

Article

# Measuring Language Dominance in Early Spanish/English Bilinguals

Cecilia Solís-Barroso <sup>1</sup>  and Sara Stefanich <sup>2,\*</sup>

<sup>1</sup> Department of Hispanic and Italian Studies, University of Illinois at Chicago, Chicago, IL 60607, USA

<sup>2</sup> Department of Spanish and Portuguese, Northwestern University, Evanston, IL 60208, USA

\* Correspondence: sara.stefanich@northwestern.edu

Received: 31 May 2019; Accepted: 2 August 2019; Published: 13 August 2019



**Abstract:** This paper analyzes the comparability of language dominance assessments with the purpose of determining whether they yield similar results. Language dominance is an important construct in the field of bilingualism as it allows for a more thorough classification of bilinguals and is thought to play a role in both bilingual production and perception. Yet, there is no unified methodology for assessing language dominance. To that end, we ask the following research question: Do different language dominance measures predict the results of one another? Twenty-nine Spanish/English early bilinguals completed four language dominance assessments. Results indicate that three of the four assessments are highly correlated with each other while the fourth, a repetition task, is not significantly correlated with any of the assessments. Further, twenty of the participants were categorized differently across the individual measures; the more “balanced” a bilingual was, the greater likelihood of being categorized differently. These results indicate that certain language dominance assessments are not comparable with one another and suggest that it could be the case they do not even measure the same variable.

**Keywords:** bilingualism; language dominance; Spanish; methodology; heritage speakers

## 1. Introduction

The aim of this paper is to examine the comparability of different language dominance assessments in their ability to identify and classify bilinguals with respect to a dominant language. Language dominance, or “observed asymmetries of skill in or use of one language over the other” (Birdsong 2014, p. 374), is a factor that is used by researchers to account for and explain variation in bilingual behavior (i.e., production, perception, processing, etc.). For instance, language dominance has been used to account for variation in studies including but not limited to: the production and perception of mid vowel contrasts (Amengual and Chamorro 2015), differences in voice onset time (VOT) in codeswitching contexts (Antoniou et al. 2011; Tsui et al. 2019), use of null/overt pronominal subjects and pre/post overt subjects (Argyri and Sorace 2007), child bilingual acquisition (Yip and Matthews 2006), and codeswitching patterns (Pérez-Leroux et al. 2014). Further, language dominance is also a variable of interest for educators and administrators and clinical research (Gertken et al. 2014).

However, even though language dominance is an important factor in bilingualism research, there is currently no unified method for assessing language dominance in the field. Rather, it has been operationalized and measured via an array of different methods whose usage depends on the individual researcher/research team, phenomenon under study, and formal perspective/view that the study takes on bilingualism (e.g., psycholinguistic, theoretical, sociolinguistic, etc.). Further, the studies that have examined language dominance as a factor in bilingual behavior have produced contradictory results. To illustrate these differences in measuring language dominance and how it plays a role in bilingual outcomes, we review four studies (Amengual and Chamorro 2015; Antoniou et al. 2011; Argyri and

Sorace 2007; Tsui et al. 2019) that operationalize language dominance in different ways and that find differing results with respect to how language dominance affects a given linguistic phenomenon.<sup>1</sup> In these four studies, we will see that three different measures were used to classify bilinguals as dominant in a given language: (1) based on their self-ratings of proficiency, (2) the language of most exposure and use, and (3) the results of a sociolinguistic questionnaire. The results show both an effect for the dominant language (Amengual and Chamorro 2015; Argyri and Sorace 2007), and an effect for the non-dominant language (Antoniou et al. 2011; Tsui et al. 2019). Additionally, language dominance will be shown to affect both groups of bilinguals, i.e., bilinguals dominant in L<sub>A</sub> and bilinguals dominant in L<sub>B</sub> (Tsui et al. 2019) or only one group of bilinguals (Argyri and Sorace 2007).

First, Amengual and Chamorro (2015) studied the production and perception of Galician mid-vowel contrasts (e.g., /e/~ε/ and /o/~ɔ/) in early Spanish/Galician bilinguals. They measured language dominance via a sociolinguistic questionnaire, the Bilingual Language Profile (Gertken et al. 2014) and classified the bilinguals into two groups: Spanish dominant and Galician dominant. They found that the Spanish dominant bilinguals had a harder time perceiving and producing the Galician mid vowel contrasts in comparison to the Galician dominant bilinguals. In other words, for the Spanish dominant bilinguals, Amengual and Chamorro found an effect of the dominant language (here, Spanish) on the non-dominant language (here, Galician). As Amengual and Chamorro (2015) did not examine the Spanish of their participants, it remains to be seen whether a similar effect would have been found in the opposite direction (i.e., an effect of Galician on the Spanish of the Galician dominant bilinguals or on the Spanish of the Spanish dominant bilinguals).

Second, Antoniou et al. (2011) and Tsui et al. (2019) examined voice onset time (i.e., the length of time between the release of a stop consonant and the onset of voicing) in codeswitching contexts across two different groups of bilinguals. Antoniou et al. (2011) examined early Greek/English bilinguals who were all classified as English-dominant based upon their self-ratings (i.e., they rated their mastery of English higher than that of their Greek). They found a unidirectional effect such that when switching from Greek to English, the English VOT values were produced more Greek-like than in non-codeswitching contexts. No such effect was found for the Greek VOT values (i.e., they were not produced as English-like). In other words, in contrast to Amengual and Chamorro (2015), Antoniou et al. (2011) found an effect for the non-dominant language (here, Greek) on the dominant language (here, English). As Antoniou et al. (2011) did not examine Greek dominant bilinguals, it remains to be seen whether a similar effect for the non-dominant language would have been found when switching from English into Greek (i.e., an effect of English on the Greek VOT values).

Tsui et al. (2019) examined early Cantonese/English bilinguals who were classified into three different dominance groups on the basis of their self-ratings of language proficiency: (1) Cantonese dominant, (2) English dominant, and (3) balanced. Similar to Antoniou et al. (2011), Tsui et al. (2019) found an asymmetrical effect for the non-dominant language on the dominant language for both the Cantonese and English dominant bilinguals. When switching from English to Cantonese, the Cantonese dominant bilinguals produced Cantonese VOT values as more English-like, whereas when switching from Cantonese to English, the English dominant bilinguals produced English VOT values as more Cantonese-like. No effects were found when switching in the opposite direction for either group. Further, the balanced bilinguals did not display any effects on either their Cantonese or English VOT values. Based upon these results, we can hypothesize that Antoniou et al. (2011) would have also found an effect of English on the Greek VOT values if Greek dominant bilinguals had been examined.

Third, Argyri and Sorace (2007) examined whether language dominance could account for syntax-pragmatics interface phenomena, specifically the use of (1) null/overt pronominal subjects,

---

<sup>1</sup> We note that these four studies were purposefully selected to highlight the disparity in the ways in which language dominance is measured in different studies and how those studies find varying results.

(2) pre/post overt subjects, (3) object pronouns and (4) the structure of wh-embedded interrogatives, in early Greek/English bilingual children. The children were all eight years old at the time of testing. Language dominance was measured as the amount of exposure the bilinguals received in each language. In this case, the Greek dominant bilingual children were born and raised in Greece, whereas the English dominant bilingual children were born and raised in the UK. Similar to [Amengual and Chamorro \(2015\)](#), Argyri and Sorace found an asymmetrical effect of the dominant language on the non-dominant language, but only for the English dominant bilinguals. That is, the English dominant bilinguals extended their use of preverbal subjects in Greek, in contexts where pragmatically they were not the most felicitous option and also where they were grammatically inappropriate. No such effect was present in the Greek dominant bilinguals. [Argyri and Sorace \(2007\)](#) concluded that while language dominance does play a role in predicting crosslinguistic effects, it cannot be the sole factor as in this case no effect was found for Greek on the English of the Greek dominant bilinguals.

A comparison of the methods and results of these four studies raises the question as to why they found different results. Why is it that [Amengual and Chamorro \(2015\)](#) and [Argyri and Sorace \(2007\)](#) found an effect of the dominant language on the non-dominant language, but [Antoniou et al. \(2011\)](#) and [Tsui et al. \(2019\)](#) found an effect of the non-dominant language on the dominant language? Does it have to do with the language pair or linguistic phenomenon under study? [Amengual and Chamorro \(2015\)](#) examined the phonological phenomenon of Galician mid-vowel contrasts and Argyri and Sorace examined syntax-pragmatics interface phenomena, whereas [Antoniou et al. \(2011\)](#) and [Tsui et al. \(2019\)](#) examined the phonetic effects of codeswitching. Or could it have to do with the way in which language dominance was operationalized? [Amengual and Chamorro \(2015\)](#) used a sociolinguistic questionnaire (the BLP) to group participants into dominance groups, while [Antoniou et al. \(2011\)](#) and [Tsui et al. \(2019\)](#) both used self-ratings of proficiency, and [Argyri and Sorace \(2007\)](#) relied on amount of language exposure. Further, why is it that [Argyri and Sorace \(2007\)](#) only found an effect for one of the dominance groups (here, English dominant) but [Tsui et al. \(2019\)](#) found an effect for both groups (here, Cantonese dominant and English dominant)? Again, could it have to do with the linguistic phenomenon under study? Or could these differences in results have to do with the way the bilinguals were grouped into different dominance groups based upon the chosen language dominance measure? The combination of variation in language dominance assessments and the different results in studies that use them to examine language dominance (that these four studies highlight) is problematic because we have no way to determine whether any differences in results are due to the specific language dominance assessment used or another variable (e.g., language pair, type of bilingualism, linguistic phenomenon, etc.). This then makes it difficult to synthesize the effects of language dominance on bilingual behavior across different studies, which is not ideal as being able to compare the effect of language dominance across different studies is an essential step in understanding not only the bilingual construct of language dominance, but also bilingual outcomes. As there is such variation in the way that language dominance is measured (see [Section 3](#) for more details), one question that arises is whether all these methods are comparable and whether we would get the same results regardless of the method. In this paper, we compare four different but commonly used language dominance assessments in order to shed light on this issue. Specifically, we compare results of the Bilingual Language Profile ([Gertken et al. 2014](#)), the Bilingual Dominance Scale ([Dunn and Tree 2009](#)), self-ratings of verbal and written proficiency and a sentence repetition task ([Flege et al. 2002](#)) on a single group of early Spanish/English bilinguals in order to determine if these different assessments classify bilinguals in the same way. In other words, does it make a difference which language dominance assessment is used? We conclude that the language dominance assessment used does make a difference as our results show that the same bilingual will be classified into different groups (here, Spanish dominant, English dominant or balanced) based upon a given assessment. Further, our results suggest that the more “balanced” bilinguals are more difficult to consistently classify into a dominance group as they demonstrate the most variability in classification across the different assessments.

This article is organized as follows: in Section 2 we define language dominance and discuss the different ways it has been operationalized in the literature. Section 3 presents three previous studies that have compared different language dominance assessments and introduces our research question. In Section 4 we outline the methods used in our current study and describe the different assessments under review. Sections 5 and 6 present the results and discussion, respectively. In Section 7 we conclude the paper and illuminate avenues for future research.

## 2. What Is Language Dominance and How Has It Been Measured?

A wide-held assumption in bilingualism research is that bilinguals, rarely, if ever, demonstrate equal capacity in both of their languages, most likely as a result of the fact that they use their languages for different purposes, in different situations, and with different people (e.g., Grosjean 2016; Treffers-Daller 2019). This difference in a bilingual's linguistic ability (i.e., proficiency), language processing ability and/or language use is often characterized as language dominance (see Treffers-Daller 2019 for a review). That is, bilinguals are thought to be dominant in one of their languages with respect to the other. Recall from the introduction that Birdsong (2014) states "in the context of bilingualism, dominance refers to observed asymmetries of skill in, or use of one language over the other" (p. 374). Following Birdsong (2014), asymmetries of skill can be observed in linguistic competence, language production, and language processing (what he calls dimensions of dominance), whereas asymmetries of use can be considered the varying situations and contexts in which bilinguals use their two languages (what he calls domains of dominance). Further, these dimensions and domains of language dominance (to use Birdsong's terminology) are not entirely independent of one another.

Language dominance has been viewed as a link between the sociolinguistic factors of language use and language exposure and the psycholinguistic factors of language processing (e.g., Dubiel 2019; De Bot 2001). Dubiel (2019) explains, "the language that bilinguals are exposed to and use more frequently becomes the language in which they can access words without pauses and hesitations, and thus it is the language they are more dominant in." (p. 96). This connection between sociolinguistic and psycholinguistic factors can be seen clearly in the realm of child bilingualism where language dominance can be defined as "a situation where one of a child's languages is more advanced or developing faster than the other" (Yip and Matthews 2006, p. 4). The child's language that is more advanced or developing faster is the language to which they are the most exposed, which then causes it to be the language the child uses more (De Bot 2001). When measures are employed to determine the child's language dominance, the language of most exposure and most use becomes the language in which they perform "better" in terms of psycholinguistic factors such as reaction or response time. In adult bilinguals, support for a similar effect has been found by the Weaker Links Hypothesis (e.g., Gollan et al. 2008), and The Activation Threshold Hypothesis (Paradis 2004). The Weaker Links Hypothesis claims that due to not being able to speak both languages at the same time, bilinguals inevitably use their two languages less than a monolingual uses his/her one language. This reduced use leads to weaker links between semantics and phonology in the bilingual linguistic system, which in turn makes bilinguals slower to recall and produce words than monolinguals. In a similar vein, the Activation Threshold Hypothesis claims that the more a language is used, the lower the activation threshold, i.e., the easier and/or faster it is to comprehend/produce that language. In contrast, languages with a higher activation threshold are subject to attrition or language loss. In these studies, the language of most exposure and use is the language in which bilinguals are able to react or respond faster and produce more infrequent lexical items.

From these definitions, it becomes readily apparent that language dominance is not a simple unidimensional construct. Instead, language dominance can be thought of as an umbrella term for possible and/or observed asymmetries that are a result of the many linguistic, social and cognitive factors and realities of bilingualism. That being said, there is a general consensus in the field that language dominance is something that is relative, continuous and fluid (Gertken et al. 2014). Language dominance is relative in that we compare dominance in one language ( $L_A$ ) with dominance in the other

(L<sub>B</sub>). It is continuous, as opposed to categorical, in that bilinguals are not simply dominant in L<sub>A</sub> but more or less dominant in L<sub>A</sub> with respect to the other language and with respect to other L<sub>A</sub> dominant bilinguals. It is also a fluid construct in that it can change over time as the language use and linguistic abilities of a bilingual change.

Given the multifaceted nature of language dominance, researchers have employed a wide variety of methods to operationalize and measure the observed asymmetries that it encompasses. Table 1 illustrates 19 different methods that have been previously employed as a measurement of language dominance.<sup>2</sup>

**Table 1.** Language Dominance Assessments <sup>1</sup>.

Method (Study)	Objective vs. Subjective	Continuous vs. Categorical	Children vs. Adults
Self-ratings (e.g., Antoniou et al. 2011)	Subjective	Categorical	Adults
Bilingual Language Profile (e.g., Amengual and Chamorro 2015)	Subjective	Both	Adults
Bilingual Dominance Scale (e.g., Dunn and Tree 2009)	Subjective	Both	Adults
Boston Naming Task (e.g., Gollan et al. 2012)	Objective	Categorical	Adults
Multilingual Naming Task (e.g., Sheng et al. 2014)	Objective	Categorical	Adults
Lexical richness (e.g., Treffers-Daller 2011)	Objective	Categorical	Adults
Sentence Repetition Task (e.g., Flege et al. 2002)	Objective	Categorical	Adults
Category Generation Task (e.g., Bahrick et al. 1994)	Objective	Categorical	Adults
Semantic knowledge test (e.g., Bedore et al. 2012)	Objective	Categorical	Children
Patterns of language use and exposure (e.g., Argyri and Sorace 2007)	Subjective	Categorical	Both
Perceived accent (e.g., Olson 2016)	Objective	Categorical	Adults
Mean length of utterance (MLU) (e.g., Yip and Matthews 2006)	Objective	Categorical	Children
Language preference (e.g., Saunders 1988)	Subjective	Categorical	Children
Direction of language mixing (e.g., Lanza 2004)	Objective	Categorical	Children
Recognition and recall of words under noise (e.g., Golato 2002)	Objective	Categorical	Adults
Spoken Proficiency (Oral Proficiency Interview, OPI) (e.g., Gollan et al. 2012)	Objective	Categorical	Adults
Word types (e.g., Deuchar and Muntz 2003)	Objective	Categorical	Children
Morphosyntactic knowledge test (e.g., Bedore et al. 2012)	Objective	Categorical	Children
Child HALA test (e.g., Dubiel and Guilfoyle 2017)	Objective	Categorical	Children

<sup>1</sup> The identification of a given assessment as objective/subjective, categorical/continuous and used for adults/children is listed in accordance with the specific study mentioned.

Table 1 reveals several differences and commonalities between the different measurements. First, a combination of both subjective (i.e., participant providing information about self) and objective (i.e., someone/something else providing information about participants) measures have been used. For example, self-ratings, one of the most commonly used assessments, is subjective as it asks bilinguals to rate their linguistic abilities (typically presented in terms of reading, writing, speaking and listening skills) of both of their languages. A frequent criticism of self-ratings is that bilinguals may be biased for or against one of their languages. A particular group of bilinguals might be inclined to under

<sup>2</sup> We acknowledge that there are more assessments that have been used to measure language dominance than are listed in this table.

or overvalue (i.e., rate lower or higher) one of their languages or a specific facet of their languages based upon their language or educational background (see Section 7 for more discussion of this issue). Other common subjective measures of language dominance take the form of sociolinguistic questionnaires such as the Bilingual Language Profile (BLP) and the Bilingual Dominance Scale (BDS) or questionnaires about language use and exposure. These questionnaires often rely on the bilinguals' memory of their perceived language use and history.

Objective measures of language dominance can take several forms as well from tasks in which bilinguals are asked to name words as fast as they can (e.g., Boston Naming Task, Multilingual Naming Task, Category Generation Task, Child HALA test) to tasks that directly measure either written or spoken language proficiency (e.g., Semantic/morphosyntax knowledge test, Oral proficiency interview). These objective measures are conducted in both languages and then the language in which a bilingual performs best (i.e., scores higher) is considered his/her dominant language.

Second, even though, as [Gertken et al. \(2014\)](#) and [Grosjean \(1998, 2001\)](#) argue that language dominance is continuous in nature, we see that it has primarily been treated as a categorical variable. That is, researchers use the results of a language dominance assessment to classify and divide bilinguals into different language dominance groups ( $L_A$  dominant,  $L_B$  dominant, balanced) instead of placing the bilinguals on a single scale of more or less dominant than each other. Further, it is interesting to note that the two assessments in Table 1 that view language dominance as a continuous variable, The BLP and BDS, can be used to treat language dominance as a categorical variable as well. Both are questionnaires that provide a score on a given scale (see Section 4.2 for more details). Researchers can consider language dominance on a continuum by using the scores and/or they can use the scores to group the bilinguals into different dominance groups based upon the cutoff points provided. For example, studies such as [Amengual and Chamorro \(2015\)](#) used the scores provided by the BLP to classify participants into different dominance groups, while studies such as [Amengual \(2016\)](#) used the scores provided by the BLP to treat language dominance both as a categorical and a continuous variable in two separate analyses.

Third, we see that different measures are used to test language dominance in children versus adults. Assessments such as the Mean Length of Utterance (MLU) and the Child HALA are designed specifically for use with children and are not used to measure language dominance in adults, while assessments such as self-ratings (e.g., [Antoniou et al. 2011](#); [Tsui et al. 2019](#)) and perceived accent are used primarily with adults and not with children. Other assessments, such as patterns of language use and exposure (e.g., [Argyri and Sorace 2007](#)) have been used with both children and adults, the key difference being that bilingual adults report their own patterns of language use and exposure while the parents of bilingual children provide that information instead of the children themselves. Additionally, we note that some assessments (e.g., Multilingual Naming Task, MINT) were developed specifically with bilinguals in mind whereas other assessments (e.g., Boston Naming Task, BNT) were developed for monolinguals and were then later adapted for use with bilinguals.

Given the multitude of assessments used and the differences between them, an important question for researchers in bilingualism is whether or not these different measures capture the construct of language dominance in the same way and on a larger scale, whether they capture language dominance at all. That is, if we want to design a research study that looks at the effects of language dominance on a given linguistic phenomenon or even if we just want to control for language dominance as an intervening variable, does it matter which method we employ to measure language dominance? Would we get different results if we used a subjective versus objective measure? Or if we treated language dominance as a categorical versus continuous variable? And if so, how do we know which assessment we should ultimately use?

In order to begin to answer these questions, we should examine (a) whether these different assessments classify an individual the same way with respect to dominance and (b) whether using a different assessment leads to a different interpretation of the results of bilingual behavior.

### 3. Comparing Different Language Dominance Assessments

In this section, we discuss three studies, [Bedore et al. \(2012\)](#), [Gollan et al. \(2012\)](#) and [Sheng et al. \(2014\)](#) that have each analyzed a subset of language dominance assessments with respect to different groups of bilinguals.<sup>3</sup> [Bedore et al. \(2012\)](#) examine the utility of different assessments on bilingual children. [Sheng et al. \(2014\)](#) look at child and young adult bilinguals, whereas [Gollan et al. \(2012\)](#) study young adult and aging adult bilinguals.

[Bedore et al. \(2012\)](#) examined language dominance in Spanish/English bilingual children. In their study, language dominance was operationalized in two ways: (1) as relative proficiency (found by comparing a proficiency score in  $L_A$  by that of  $L_B$ ) and (2) the language of most exposure (determined by parental report). Children completed the BESOS (Bilingual English and Spanish Oral Language Screening, [Peña et al. 2010](#)) in both English and Spanish, which included semantics and morphosyntax subtests. As a part of the BESOS, interviews were completed with the parents to determine the language use of the children (combination of input and output). The scores achieved by the children on the semantics and morphosyntax tests were used to classify language dominance in that a child was considered dominant in  $L_A$  if s/he scored higher on the semantics/morphosyntax test in  $L_A$  when compared to  $L_B$ . The parental interviews were also used to calculate language dominance by averaging out the reported input and output of each language; A child was considered dominant in  $L_A$  if s/he had a higher percentage of language use in  $L_A$  when compared to  $L_B$ .

The results of their analysis determined that in general, different ways of operationalizing language dominance can result in different classifications of bilingual children. Specifically, 17.1% of the children were classified differently based on input and output. Further, 51.2% of the children were classified differently based on the results of the morphosyntax versus semantics tests. As [Bedore et al. \(2012\)](#) explain “if tested at 60% usage of English, children would appear to be English dominant if given a semantics test but Spanish dominant if given a morphosyntax test” (p. 622). In other words, using the metrics of language use and performance on a semantics test, children would be classified into an English-dominant group but using the metric of a morphosyntax test they would be classified into a Spanish-dominant group. These findings suggest that it does make a difference which assessment is used to determine and classify the language dominance of participants and that this difference could lead to different interpretations of the results.

[Gollan et al. \(2012\)](#) looked at language dominance in young adult and aging Spanish/English bilinguals, while [Sheng et al. \(2014\)](#) examined Mandarin/English child and adult bilinguals. As [Sheng et al. \(2014\)](#) was a replication of [Gollan et al. \(2012\)](#), both studies operationalized language dominance as relative proficiency calculated from the results of four assessments: self-ratings, scores on oral proficiency interviews (OPI), the Boston Naming Task (BNT) and the Multilingual Naming Task (MINT). Both [Gollan et al. \(2012\)](#) and [Sheng et al. \(2014\)](#) found mismatches in the dominance classification of participants across the four measures. For instance, participants tended to be classified as English dominant more via the BNT or MINT than the self-ratings and OPI. The researchers concluded that in general, the different assessments are comparable in terms of classifying bilinguals into different groups but that they are unable to capture the degree of dominance.<sup>4</sup> From this, we can infer that the bilinguals who tend to fall more towards the middle of the scale (i.e., they are more balanced) are more difficult to accurately classify across different language measures.

The results of [Bedore et al. \(2012\)](#), [Gollan et al. \(2012\)](#) and [Sheng et al. \(2014\)](#) indicate that in research studies on bilingualism and language dominance it is important which metric is used to operationalize and measure language dominance, as different measures seem to classify bilinguals

<sup>3</sup> We refer the reader to [Treffers-Daller \(2019\)](#) for a discussion of three other studies ([Jia et al. 2002](#); [Lim et al. 2008](#); [Unsworth 2016](#)) that also examined whether different measures of language dominance correlate with one another.

<sup>4</sup> Note that it is most likely that the inability to capture degree of dominance is what leads to some of the participants being classified differently across the different tests.

differently, which in turn can lead to variable interpretation of the results. Together, these studies examined only six of the many different types of assessments previously used in the literature (see Table 1). As such, their findings highlight the need for more research comparing the differences between language dominance assessments and the reasons behind why different assessments that are all supposedly measuring the same thing, language dominance, classify bilinguals differently with respect to their dominant language. This discussion brings us to the current study which aims to analyze the comparability of a different set of commonly used language dominance assessments. To that effect, we ask the following research question:

RQ: Do different language dominance measures predict the results of one another?

In order to answer this question, we examined four language dominance assessments: The Bilingual Language Profile (Gertken et al. 2014), the Bilingual Dominance Scale (Dunn and Tree 2009), self-ratings of verbal and of written ability and a repetition task (Flege et al. 2002). Following the results of Bedore et al. (2012), Gollan et al. (2012) and Sheng et al. (2014) as well as Birdsong (2014)'s definition of language dominance, we hypothesized that dominance measures that do not test the same domain/dimension would not predict the results of one another with respect to dominance classification.

## 4. Materials and Methods

### 4.1. Participants

Twenty-nine Spanish/English bilinguals participated in this study. At the time of the study all participants were undergraduate students residing in the Chicagoland area. The age of the participants ranged between 18–24 ( $M = 20.89$ ,  $SD = 2.00$ ). All participants reported learning Spanish before the age of 4 ( $M = 0.24$ ,  $SD = 0.91$ ) and English before the age of 10 ( $M = 3.58$ ,  $SD = 2.88$ ). The participants grew up in Mexican ( $n = 28$ ) and Salvadoran ( $n = 1$ ) households where both Spanish and English were spoken. On any given week, participants reported speaking English with their friends an average of 71% ( $SD = 18.19$ ) of the time and Spanish 29% ( $SD = 20.51$ ) of the time. With family members or in a school environment, participants reported speaking English 41% ( $SD = 23.33$ ) of the time and Spanish 59% ( $SD = 24.05$ ) of the time. Finally, in a school or work environment, participants reported speaking English 72% ( $SD = 17.81$ ) of the time and Spanish 28% ( $SD = 20.36$ ) of the time on average per week. We further note that all participants are self-reported codeswitchers and that they reported at least one person with whom they regularly use both Spanish and English in the same conversation.

The participants were asked to self-rate their Spanish and English proficiency in speaking, understanding, reading and writing using a scale from 1 (not well) to 6 (very well). The participants' average self-rated proficiency is reported in Table 2. Note that the self-ratings were also used as one of the language dominance assessments (see Section 4.2); here the raw values are presented.

**Table 2.** Spanish and English self-rated proficiency.

	Spanish		English	
	M	SD	M	SD
Speaking	4.89	0.85	5.58	0.56
Understanding	5.44	0.63	5.86	0.35
Reading	4.89	1.11	5.65	0.61
Writing	4.44	1.15	5.58	0.62
Overall	4.92	0.84	5.67	0.44

### 4.2. Language Dominance Assessments

In our study, the four language dominance assessments administered were the Bilingual Language Profile (BLP; Gertken et al. 2014), the Bilingual Dominance Scale (BDS; Dunn and Tree 2009), self-ratings

of verbal/written ability and a repetition task (Flege et al. 2002). These four dominance assessments were chosen due to their prevalence in the language dominance literature.

The BLP is a computer-based questionnaire that calculates dominance on the basis of 19 multiple-choice and fill-in-the-blank items related to language history, language use, language proficiency, and language attitudes. To administer this assessment, participants complete the online questionnaire and a dominance score is automatically calculated on a scale from  $-218$  to  $218$  (see Gertken et al. 2014 for details on how the score is calculated on the basis of participant responses). Lower (negative) scores indicate greater Spanish dominance, higher (positive) scores indicate English dominance and a score of '0' indicates a perfect balance between the two.

The BDS is also a computer-based questionnaire that calculates dominance based on language history, language use, language proficiency, and language attitudes. This questionnaire, however, consists of 12 fill-in-the-blank items. The dominance score is calculated on a scale of  $-32$  to  $32$ . Like the BLP, lower (negative) scores indicate greater Spanish dominance, higher (positive) scores indicate English dominance and '0' indicates a perfect balance between the two. One of the most significant differences between the BLP and the BDS is that the BDS weighs the questions thus assigning more value to some domains. For example, questions related to language attitudes and language use at school are assigned less value than questions related to language history and language use at home (Dunn and Tree 2009).

Self-ratings of proficiency are one of the most commonly used language dominance assessments in the field (Flege et al. 2002; Gertken et al. 2014). To assess language dominance using self-ratings, participants rate their own ability to speak, understand, read, and write in both languages, using a scale from 1 (not well) to 7 (very well).<sup>5</sup> Following Flege et al. (2002), for this dominance assessment, two scores are calculated: one for verbal ability (i.e., speaking and understanding) and one for written ability (i.e., reading and writing). To calculate a dominance score, the ratio of Spanish and English self-ratings is calculated. A score lower than one indicates English dominance, a score higher than one indicates Spanish dominance, and a score of '1' indicates a perfect balance between the two.

The fourth language dominance assessment is a repetition task. This task was used by Flege et al. (2002) to measure language dominance in early Italian-English bilinguals. In this task, participants hear 12 sentences, once in each language, and after the participant hears each of the sentences, one at a time, s/he repeats the sentence. The underlying assumption is that participants are more likely to repeat sentences faster in their dominant language. The productions are recorded, and each sentence is measured to the nearest millisecond from onset to offset. A dominance score is generated based on the ratio of the two languages. A score lower than one indicates English dominance, a score higher than one indicates Spanish dominance, and a score of '1' indicates a perfect balance between the two.

To ensure comparability with the original study, the design and items of the repetition task were taken from Flege et al. (2002) and adapted for the language pair of this study. The 12 Italian sentences of the original task were translated into Spanish and the 12 English sentences were kept in the original form. As in the original task, all 12 pairs of sentences were matched for the number of words and syllables and had identical or nearly identical meanings. The 12 English items and 12 Spanish items were recorded by a Spanish/English bilingual (the first author) in separate sessions, to avoid code-switching effects, using a Shure SM81 microphone. The production was recorded digitally to disk using MOTU Ultralite external interface. The productions were digitized and normalized for peak intensity. Using Praat software, each sentence was measured to the nearest millisecond from onset to offset of acoustic energy associated with the phonetic segment. When recorded, the Spanish sentences ( $M = 2590.625$  ms) were slightly shorter than the English sentences ( $M = 2651.25$  ms), possibly

---

<sup>5</sup> The width of this scale varies by study. In the current study we employed a scale of 1–6 as the self-ratings were extracted from the BLP, which uses a 1–6 scale.

attributed to the first author's speech rate, but a similar effect occurred in [Flege et al. \(2002\)](#) original task (Italian  $M = 2548$ , English  $M = 2796$  ms).

#### 4.3. Experimental Design and Procedure

The data for this study were collected on two separate days as part of a larger research project (see [Stefanich 2019](#)). On Day 1, participants first provided their informed consent to participate in the study (IRB protocol approval #2015-0040). In each session, a Spanish/English bilingual with the same background as the participants conducted a 10-min interview in Spanish/English codeswitching (CS) in order to activate both of the participant's languages (i.e., establish a bilingual mode). Importantly, the interviewer did not request or ask the participants to codeswitch but rather she engaged in codeswitching in her role as the interviewer and let the participants respond and participate however they felt most comfortable. All instructions for the tasks were given in English/Spanish CS. While establishing a bilingual mode via the CS interview and the code-switched instructions was a required methodological consideration for the larger study, it also served as a control to help prevent participants' biases towards a particular language during completion of the language dominance assessments.

During each session the participants first completed the Spanish/English CS interview, then an aural acceptability judgment task (AJT) that was related to the larger study (see [Stefanich 2019](#) for details). After the AJT, in the first session, participants completed the repetition task and then filled out the BLP. In the second session, after the AJT, the participants completed the BDS. The repetition task was administered in a sound-attuned booth using the stimuli presentation software, E-prime 2.0 (Psychology Software Tools, Inc., Pennsylvania, USA). In order to replicate [Flege et al. \(2002\)](#), participants were not given any instruction regarding their rate of speech. Participants were informed that they would be aurally presented with each item twice and that their task was to repeat the sentence following a tone. The production was recorded directly to disk using MOTU Ultralite external interface using a Shure SM81 microphone. Participants completed the Spanish/English BLP and the BDS on a computer and were instructed to respond to the best of their knowledge.

Note that self-ratings of verbal and written ability are a part of the BLP and therefore we didn't ask the participants to do the same thing twice. Instead, we administered the BLP and later extracted the self-rating scores from the BLP to use as an individual assessment.

#### 4.4. Analysis

In total five<sup>6</sup> language dominance scores were calculated for each participant, following the guidelines of the original assessments ([Gertken et al. 2014](#); [Dunn and Tree 2009](#); [Flege et al. 2002](#)). The BLP dominance score was automatically generated and the BDS score was manually calculated following [Dunn and Tree \(2009\)](#). To calculate the score for self-ratings of verbal and of written ability, first the self-ratings were extracted from the BLP. To calculate the verbal ability ratings for Spanish and English, the self-ratings of speaking and understanding were averaged together. To calculate the writing ability ratings, the self-ratings of reading and writing were averaged together. Finally, to calculate the dominance scores of verbal and written ability, the mean Spanish self-rating was divided by the mean English self-rating.

To calculate the repetition task dominance score, each sentence was measured to the nearest millisecond from onset to offset. The average duration of the 12 Spanish sentences and the 12 English sentences was calculated. To generate the dominance score, the mean duration of the Spanish sentences was divided by the mean duration of the English sentences. For the second step of the analysis, a series of correlations were run using SPSS, comparing each of the five assessments with each other.

---

<sup>6</sup> Recall, self-ratings were divided into two types, verbal and written ability, thus when combined with the BLP, the BDS, and Repetition task we have five assessments.

### 5. Results

Our research question asked whether different language dominance assessments predict the results of one another with respect to the classification of a bilingual’s dominant language. In this section, we first present the dominance score and classification distribution for each assessment. Then, we turn to an analysis of the individual participants to determine how they were classified across the different measures based upon their individual dominance scores.

The Bilingual Language Profile calculates a language dominance score on a scale from –218 (Spanish dominance) to 218 (English dominance). In this experiment all participants scored between –46 and 102 ( $M = 15.077, SD = 45.073$ ). Based on their score, 19 participants were categorized as English dominant, nine were classified as Spanish dominant, and one participant was classified as Balanced. Individual scores are provided in Appendix A.

The Bilingual Dominance Scale calculates a dominance score on a scale from –32 (Spanish dominance) to 32 (English dominance). All of the participants in this experiment scored between –5 and 18 ( $M = -4.727, SD = 9.424$ ). According to their dominance score, 20 participants were classified as English dominant, eight were classified as Spanish dominant, and one participant was classified as Balanced. Individual scores are provided in Appendix B.

The self-ratings of proficiency provided two separate scores, one of verbal ability and one of written ability, which will be presented individually. In the self-ratings of verbal ability, all participants scored between 0.64 and 1.09 ( $M = 0.90, SD = 0.10$ ). According to their scores, 17 participants were classified as English dominant, one was classified as Spanish dominant and 11 participants were classified as Balanced. Individual scores are provided in Appendix C.

In the self-ratings of written ability, participants scored between 0.5 and 1.2 ( $M = 0.80, SD = 0.22$ ). According to the self-ratings of verbal ability, 21 participants were classified as English dominant, two were classified as Spanish dominant and six participants were classified as Balanced. Individual scores are provided in Appendix D.

In the repetition task, participants scored between 0.92 and 1.19 ( $M = 1.00, SD = 0.074$ ). Based on this assessment, 19 participants were classified as English dominant, nine participants were classified as Spanish dominant, and one was classified as Balanced. Individual scores are provided in Appendix E.

A visual summary of the dominance classification distribution between the five assessments is provided in Figure 1.

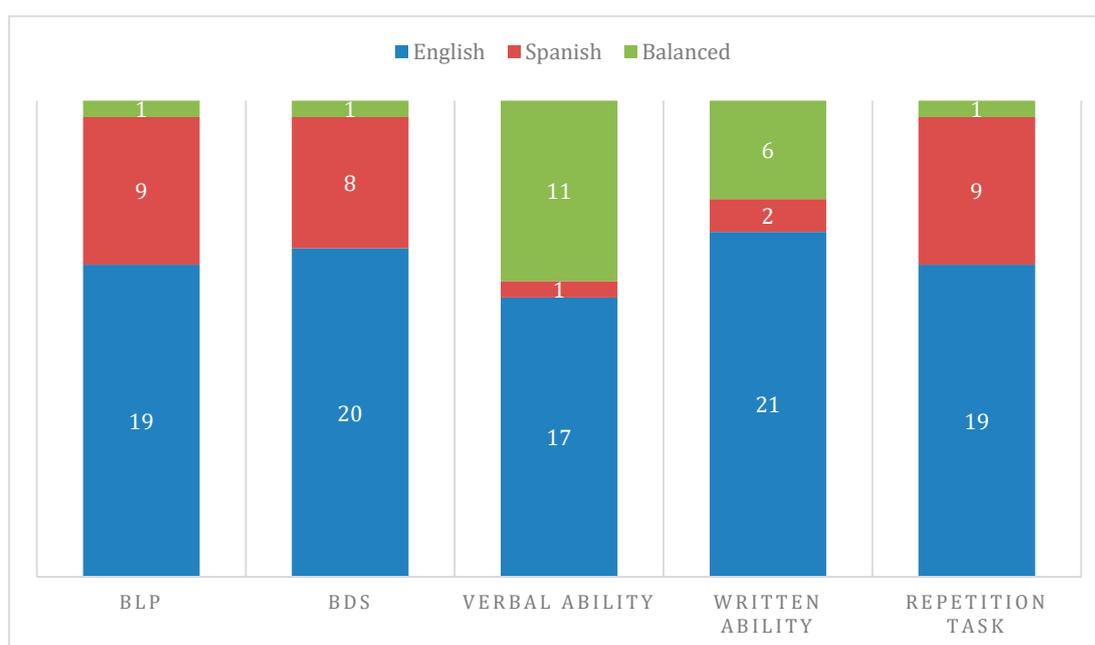


Figure 1. Language dominance classifications.

To corroborate whether or not participants were consistently classified into the same dominance group across the five assessments, the individual dominance classifications of each participant were examined (Table 3). The results show that 20 out of the 29 participants were classified differently in one or more of the assessments. 13 out of the 29 participants were categorized into two different dominance groups and seven out of the 29 participants were categorized into three different dominance groups.

**Table 3.** Individual dominance classifications.

Participant	BLP	BDS	Verbal Ratings	Written Ratings	Repetition Task
2001	English	English	English	English	English
2002	Balanced	English	Balanced	English	Spanish
2003	Spanish	Balanced	Balanced	English	English
2004	English	English	English	English	English
2005	Spanish	Spanish	Balanced	Balanced	English
2006	English	Spanish	English	English	English
2007	Spanish	Spanish	Spanish	Spanish	Spanish
2008	English	English	English	English	English
2009	Spanish	Spanish	Balanced	English	English
2010	Spanish	English	English	English	Spanish
2011	English	English	English	English	English
2012	English	English	English	English	English
2013	English	English	Balanced	Balanced	Balanced
2014	English	English	English	English	English
2015	English	English	Balanced	English	English
2016	Spanish	Spanish	English	English	English
2017	English	English	English	English	Spanish
2018	Spanish	Spanish	Balanced	Spanish	English
2019	English	English	English	English	Spanish
2020	English	English	Balanced	Balanced	English
2022	English	Spanish	English	English	Spanish
2023	English	English	English	English	Spanish
2024	Spanish	English	Balanced	Balanced	Spanish
2025	Spanish	Spanish	Balanced	English	English
2026	English	English	English	Balanced	English
2027	English	English	English	English	Spanish
2028	English	English	Balanced	Balanced	English
2029	English	English	English	English	English
2030	English	English	English	English	English

After establishing the language dominance scores for all assessments, a series of correlations ( $n = 10$ ) were run between the BLP scores, the BDS scores, the self-ratings of verbal ability scores, the self-ratings of written ability scores and the repetition task scores. To correct for multiple comparisons, the Holm-Bonferroni correction was used. Table 4 shows the pairwise comparisons with the adjusted alpha levels.

The results indicated a moderate, positive association between the BLP and the BDS ( $r(29) = 0.692$ ,  $p < 0.001$ ). Further, the BLP had a strong, negative association ( $r(29) = -0.761$ ,  $p < 0.001$ ) with the self-rating of verbal ability. The BDS and the self-ratings of verbal ability had a moderate, negative association ( $r(29) = -0.527$ ,  $p = 0.003$ ), whereas the self-ratings of verbal and written ability had a moderate, positive association ( $r(29) = 0.534$ ,  $p = 0.003$ ). No significant correlations were found between the self-ratings of written ability and the BDS ( $p = 0.210$ ), or the BLP ( $p = 0.015$ ). Further, no significant correlations were found between the repetition task and any of the other assessments.

**Table 4.** Pairwise comparisons of Language Dominance Assessments with Holm-Bonferroni correction.

Comparison	Holm-Bonferroni Adjusted Alpha Value ( $\alpha$ )	<i>p</i> -Value
BLP $\times$ BDS	0.005	$p < 0.001$
BLP $\times$ self-ratings of verbal ability	0.0056	$p < 0.001$
BDS $\times$ self-ratings of verbal ability	0.00625	$p = 0.003$
Self-ratings of verbal ability $\times$ self-ratings of written ability	0.007	$p = 0.003$
BLP $\times$ self-ratings of written ability	0.008	n.s $p = 0.015$
BLP $\times$ repetition task	0.01	n.s $p = 0.126$
BDS $\times$ self-ratings of written ability	0.125	n.s $p = 0.210$
BDS $\times$ repetition task	0.0167	n.s $p = 0.311$
Self-ratings of written ability $\times$ repetition task	0.025	n.s $p = 0.507$
Self-ratings of verbal ability $\times$ repetition task	0.05	n.s $p = 0.540$

## 6. Discussion

The goal of this study was to investigate whether different language dominance assessments produce comparable results. To answer the research question, 29 early Spanish/English bilinguals completed four dominance assessments: BLP (Gertken et al. 2014), BDS (Dunn and Tree 2009), self-ratings of proficiency and a repetition task (Flege et al. 2002) and the language dominance scores from each assessment were compared to one another. The language dominance measures used in this study varied in regard to the dimensions and domains of dominance they examined and as was predicted, not all of the language dominance measures provided comparable results. This finding supports the conclusions drawn by previous studies (Bedore et al. 2012; Gollan et al. 2012; Sheng et al. 2014), who found a similar result in their analyses of different sets of language dominance assessments.

In our study, the BLP did significantly predict the dominance scores provided by the BDS and the self-ratings of verbal proficiency. A relationship between these three measures is to be expected because the BLP and BDS are both language questionnaires and they determine dominance based upon the same factors: language history, language use, language proficiency, and language attitudes. Further, self-ratings of verbal ability are a part of the BLP and the BDS, so it is to be expected that when self-ratings of verbal ability are considered their own assessment, they relate to the BLP and BDS.

However, unlike the self-ratings of verbal ability, the self-ratings of written ability did not correlate with either the BLP or the BDS. With respect to the BLP, this outcome is unexpected considering self-ratings of written ability are included within the BLP questionnaire and because the self-ratings of written ability correlated with those of verbal ability. In contrast, the BDS only accounts for verbal language proficiency and not written language proficiency like the BLP (Dunn and Tree 2009). The discrepancy between the dominance scores of the verbal and written self-ratings suggests that self-ratings of verbal ability capture a different facet of a bilingual's competence than self-ratings of written ability. This might be especially relevant for heritage speakers who are well known for being more proficient in terms of speaking/listening than reading/writing the heritage language.

In this study, the dominance scores from the repetition task did not predict the scores of any other assessment. There are several notable differences between the repetition task and the other assessments that could explain this lack of association. First, the repetition task assessed dominance via a processing/production task while the other tasks assessed language dominance based on their language background/history. Second, given current research on the cost effects of codeswitching it could be the case that the design of the repetition task itself is not ideal to capture the linguistic processing dimension of dominance. In this task, participants heard and repeated sentences in both Spanish and English and the order of the languages was counterbalanced within the repetition task. That is, participants repeated two sentences in English, then two sentences in Spanish, etc. In

other words, participants were, in effect, switching between their two languages throughout the task. This was done in order to replicate the repetition task as presented in the original study (Flege et al. 2002). More recent work on code-switching has found evidence that switching from one language to another leads to processing burdens (see Van Hell et al. 2015 for review). Bilinguals are thought to be slower to switch from their non-dominant language into their dominant language given that they have to work harder to suppress their dominant language when speaking the non-dominant language (e.g., Meuter and Allport 1999).

Given that the underlying assumption behind the repetition task in Flege et al. (2002) was that bilinguals would be more likely to repeat sentences faster in their dominant language, we see a clear conflict with the use of the switching between the two languages in the task. If the participants demonstrate the cost effects of language switching, then they would actually be slower to repeat half of the sentences in their dominant language (i.e., the ones directly following a sentence in the opposite language), which would then skew the measurements used to calculate the dominance score and therefore the classification. For the purposes of our study, we aimed to replicate the use of each assessment as close as possible to the way they have been used previously. It remains to be seen whether the output of a repetition task that did not include the convoluting factor of language switching would correlate with that of other language dominance measures. Future work on language dominance could employ a modified version of Flege’s repetition task in which participants first repeat sentences in  $L_A$  and then in  $L_B$  (and vice versa) or in which the sentences are divided into two sessions so that the participants are not forced to switch between their languages. Third, some important information needed to accurately replicate the analysis of the repetition task was not provided in Flege et al. (2002). For instance, Flege et al. (2002) mentioned that “non-fluently produced sentences were excluded from analysis” (p. 580) but no criteria about what was considered non-fluently produced sentences were given.

Even though the results of the correlation analysis indicated that the dominance scores provided by the BLP, BDS, and self-ratings of verbal ability are predictive of one another, 20 of the 29 participants were classified into different groups based upon those scores. Even if we remove the repetition task from the analysis (as it did not correlate with any other assessment), 16 of the participants were still classified differently, even across the BLP, BDS, and self-ratings of verbal ability (see Table 3). The average dominance scores for the 16 participants whose dominance classification changed across the four measures and for the 13 participants whose classification did not change are provided in Table 5.

**Table 5.** Average dominance scores by assessment.

	BLP −218 (Spanish) to 218 (English)		BDS −32 (Spanish) to 32 (English)		Self-Ratings Verbal Ability <1 (English) to >1 (Spanish)		Self-Ratings Written Ability <1 (English) to >1 (Spanish)	
	M	SD	M	SD	M	SD	M	SD
Participants with consistent dominance classification	42.86	34.26	9	7.19	0.83	0.11	0.72	0.18
Participants with inconsistent dominance classification	1.33	23.52	2.37	6.54	0.97	0.05	0.92	0.12

We note, descriptively speaking, that the bilinguals who were more inconsistently classified are the bilinguals whose dominance scores tended to fall towards the middle of the scale (i.e., were more balanced). Recall that a perfectly balanced bilingual would score 0 on the BLP and BDS and 1 on the self-ratings. Bilinguals whose scores were at more extreme ends of the scale (i.e., more Spanish dominant or more English dominant) tended to maintain their dominance classification across the five measures. For example, Participant 2001 was classified as English dominant across the board with scores of 102, 18, 0.75 and 0.58 on the BLP, BDS, self-ratings of verbal ability and self-ratings of

written ability respectively. In contrast, Participant 2003 was classified as Spanish dominant, English dominant and balanced with scores of  $-11$ ,  $0.916$ ,  $0$ , and  $1$  on the BLP, self-ratings of written ability, BDS, and self-ratings of verbal ability respectively. These results suggest that more balanced bilinguals are more difficult to consistently classify across different language dominance assessments.

This discrepancy in the individual dominance classification of certain bilinguals echoes the results of Gollan et al. (2012) and Sheng et al. (2014), who found similar mismatches across four different dominance measures: The Boston Naming Task, The Multilingual Naming Task, self-ratings, and oral proficiency interviews. They concluded that the dominance measures in their studies were unable to capture “degree of dominance”, thus begging the question as to whether the dominance assessments used in our current study were also unable to capture the degree of dominance of a given bilingual (i.e., the relationship of how dominant a bilingual is in a given language in comparison to other bilinguals dominant in that same language).

This is a significant finding as it suggests a difference in treating language dominance as a categorical vs. continuous variable. If you recall, the majority of assessments provided in Table 1 treat language dominance as a categorical variable. The results of the current study show that treating language dominance as a categorical variable is problematic given that an individual bilingual will not be consistently placed into the same dominance group depending on which assessment is given. When we think about this abstractly, we can see that it makes sense. Consider the hypothetical case of three different Spanish/English bilinguals who were all given the BLP. Bilingual 1 received a score of  $-5$  and was placed into the Spanish dominant group. Bilingual 2 received a score of  $5$  and was placed into the English dominant group and Bilingual 3 received a score of  $67$  and was also placed into the English dominant group. In this scenario, there is a good chance that based on their scores Bilinguals 1 and 2 are more similar to each other than Bilinguals 2 and 3. Yet, Bilinguals 2 and 3 will be placed into the same group whereas Bilingual 1 is in a different group. In our study, this hypothetical scenario did, in fact, take place. Following the BLP, Participant 2026, received a score of  $8.72$ , and was classified as English dominant. Participant 2001 received a score of  $102$  and was also classified as English dominant. In contrast, Participant 2010 was classified as Spanish dominant with a score of  $-4$ . When we view these scores on a continuum, Participant 2026’s score is closer to that of Participant 2010 than that of Participant 2001.

Now let’s consider what would happen if these three bilinguals were a part of a research study that was examining the effect of language dominance on a given linguistic phenomenon. Using the dominance scores on the basis of a continuum, we would logically predict that Participant 2026 will behave more similarly to Participant 2010 on a given linguistic task than Participant 2001. Or we might predict that Participant 2026’s behavior should fall somewhere in between that of Participant 2010 and 2001. However, if we use these scores to classify the participants into different dominance groups, Participant 2026 will end up in the same group as Participant 2001. Following this dominance classification, we would predict that Participant 2026 and Participant 2001 should pattern more similarly on a given task than either of them should with Participant 2010 (given that they belong in two different groups). Based on this simplified example, we can see that a given bilingual’s linguistic behavior is predicted to be different based upon whether language dominance is treated as a categorical versus a continuous variable.

Further, a natural consequence of treating language dominance as categorical is the potential heterogeneity of each group. In our example, Participant 2026 and Participant 2001 are both classified as English dominant even though their individual scores fall quite far from each other. Collapsing participants across the group level (i.e., English dominant versus Spanish dominant) could potentially obscure patterns that would be otherwise more visible if the participant’s linguistic behavior was examined based upon a continuum of their dominance scores. This would allow us a more fine-grained analysis of the construct of language dominance in which we could answer questions as to whether the more dominant a bilingual is in  $L_A$  the more likely they are to behave a certain way with respect to their production, perception or processing of a given linguistic phenomenon.

## 7. Conclusions

This study provides evidence that not all language dominance measures provide comparable results. Out of the five assessments tested, only the Bilingual Language Profile (BLP), Bilingual Dominance Scale (BDS) and self-ratings of verbal ability were found to predict the scores of one another, and the self-ratings of verbal ability predicted the scores of the self-ratings of written ability. The self-ratings of written ability did not predict the scores of the BLP nor the BDS and the repetition task did not predict the scores of any other assessment. Further, across the five language dominance measures 20 out of the 29 participants in this study were classified differently in one or more measure. Given the fact that language dominance is a construct that represents asymmetries in language ability, processing, and use, these results suggest that language dominance measures that tap into one specific domain/dimension of dominance do not necessarily predict the dominance of another domain/dimension. In other words, it could be the case that a bilingual's dominant language is not all encompassing but rather his/her dominant language varies as a result of the domain and/or dimension in question. The implication of this result is that when comparing across studies that examine language dominance, we cannot assume, for example, that all Spanish dominant bilinguals are alike. Rather, we must ask the follow-up question "Spanish dominant in what way?" (see (Grosjean 2016) for more discussion of the domains of language dominance). Thus, when it comes to operationalizing language dominance, we must consider what and how each individual measure tests dominance as it seems to be the case that different measures approach it differently. For instance, the BLP and the BDS are questionnaires and calculate dominance based on language history, language use, language proficiency, and language attitudes. The self-ratings treat proficiency and dominance as interchangeable constructs, and the repetition task measures dominance based on a processing task.

The difference in results between the correlation analysis of the dominance scores and the dominance classification based upon those scores, has implications for treating language dominance as a continuous versus a categorical variable. When language dominance is viewed as a continuous variable (here, the correlation analysis), a relationship was found between three of the assessments: the BLP, the BDS, and the self-ratings of verbal ability, and also between the self-ratings of verbal and written ability. However, when language dominance is viewed as a categorical variable (here, the dominance classifications), 69% of the participants were classified into more than one dominance group across the five different assessments, including the BLP, BDS, and self-ratings of verbal ability. In other words, treating language dominance as categorical is problematic because a given bilingual could be classified differently depending on what language dominance assessment is being used. Particularly, the more balanced a bilingual is the greater chance s/he has of being inconsistently classified. Different dominance classifications of a group of bilinguals leads to different predictions as to how said bilinguals will behave on a given linguistic task. The potential heterogeneity of each group based upon these inconsistent classifications feeds different and possibly contradictory results which in turn obstructs successful synthesis of studies that examine the effects of language dominance on linguistic phenomena. Given that these different assessments measure different aspects of language dominance and that these aspects result in different dominance profiles, we can ask several theoretical questions. First, are all these assessments actually measuring language dominance? Second, if they are, then what is language dominance? Can we simply consider a bilingual to be dominant in  $L_A$ ? Or must we specify that a bilingual is dominant in  $L_A$  with respect to a specific domain/dimension?

Based on the results of the current study it seems to be the case that some dominance measures are better equipped for assessing dominance in particular bilingual populations. When analyzing dominance in early and late bilinguals together, Flege et al. (2002) found that the repetition task and self-ratings provided consistent dominance scores. In the current study in which only early bilinguals were tested, the same result did not occur. There was no significant correlation between the dominance scores provided by the repetition task and the self-ratings. Although self-ratings of proficiency are a widely used language dominance measure (Gertken et al. 2014), this assessment has been found to be less valuable for some bilingual populations (see Tomoschuk et al. 2019 for discussion). Due to their

linguistic environment and the limited formal instruction that Spanish/English heritage bilinguals receive, it could be the case that heritage bilinguals underrate their verbal and writing proficiency in Spanish. As a result, their self-ratings might not accurately be representative of their proficiency which might make this assessment unreliable for determining dominance in early heritage bilinguals.

In sum, in this study, we have shed light on some methodological issues regarding language dominance. For the purpose of replicability and comparability, we suggest that it is important to determine what language dominance is (i.e., work towards a consensus on a unified definition) and to develop a unified methodology for assessing it. In order to do so, it is important to consider the bilingual population and the domain and dimension being tested. We suggest that treating language dominance as a continuous variable, rather than categorical, can help mitigate some of the possible effects of using a variety of assessments across the field.

Future work will test three other dominance measures in early Spanish/English bilinguals including the Multilingual Naming Test (e.g., Gollan et al. 2012), a category generation task (e.g., Bahrick et al. 1994) and measures of spoken proficiency in the form of oral interviews (e.g., OPI). By assessing more language dominance measurements, we can begin to get an idea of which types of assessments pattern together in order to delve deeper into the construct of language dominance. Additionally, future research will test different types of bilinguals (e.g., late bilinguals) to see if we find the same issues with classifying bilinguals into dominance groups.

**Author Contributions:** C.S.-B. performed the experiments, analyzed the data and wrote the paper. S.S. conceived and designed the experiment, supervised data collection and analysis, acquired funding and wrote the paper.

**Funding:** The data presented in this study was collected as part of a larger research project that was funded by the National Science Foundation Doctoral Dissertation Research Improvement Grant, BCS #1823909.

**Acknowledgments:** First and foremost, the authors would like to thank Luis López for his guidance and feedback on previous versions of the manuscript. The authors would also like to thank Sahian López for her assistance in data collection and Kara Morgan Short for her valuable feedback. Further thanks are extended to members of the Bilingualism Research Laboratory and the Multilingual Phonology Lab for their feedback on previous versions of the manuscript.

**Conflicts of Interest:** The authors declare no conflict of interest. The funding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

## Appendix A

**Table A1.** Bilingual Language Profile (BLP) individual dominance scores on a scale from −218 (Spanish) to 218 (English).

Participant	Dominance Score	Dominant Language
2001	102	English
2002	0	Balanced
2003	−11	Spanish
2004	52	English
2005	−42	Spanish
2006	34	English
2007	−46	Spanish
2008	28	English
2009	−11	Spanish
2010	−4	Spanish
2011	63	English
2012	71.2	English
2013	22	English
2014	20.25	English
2015	21.98	English
2016	−20.24	Spanish

Table A1. Cont.

Participant	Dominance Score	Dominant Language
2017	66.29	English
2018	-24.61	Spanish
2019	33.15	English
2020	15.62	English
2022	9.18	English
2023	39.83	English
2024	-33.77	Spanish
2025	-20.06	Spanish
2026	8.72	English
2027	45.77	English
2028	34.96	English
2029	40.41	English
2030	41.23	English

## Appendix B

**Table A2.** Bilingual Dominance Scale (BDS) individual dominance scores on a scale from -32 (Spanish) to 32 (English).

Participant	Dominance Score	Dominant Language
2001	18	English
2002	17	English
2003	0	Balanced
2004	16	English
2005	-2	Spanish
2006	-2	Spanish
2007	-8	Spanish
2008	7	English
2009	-6	Spanish
2010	5	English
2011	6	English
2012	15	English
2013	4	English
2014	9	English
2015	6	English
2016	-3	Spanish
2017	4	English
2018	-2	Spanish
2019	13	English
2020	7	English
2022	-3	Spanish
2023	15	English
2024	1	English
2025	-5	Spanish
2026	9	English
2027	4	English
2028	12	English
2029	14	English
2030	4	English

### Appendix C

**Table A3.** Self-ratings of Verbal Ability individual dominance scores on a ratio of <1 (English) to >1 (Spanish).

Participant	Dominance Score	Dominant Language
2001	0.75	English
2002	1	Balanced
2003	1	Balanced
2004	0.833	English
2005	1	Balanced
2006	0.916	English
2007	1.09	Spanish
2008	0.833	English
2009	1	Balanced
2010	0.818	English
2011	0.833	English
2012	0.64	English
2013	1	Balanced
2014	0.82	English
2015	1	Balanced
2016	0.92	English
2017	0.75	English
2018	1	Balanced
2019	0.9	English
2020	1	Balanced
2022	0.9	English
2023	0.83	English
2024	1	Balanced
2025	1	Balanced
2026	0.9	English
2027	0.73	English
2028	1	Balanced
2029	0.83	English
2030	0.92	English

### Appendix D

**Table A4.** Self-ratings of Written Ability individual dominance scores on a ratio of <1 (English) to >1 (Spanish).

Participant	Dominance Score	Dominant Language
2001	0.583	English
2002	0.916	English
2003	0.916	English
2004	0.727	English
2005	1	Balanced
2006	0.833	English
2007	1.2	Spanish
2008	0.833	English
2009	0.916	English
2010	0.666	English
2011	0.666	English
2012	0.7	English
2013	1	Balanced
2014	0.6	English

Table A4. Cont.

Participant	Dominance Score	Dominant Language
2015	0.83	English
2016	0.83	English
2017	0.5	English
2018	1.2	Spanish
2019	0.7	English
2020	1	Balanced
2022	0.83	English
2023	0.83	English
2024	1	Balanced
2025	0.78	English
2026	1	Balanced
2027	0.56	English
2028	1	Balanced
2029	0.67	English
2030	0.83	English

### Appendix E

Table A5. Repetition Task individual dominance scores on a ratio of <1 (English) to >1 (Spanish).

Participant	Dominance Score	Dominant Language
2001	0.997	English
2002	1.026	Spanish
2003	0.998	English
2004	0.919	English
2005	0.986	English
2006	0.938	English
2007	1.191	Spanish
2008	0.986	English
2009	0.938	English
2010	1.056	Spanish
2011	0.973	English
2012	0.941	English
2013	1.00	Balanced
2014	0.924	English
2015	0.937	English
2016	0.981	English
2017	1.011	Spanish
2018	0.979	English
2019	1.034	Spanish
2020	0.955	English
2022	1.03	Spanish
2023	1.09	Spanish
2024	1.018	Spanish
2025	0.995	English
2026	0.989	English
2027	1.069	Spanish
2028	0.953	English
2029	0.971	English
2030	0.947	English

## References

- Amengual, Mark, and Pilar Chamorro. 2015. The effects of language dominance in the perception and production of the Galician mid vowel contrasts. *Phonetica* 72: 207–36. [[CrossRef](#)] [[PubMed](#)]
- Amengual, Mark. 2016. The perception and production of language-specific mid-vowel contrasts: Shifting the focus to the bilingual individual in early language input conditions. *International Journal of Bilingualism* 20: 133–52. [[CrossRef](#)]
- Antoniou, Mark, Catherine T. Best, Michael D. Tyler, and Christian Kroos. 2011. Inter-language interference in VOT production by L2-dominant bilinguals: Asymmetries in phonetic code-switching. *Journal of Phonetics* 39: 558–70. [[CrossRef](#)] [[PubMed](#)]
- Argyri, Efrosyni, and Antonella Sorace. 2007. Cross linguistic influence and language dominance in older bilingual children. *Bilingualism: Language and Cognition* 10: 79–99. [[CrossRef](#)]
- Bahrack, Harry P., Lynda K. Hall, Judith P. Goggin, Lorraine E. Bahrack, and Stephanie A. Berger. 1994. Fifty years of language maintenance and language dominance in bilingual Hispanic immigrants. *Journal of Experimental Psychology* 123: 264–83. [[CrossRef](#)] [[PubMed](#)]
- Bedore, Lisa M., Elizabeth D. Peña, Connie L. Summers, Karin M. Boerger, Maria D. Resendiz, Kai Greene, Thomas M. Bohman, and Ronald B. Gillam. 2012. The measure matters: Language dominance profiles across measures in Spanish–English bilingual children. *Bilingualism: Language and Cognition* 15: 616–29. [[CrossRef](#)] [[PubMed](#)]
- Birdsong, David. 2014. Dominance and age in bilingualism. *Applied Linguistics* 35: 374–92. [[CrossRef](#)]
- De Bot, Kees. 2001. Language use as an interface between sociolinguistic and psycholinguistic processes in language attrition and language shift. In *Theories on Maintenance and Loss of Minority Languages. Towards a More Integrated Explanatory Framework*. Edited by Jetske Folmer and Piet van Avermaet. Munster and New York: Waxmann, pp. 65–81.
- Deuchar, Margaret, and Rachel Muntz. 2003. Factors accounting for code-mixing in an early developing bilingual. In *Vulnerable Domains in Multilingualism*. Edited by Natascha Muller. Amsterdam: John Benjamins Publishing Company, pp. 161–90.
- Dubiel, Bozena. 2019. The assessment of language maintenance in bilingual children. *TEANGA, The Journal of the Irish Association for Applied Linguistics* 10: 94–112. [[CrossRef](#)]
- Dubiel, Bozena, and Eithne Guilfoyle. 2017. Language strength in bilingual children: The child HALA test. *Heritage Language Journal* 14: 1–29.
- Dunn, Alexandra L., and Jean E. Fox Tree. 2009. A quick, gradient bilingual dominance scale. *Bilingualism: Language and Cognition* 12: 273–89. [[CrossRef](#)]
- Flege, James E., Ian R. A. Mackay, and Thorsten Piske. 2002. Assessing bilingual dominance. *Applied Psycholinguistics* 23: 567–98. [[CrossRef](#)]
- Gertken, Libby M., Mark Amengual, and David Birdsong. 2014. Assessing language dominance with the bilingual language profile. In *Measuring L2 Proficiency: Perspectives from SLA*, 1st ed. Edited by Pascale Leclercq, Amanda Edmonds and Heather Hilton. Bristol: Cambridge University Press, pp. 208–25.
- Golato, Peter. 2002. Operationalizing “language dominance” in late bilinguals. *Working Papers in Linguistics* 1: 26–85.
- Gollan, Tamar H., Rosa Montoya, Cynthia Cera, and Tiffany Sandoval. 2008. More use almost always means a smaller frequency effect: Aging, bilingualism, and the weaker links hypothesis. *Journal of Memory and Language* 58: 787–814. [[CrossRef](#)] [[PubMed](#)]
- Gollan, Tamar H., Gali H. Weissberger, Elin Runnqvist, Rosa I. Montoya, and Cynthia M. Cera. 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: 594–615. [[CrossRef](#)] [[PubMed](#)]
- Grosjean, François. 1998. Studying bilinguals: Methodological and conceptual issues. *Bilingualism: Language and Cognition* 1: 131–49. [[CrossRef](#)]
- Grosjean, François. 2001. The bilingual’s language modes. In *One Mind, Two Languages: Bilingual Language Processing*. Edited by Janet L. Nicol. Malden: Blackwell, pp. 1–22.
- Grosjean, François. 2016. The Complementarity Principle and its impact on processing, acquisition, and dominance. In *Language Dominance in Bilinguals: Issues of Measurement and Operationalization*. Edited by Carmen Silva-Corvalán and Jeanine Treffers-Daller. Cambridge: Cambridge University Press, pp. 66–84.

- Jia, Gisela, Doris Aaronson, and Yanhong Wu. 2002. Long-term language attainment of bilingual immigrants: Predictive variables and language group differences. *Applied Psycholinguistics* 23: 599–621. [CrossRef]
- Lanza, Elizabeth. 2004. *Language Mixing in Infant Bilingualism: A Sociolinguistic Perspective*. Oxford: Oxford University Press.
- Lim, Valerie, Susan Rickard Liow, Michelle Lincoln, and Mark Onsolow. 2008. Determining language dominance in English-Mandarin bilinguals: Development of a self-report classification tool for clinical use. *Applied Psycholinguistics* 29: 389–412. [CrossRef]
- Meuter, Renata F. I., and Alan Allport. 1999. Bilingual language switching in naming: Asymmetrical costs of language selection. *Journal of Memory and Language* 40: 25–40. [CrossRef]
- Olson, Daniel J. 2016. The impact of code-switching, language context, and language dominance on suprasegmental phonetics: Evidence for the role of predictability. *International Journal of Bilingualism* 20: 453–72. [CrossRef]
- Paradis, Michel. 2004. *A Neurolinguistic Theory of Bilingualism*. Amsterdam: John Benjamins Publishing Company.
- Peña, Elizabeth D., Lisa M. Bedore, Vera F. Gutiérrez-Clellen, Aquiles Iglesias, and Brian Goldstein. 2010. *Bilingual English Spanish Oral Language Screener*. Experimental Version taken from BESA: Bilingual English-Spanish Assessment. Baltimore: Brookes Publishing.
- Pérez-Leroux, Ana Teresa, Erin O'Rourke, and Gretchen Sunderman. 2014. Language dominance and codeswitching asymmetries. In *Grammatical Theory and Bilingual Codeswitching*. Edited by Jeff MacSwan. Cambridge: MIT Press, pp. 283–312.
- Saunders, George. 1988. *Bilingual Children: From Birth to Teens*. Clevedon: Multilingual Matters.
- Sheng, Li, Ying Lu, and Tamar H. Gollan. 2014. Assessing language dominance in Mandarin-English bilinguals: Convergence and divergence between subjective and objective measures. *Bilingualism: Language and Cognition* 17: 364–83. [CrossRef]
- Stefanich, Sara. 2019. A Morphophonological Account of Spanish/English Word-Internal Codeswitching. Ph.D. Thesis, University of Illinois at Chicago, Chicago, IL, USA.
- Tomoschuk, Brendan, Victor S. Ferreira, and Tamar H. Gollan. 2019. When a seven is not a seven: Self-ratings of bilingual language proficiency differ between and within language populations. *Bilingualism: Language and Cognition* 22: 516–36. [CrossRef]
- Treffers-Daller, Jeanine. 2011. Operationalizing and measuring language dominance. *International Journal of Bilingualism* 15: 147–63. [CrossRef]
- Treffers-Daller, Jeanine. 2019. What defines language dominance in bilinguals? *Annual Review of Linguistics* 5: 375–93. [CrossRef]
- Tsui, Rachel Ka-Ying, Xiuli Tong, and Chuck Siu Ki Chan. 2019. Impact of language dominance on phonetic transfer in Cantonese-English bilingual language switching. *Applied Psycholinguistics* 40: 29–58. [CrossRef]
- Unsworth, Sharon. 2016. Amount of exposure as a proxy for dominance in bilingual language acquisition. In *Language Dominance in Bilinguals: Issues of Measurement and Operationalization*. Edited by Carmen Silva-Corvalán and Jeanine Treffers-Daller. Cambridge: Cambridge University Press, pp. 156–73.
- Van Hell, Janet G., Kaitlyn A. Litcofsky, and Caitlin Y. Ting. 2015. Intra-sentential code-switching: Cognitive and neural approaches. In *The Cambridge Handbook of Bilingual Processing*. Edited by John W. Scwiete. Cambridge: Cambridge University Press, pp. 459–82.
- Yip, Virginia, and Stephen Matthews. 2006. Assessing language dominance in bilingual acquisition: A case for mean length utterance differentials. *Language Assessment Quarterly* 3: 97–116. [CrossRef]

