/j.jcomdis.2010.05.002

Beier, E. J., Chantavarin, S., & Ferreira, F. (2023). Do disfluencies increase with age? Evidence from a sequential corpus study of disfluencies. *Psychology and Aging*, **38**(3), 203–218. https://doi.org/10.1037/pag0000741

Berisha, V., Wang, S., LaCross, A., & Liss, J. (2015). Tracking discourse complexity preceding Alzheimer's disease diagnosis: A case study comparing the press conferences of presidents ronald reagan and george herbert walker bush. *Journal of Alzheimer's Disease*, **45**(3), 959–963. https://doi.org/10.3233/JAD-142763

Blumenfeld, H. K., Quinzon, S. J. R., Alsol, C., & Riera, S. A. (2017). Predictors of successful learning in multilingual older adults acquiring a majority language. *Frontiers in Communication*, **2**, 23. https://doi.org/10.3389/fcomm.2017.00023

Bradlow, A. R., Kim, M., & Blasingame, M. (2017). Language-independent talker-specificity in first-language and second-language speech production by bilingual talkers: L1 speaking rate predicts L2 speaking rate. *The Journal of the Acoustical Society of America*, **141**(2), 886–899. https://doi.org/10.1121/1.4976044

Bradlow, A. R., Blasingame, M., & Lee, K. (2018). Language-independent talker-specificity in bilingual speech intelligibility: Individual traits persist across first-language and second-language speech. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, **9**(1), 17. https://doi.org/10.5334/labphon.137

Bradlow, A. R. (2019). Speaking Rate, Information Density, and Information Rate in First-Language and Second-Language Speech. *Interspeech* **2019**, 3559–3563. https://doi.org/10.21437/Interspeech.2019-1150

Corrada, M. M., Brookmeyer, R., Paganini-Hill, A., Berlau, D., & Kawas, C. H. (2010). Dementia incidence continues to increase with age in the oldest old: The 90+ study. *Annals of Neurology*, **67**(1), 114–121. https://doi.org/10.1002/ana.21915

De Jong, N. H., Steinel, M. P., Florijn, A. F., Schoonen, R., & Hulstijn, J. H. (2012). Facets of speaking proficiency. *Studies in Second Language Acquisition*, **34**(1), 5–34. https://doi.org/10.1017/S0272263111000489

Duchin, S. W., & **Mysak, E. D**. (1987). Disfluency and rate characteristics of young adult, middle-aged, and older males. *Journal of Communication Disorders*, **20**(3), 245–257. https://doi.org/10.1016/0021-9924(87)90022-0

Engelhardt, P. E., **McMullon, M. E**., & **Corley, M**. (2019). Individual differences in the production of disfluency: A latent variable analysis of memory ability and verbal intelligence. *Quarterly Journal of Experimental Psychology*, **72**(5), 1084–1101. https://doi.org/10.1177/1747021818778752

Engelhardt, P. E., **Nigg, J. T**., & **Ferreira, F**. (2013). Is the fluency of language outputs related to individual differences in intelligence and executive function? *Acta Psychologica*, **144**(2), 424–432. https://doi.org/10.1016/j.actpsy.2013.08.002

Fagan, W. T. (1982). The relationship of the "Maze" to language planning and production. *Research in the Teaching of English*, **16**(1), 85–95.

Feldstein, S., & **Sloan, B**. (1984). Actual and stereotyped speech tempos of extraverts and introverts. *Journal of Personality*, **52**(2), 188–204. https://doi.org/10.1111/j.1467-6494.1984.tb00352.x

Folstein, M. F., **Folstein, S. E**., & **McHugh, P. R**. (1975). Mini-mental state. *Journal of Psychiatric Research*, **12**(3), 189–198. https://doi.org/10.1016/0022-3956(75)90026-6

Garcia, D. L., & **Gollan, T. H**. (2021). The MINT Sprint: Exploring a Fast Administration Procedure with an Expanded Multilingual Naming Test. *Journal of the International Neuropsychological Society*: **JINS**, 1–17. https://doi.org/10.1017/S1355617721001004

Garcia, D. L., & **Gollan, T. H**. (2024). Language switching and speaking a nondominant language challenge executive control: Preliminary data for novel behavioral markers of Alzheimer's risk in Spanish–English bilinguals. *Neuropsychology*, **38**(4), 322–336. https://doi.org/10.1037/neu0000943

Gold, D., **Andres, D**., **Arbuckle, T**., & **Schwartzman, A**. (1988). Measurement and Correlates of Verbosity in Elderly People. *Journal of Gerontology*, **43**(2), P27–P33. https://doi.org/10.1093/geronj/43.2.P27

Gollan, T. H., **Weissberger, G. H**., **Runnqvist, E**., **Montoya, R. I**., & **Cera, C. M**. (2012). Self-ratings of spoken language dominance: A Multilingual Naming Test (MINT) and preliminary norms for young and aging Spanish–English bilinguals. *Bilingualism: Language and Cognition*, **15**(3), 594–615. https://doi.org/10.1017/S1366728911000332

Guion, S. G., **Flege, J. E**., **Liu, S. H**., & **Yeni-Komshian, G. H**. (2000). Age of learning effects on the duration of sentences produced in a second language. *Applied Psycholinguistics*, **21**(2), 205–228. https://doi.org/10.1017/S0142716400002034

Hanulová, J., **Davidson, D. J**., & **Indefrey, P**. (2011). Where does the delay in L2 picture naming come from? Psycholinguistic and neurocognitive evidence on second language word production. *Language and Cognitive Processes*, **26**(7), 902–934. https://doi.org/10.1080/01690965.2010.509946

Hitczenko, K., **Mittal, V. A**., & **Goldrick, M**. (2021). Understanding language abnormalities and associated clinical markers in psychosis: The promise of computational methods. *Schizophrenia Bulletin*, **47**(2), 344–362. https://doi.org/10.1093/schbul/sbaa141

Horton, W. S., **Spieler, D. H**., & **Shriberg, E**. (2010). A corpus analysis of patterns of age-related change in conversational speech. *Psychology and Aging*, **25**(3), 708–713. https://doi.org/10.1037/a0019424

Hulstijn, J. (2024). Predictions of individual differences in the acquisition of native and non-native languages: An update of BLC theory. *Languages*, **9**(5), 173. https://doi.org/10.3390/languages9050173

Hulstijn, J. H. (2012). The construct of language proficiency in the study of bilingualism from a cognitive perspective. *Bilingualism: Language and Cognition*, **15**(2), 422–433. https://doi.org/10.1017/S1366728911000678

James, L. E., **Burke, D. M**., **Austin, A**., & **Hulme, E**. (1998). Production and perception of "verbosity" in younger and older adults. *Psychology and Aging*, **13**(3), 355–367. https://doi.org/10.1037/0882-7974.13.3.355

Kavé, G., & **Goral, M**. (2016). Word retrieval in picture descriptions produced by individuals with Alzheimer's disease. *Journal of Clinical and Experimental Neuropsychology*, **38**(9), 958–966. https://doi.org/10.1080/13803395.2016.1179266

Kern, S., **Zetterberg, H**., **Kern, J**., **Zettergren, A**., **Waern, M**., **Hoglund, K**., **Andreasson, U**., **Wetterberg, H**., **Borjesson-Hanson, A**., **Blennow, K**., & **Skoog, I**. (2018). Prevalence of preclinical Alzheimer disease: Comparison of current classification systems. *Neurology*, **90**(19). https://doi.org/10.1212/WNL.0000000000005476

Kuperman, V., **Siegelman, N**., **Schroeder, S**., **Acartürk, C**., **Alexeeva, S**., **Amenta, S**., **Bertram, R**., **Bonandrini, R**., **Brysbaert, M**., **Chernova, D**., **Da Fonseca, S. M**., **Dirix, N**., **Duyck, W**., **Fella, A**., **Frost, R**., **Gattei, C. A**., **Kalaitzi, A**., **Lõo, K**., **Marelli, M**., … **Usal, K. A**. (2023). Text reading in English as a second language: Evidence from the multilingual eye-movements corpus. *Studies in Second Language Acquisition*, **45**(1), 3–37. https://doi.org/10.1017/S0272263121000954

Lake, J. K., **Humphreys, K. R**., & **Cardy, S**. (2011). Listener vs. speaker-oriented aspects of speech: Studying the disfluencies of individuals with autism spectrum disorders. *Psychonomic Bulletin & Review*, **18**(1), 135–140. https://doi.org/10.3758/s13423-010-0037-x

Loban, W. (1976). Language development: Kindergarten through grade twelve *(18)*. National Council of Teachers of English.

MacKenzie, C. (2000). Adult spoken discourse: The influences of age and education. *International Journal of Language & Communication Disorders*, **35**(2), 269–285. https://doi.org/10.1080/136828200247188

Marini, A., **Boewe, A**., **Caltagirone, C**., & **Carlomagno, S**. (2005). Age-related differences in the production of textual descriptions. *Journal of Psycholinguistic Research*, **34**(5), 439–463. https://doi.org/10.1007/s10936-005-6203-z

Mattis, S. (1988). *Dementia rating scale: Professional manual. Published Online.*

McDonough, I. M., **Bischof, G. N**., **Kennedy, K. M**., **Rodrigue, K. M**., **Farrell, M. E**., & **Park, D. C**. (2016). Discrepancies between fluid and crystallized ability in healthy adults: A behavioral marker of preclinical Alzheimer's disease. *Neurobiology of Aging*, **46**, 68–75. https://doi.org/10.1016/j.neurobiolaging.2016.06.011

Miller, J., & **Iglesias, A**. (2020). *Systematic analysis of language transcripts (SALT)* (Version 20 [Computer Software]).Middleton, WI: SALT Software, LLC.

Monsch, A. U., **Bondi, M. W**., **Salmon, D. P**., **Butters, N**., **Thal, L. J**., **Hansen, L. A**., **Weiderholt, W. C**., **Cahn, D. A**., & **Klauber, M. R**. (1995). Clinical Validity of the Mattis Dementia Rating Scale in Detecting Dementia of the Alzheimer Type: A Double Cross-Validation and Application to a Community-Dwelling Sample. *Archives of Neurology*, **52**(9), 899. https://doi.org/10.1001/archneur.1995.00540330081018

Mortensen, L., **Meyer, A. S**., & **Humphreys, G. W**. (2006). Age-related effects on speech production: A review. *Language and Cognitive Processes*, **21**(1–3), 238–290. https://doi.org/10.1080/01690960444000278

Mulder, K., & **Hulstijn, J. H**. (2011). Linguistic skills of adult native speakers, as a function of age and level of education. *Applied Linguistics*, **32**(5), 475–494. https://doi.org/10.1093/applin/amr016

Neveu, A., & **Gollan, T. H**. (2024). New insights on what leads bilinguals to be able to name some pictures only in their nondominant language: Immersion, dominance reversal, and balanced bilingualism. *Journal of the International Neuropsychological Society*, **30**(6), 546–552. https://doi.org/10.1017/S1355617724000067

Peña, E. D. (2007). Lost in translation: Methodological considerations in cross-cultural research. *Child Development*, **78**(4), 1255–1264. https://doi.org/10.1111/j.1467-8624.2007.01064.x

Pereira, N., **Gonçalves, A. P. B**., **Goulart, M**., **Tarrasconi, M. A**., **Kochhann, R**., & **Fonseca, R. P**. (2019). Age-related differences in conversational discourse abilities A comparative study. *Dementia & Neuropsychologia*, **13**(1), 53–71. https://doi.org/10.1590/1980-57642018dn13-010006

Pezold, M. J., **Imgrund, C. M**., & **Storkel, H. L**. (2020). Using computer programs for language sample analysis. *Language, Speech, and Hearing Services in Schools*, **51**(1), 103–114. https://doi.org/10.1044/2019_LSHSS-18-0148

Pirvulescu, M., **Pérez-Leroux, A.-T**., **Roberge, Y**., **Strik, N**., & **Thomas, D**. (2014). Bilingual effects: Exploring object omission in pronominal languages. *Bilingualism: Language and Cognition*, **17**(3), 495–510. https://doi.org/10.1017/S1366728913000631

Pushkar, D., **Basevitz, P**., **Arbuckle, T**., **Nohara-LeClair, M**., **Lapidus, S**., & **Peled, M**. (2000). Social behavior and off-target verbosity in elderly people. *Psychology and Aging*, **15**(2), 361–374. https://doi.org/10.1037/0882-7974.15.2.361

Ramig, L. A. (1983). Effects of physiological aging on speaking and reading rates. *Journal of Communication Disorders*, **16**(3), 217–226. https://doi.org/10.1016/0021-9924(83)90035-7

Reeves, S. M., Williams, V., Blacker, D., & Woods, R. L. (2023). Further evaluation of narrative description as a measure of cognitive function in Alzheimer's disease. *Neuropsychology*, **37**(7), 801–812. https://doi.org/10.1037/neu0000884

Runnqvist, E., Strijkers, K., Sadat, J., & Costa, A. (2011). On the temporal and functional origin of L2 disadvantages in speech production: A critical review. *Frontiers in Psychology*, **2** 1–8. https://doi.org/10.3389/fpsyg.2011.00379

Sliwinski, M., Lipton, R. B., Buschke, H., & Stewart, W. (1996). The effects of preclinical dementia on estimates of normal cognitive functioning in aging. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, **51B**(4), P217–P225. https://doi.org/10.1093/geronb/51B.4.P217

Treffers-Daller, J. (2011). Operationalizing and measuring language dominance. *International Journal of Bilingualism*, **15**(2), 147–163. https://doi.org/10.1177/1367006910381186

Tremblay, P., Sato, M., & Deschamps, I. (2017). Age differences in the motor control of speech: An fMRI study of healthy aging. *Human Brain Mapping*, **38**(5), 2751–2771. https://doi.org/10.1002/hbm.23558

Vitaliano, P. P., Breen, A. R., Russo, J., Albert, M., Vitiello, M. V., & Prinz, P. N. (1984). The clinical utility of the Dementia Rating Scale for assessing Alzheimer patients. *Journal of Chronic Diseases*, **37**(9–10), 743–753. https://doi.org/10.1016/0021-9681(84)90043-2

Weissberger, G. H., Salmon, D. P., Bondi, M. W., & Gollan, T. H. (2013). Which neuropsychological tests predict progression to Alzheimer's disease in Hispanics? *Neuropsychology*, **27**(3), 343–355. https://doi.org/10.1037/a0032399

Weissberger, G. H., Wierenga, C. E., Bondi, M. W., & Gollan, T. H. (2012). Partially overlapping mechanisms of language and task control in young and older bilinguals. *Psychology and Aging*, **27**(4), 959–974. https://doi.org/10.1037/a0028281