AUTHORS' ACCEPTED VERSION

To appear in: Luk, G., Anderson J.A.E., & Grundy, J. (eds.). *Understanding Language and Cognition through Bilingualism. In honor of Ellen Bialystok.*

Chapter 3: Defining bilingualism as a continuum: Some tools and consequences for the study of bilingual mind and brain effects

Jason Rothman[1,2], Fatih Bayram[1], Vincent DeLuca[1], Jorge González Alonso[1], Maki Kubota[1], Eloi Puig-Mayenco[3]

UiT the Arctic University of Norway[1], Universidad Nebrija[2], King's College London[3]

1. Introduction

Prior to the early 1960s, the idea that children in bilingual contexts were somehow disadvantaged for linguistic and cognitive development was popularly held. Without consideration of socio-economic inequalities across groups, often co-occurring with ethnic/racial minority status, it was argued that simultaneous language exposure resulted in confusion, delaying the process of language acquisition and cognitive development, if not, in the extreme, causing mental retardation (Goodenough, 1926; Saer, 1923). In light of what we know today— i.e., minimally, that bilingualism provides no disadvantage relative to monolingualism, a discussion to which we return in greater detail below (see e.g., Meisel, 2011, Serratrice, 2013, Bialystok, 2016, 2017)— it boggles the mind how such ideas could have originated, much less propagated. Perplexity, as is often the case, is only made possible by the clarity of hindsight. In some sense, claims about the extreme disadvantages of bilingualism were effectively inevitable. Simply put, the science of the times did not know better; it was fundamentally flawed.

So, what changed? Was it the case that the social milieu of bilingualism and attitudes toward it have dramatically changed in many monolingual-dominant, Western societies? By some measures, they have at least in some instances. The post Franco dictatorship era of Spain, beginning in 1975, is a good example of the walls of bilingualism oppression falling. In the previous 50 years, from the mid-1920s to the mid-1970s, there was an enforced "Spanish-only" policy that delegitimized (even criminalized) regional languages such as Catalan, Galician, and Basque. Although not without challenges, the overall response to linguistic persecution was a sharp resurgence of institutionalized bilingual language protection, identity, and pride (e.g., Hoffman, 2000; Azurmendi, Larranaga & Apalategi, 2008; Loureiro-Rodriguez, 2008; Urla, 2012). However, such examples do not reflect a more generalized trend overall—there are many areas, perhaps more in recent times, where functional monolingualism and/or attitudes against the utility of bilingualism have become more entrenched. Even if it were the case that a less polarized, more (alternatively) normative view of bilingualism were more widespread in the monolingual parts of the world, it is doubtful that this alone would explain the change in tides and ensuing results of scientific bilingualism research since the 1960s. Most researchers would agree that a (societal) context in which a pre-emptively definitive conclusion has not already been drawn is the best grounds for any scientific inquiry. Yet, the scientific benefit of such is really only the neutrality it provides for the object of inquiry itself. So, what, then, is responsible for the more balanced view on bilingualism, at least within research circles, we have experienced since the 1960s?

In our view, the answer to the above, as it could only be, has been higher standards in the scientific study of bilingualism itself. Peal and Lambert (1962) marks the beginning of the turn. In response to claims of the times, they pondered the extent to which so-called evidence

reflected true deficits within bilinguals or was a byproduct of the modality they were being tested in, all of which related to intelligence measured through language. If it were the case that bilinguals were "confused" and thus had impaired trajectories of cognitive development,
then this should be confirmed in complementary tests of nonverbal intelligence as well. Alternatively, Peal and Lambert hypothesized that the verbal modality tested in the societal dominant language promoted monolinguals to seemingly better success. If on the right track, then, they expected monolingual and bilingual children to show no differences on measures of nonverbal intelligence, but the monolinguals might show higher scores with verbal measures. Their predictions, however, only partially obtained. Surprisingly, the bilingual children outperformed monolingual peers not only on tests related to nonverbal intelligence, but rather on virtually all measures. Peal and Lambert (1962) is the first study to show a potentially enhanced "mental flexibility" of bilinguals. More importantly, at least for the discussion of the present chapter, is its underscoring of the deterministic role of *how* we do the science. Their framing of the (nuanced) questions themselves, controlling for variables such as socioeconomic compatibility and the importance of task (modality) effects opened new doors. Since Peal and Lambert (1962), the controlling of many variables and the responsiveness, ingenuity and flexibility of empiricism have combined with modern advances in technology to immensely increase the ecological validity of studying bilingualism. At present, the study of bilingualism across (psycho)linguistics, psychology and cognitive neuroscience has never been more sophisticated or better aligned with the rigors of the natural sciences. However, this does not mean there is no room for improvement.

The goal of this chapter has the above notion of "room for improvement" at its core. It will do so by honing in on one of potentially many, variables that has substantial room for refinement: *the defining of bilingualism in juxtaposition to the wide spectrum of experiences that give rise to it.* Doing so is timely, as we will contextualize below, on several fronts. Significant strides in unpacking bilingual experiences and charting their correlations to outcomes have recently been made. By experience, we are referring to moments of meaningful engagement with language and, crucially, points in time where a choice (unconscious and conscious alike) between the languages a bilingual has competence in is afforded. The idea is that the distribution of experiences proxies for the opportunity over the lifespan of a bilingual to develop and maintain linguistically and to exercise the mind/brain in ways it has been argued bilingualism can. The relationship(s) between input (quality and quantity), use and proficiency is/are not entirely clear for several reasons, not least given the multifarious ways in which each of these constructs can be conceived. Nevertheless, it should be uncontroversial to state that at some crucial level(s) these three constructs overlap and interact and that their relationships are as dynamic as they are non-linear. Suffice it to say, exposure to language (input) in ample quantity and quality is necessary, although not sufficient, for the growth of competence (on a spectrum of proficiency), which in turn affects among many other factors language use. It is especially fitting that such a discussion should be in a book that honors the seminal contributions of a pioneer in the study of bilingualism *par excellence,* Ellen Bialystok. While she is perhaps best known for her work that created the modern study of the neurocognition of bilingualism since the early 2000s, the impact, ubiquity and breadth of Professor Bialystok's work has laid cornerstones, since the 1970s, in virtually all inquiries into bilingual language acquisition and processing, education, and literacy as well. While in the remainder of this chapter we will focus our discussion on consequences for the neurocognitive study of bilingualism, advocating a more nuanced view of bilingualism as a spectrum of experiences, leading to variation in outcomes is equally important for the sister fields of language acquisition, linguistic processing, and literacy/education.

Not least in light of concerns regarding replication and generalizability extensions for work showing so-called "bilingual advantages", Professor Bialystok was among the first and most vociferous researchers to acknowledge the reality and implications of the existence of null effects alongside positive ones. She welcomed principled and systematic changes to methodological practice that would enable us to better understand the very object of study while increasing predictive validity regarding the conditions under which bilingualism can be expected (and not) to show effects (see e.g., Bialystok, 2016, 2017). Some meta-analyses and systematic reviews have led several to the conclusion that claims of generalized bilingual effects are exaggerated, if warranted at all (e.g., Paap, Johnson & Sawi, 2015, Lehtonen et al., 2018). Others offer evidence casting doubt regarding the existence of generalizable effects for specific domains, such as *interference resolution,* while showing how the totality of data still "suggest bilinguals do enjoy a more widespread cognitive advantage (a bilingual executive processing advantage) that is likely observable on a variety of cognitive assessment tools" (Hilchey & Klein, 2011: 625). Putting aside for the moment other methodological issues related to cross-study comparability such as effect sizes, power within studies, the tendency to not publish null results and the like, a recent meta-analysis by van den Noort et al. (2019) seems particularly relevant to the goal of the current paper. They reviewed 46 original studies on bilingualism and cognitive control tasks, finding that the majority, 54.3%, suggest beneficial effects, 28.3% null effects and 17.4% offer evidence against any type of bilingual facilitation. Of equal importance, they revealed that issues of compatibility across the studies, mostly methodological, among them especially participant selection and individual differences *not* considered, had good explanatory power for the contradictory results across the studies.

Historically, the main criterion for inclusion in bilingualism studies of all types, but especially within the neurocognitive literature, has been some sort of qualifying proficiency measure that relegates participants to one or another group (typically monolinguals versus bilinguals). While it is already troubling for cross-study comparison reasons that the quality of such measures vary considerably, from a single 'yes'/'no' question (are you bilingual?), to self-reporting of perceived language skills, to formalized proficiency measures of vocabulary and/or grammar, it is far from clear that proficiency per se should be a particularly good variable for determining the status of one's *-lingualism*[1]. Pertaining to the knock-on effects this has for meta-analyses, Leivada, Duñabetia, Westergaard & Rothman (2021) call our attention to the following:

> *On the whole, recent meta analyses and systematic reviews give serious cause for reflection, if not concern [...] That which can be understood (better) from a meta-analysis or systematic review is inherently related to the actual appropriateness of bringing data sets together in the first place, that is, their comparability [...] Failure to get this crucial condition right could translate to comparisons of proverbial apples to oranges, the blending of which fails in the most essential ways to ensure confidence for meaningful conclusions that sound meta-analyses provide. In light of the provisos discussed in van den Noort et al.'s (2019) work, if methodological differences reduce the similarity/comparability of data sets to a significant degree then we must consider what consequences these have [...] since bilingualism itself is defined distinctly in many studies, i.e. often not treated as the spectrum it is, we must ponder what the consequences are of collapsing data across studies with participants of vastly different bilingual profiles in these meta-analyses and systematic reviews (Leivada et al. XXXX:XX)*

There are several reasonable conclusions in light of the above, two of which sit at the fore of the present discussion. The first is that any definitive conclusion, especially those

[1] Proficiency as such is typically assessed with a qualitative standard baseline in mind (explicitly or implicitly), to which many bilinguals do not have access. Given that the underlying mechanisms argued to result in neurocognitive bilingual effects relate to a tension between activation and suppression, one should expect effects so long as there are two systems, regardless of the subjective quality of any given system, and those

systems are engaged. Of course, higher proficiency could be collinear with greater engagement, that is, the more normative proficient one is the more domains they have to use both languages. This could then give the impression that proficiency matters, but in such proxies for a correlation with language engagement patterns (DeLuca, Rothman & Pliatsikas, 2018).

supported by the perceived strength of meta-analyses/systematic reviews, about the effects bilingualism may or may not have on neurocognition are at present premature. The second is that any future cross-study claims are destined to remain murky and inconclusive to the extent they fail to adopt a scalar definition of bilingualism, augmented by a testing battery able to capture quantities and qualities of input and opportunities to engage and use languages across time and space.

Are we to assume that a naturalistic Spanish-speaking child bilingual growing up in Catalonia versus one growing up in New York have had comparable experiences with bilingualism? Are we to take for granted that Spanish-English bilinguals in New York are all the same as well, irrespective of when and how they learned the language (at home or in school, as a child or adult), ignoring the ensuing opportunities and patterns of use that distinguish them? We presume the answer to this is an overwhelming "no". It is at a minimum an important question that must be determined empirically. Yet, within meta-analytic approaches to addressing serious debates in the literature, we have those that effectively answer the question for us by offering conclusions from bringing studies together that do not meet the minimum standards of comparability required (see Grundy, 2020 for similar argumentation and meta-analytic data that support the general concern). No one denies that there is a *now-you-see-it-now-you-don't* quality to bilingual effects in the literature; the question becomes why this is so. With others, we maintain that part of the answer to this question lies in the general lack, until recent years, of collecting and correlating data regarding input quantity, the dynamic patterns of daily language use (in real and apparent time) and the extent of the bilingual experience related to the age where actively using two languages commences. In other words, we should be treating bilingualism as a continuous variable, as opposed to a categorical one. Uncovering, then, what the parameters of the bilingual experiences are that have the best chance to confer effects is not only a crucial determinant for greater efficacy, precision and cross-study comparability in our field of research, but is an important variable for understanding if there are bilingual effects at all and what, if any, the real world implications are. In our view, unraveling the manifold hues and shades of *-lingualism*—where mono-, bi-, and multi- sit on a continuum—is an imperative.

Recent research hints at the potential utility of treating bilingualism as a spectrum, for example, studies show how specific experiences related to bilingualism (exposure, domains of use, etc.) correlate to greater probability at the individual level of behavioral effects with executive functions, more efficient neuronal recruitment and/or neuro-anatomical change (see e.g., Bialystok 2016, DeLuca et al. 2019, 2020; Dash et al., 2019; Gullifer et al. 2018; Li, Legault, & Litcofsky, 2014; Luo et al., 2019; Nichols & Joannisse, 2016; Sulpizio et al. 2019; Yamasaki, Stocco, & Prat, 2018). How studies measure these dynamic variables is of significant importance. Below, we provide the reader with a sample of five increasingly used measures that are designed to record or gauge proxies for degree of input quality, quantity, and how language is used across space and time. Having more than one measure available is ideal because each is designed to best capture distinct aspects of bilingual experiences and might be more and less appropriate for specific questions, specific types of bilinguals and/or specific ages. What they have in common is the view that understanding the complexities of bilingual outcomes requires a view of bilingualism that can be adjusted for a bespoke fit. Of course, the represented selection is not an exhaustive one. No judgement or evaluation should be interpreted from inclusion or exclusion in our selection. Rather, we have chosen these five measures because they are good examples of how one can approach collecting relevant data. In all cases, they have been widely used and/or been especially influential relative to their emergence (obviously older measures are more likely to have been used more at the time of writing this), and where they do not overlap they are complementary to each other.

2.1 What is it?
The Bilingual Language Experience Calculator (BiLEC) is a language background questionnaire developed by Sharon Unsworth and her colleagues in 2008. It is an open-source questionnaire that is freely available to download via IRIS (A Digital Repository of Instruments and Materials for Research into Second Language): https://www.iris-database.org/iris/app/home/detail?id=york%3A928327&ref=search. BiLEC was developed to gather detailed and quantified information about the language experiences, especially related to input exposure, of bilingual children (Unsworth, 2013). It can be used for both simultaneous and sequential bilinguals, as well as trilinguals from the ages of 2 to 18 years. The questionnaire requires the investigator to interview the parents about their children's language background, language history, and daily use of the language, etc. Responses are directly entered into an Excel spreadsheet by the investigator, in which specific algorithms are applied to generate a number of quantity-oriented and quality-oriented composite measures of children's language experience. For instance, quantity-oriented measures include: the child's relative exposure to each language at the current time; the child's relative use of each language at the current time; the child's relative exposure to each language over time, while quality-oriented measures include: the average quality of exposure to each language at the current time; the number of different (native/non-native) speakers providing input; and the number of different interlocutors with whom the child communicates exclusively in a given language.

2. The Bilingual Language Experience Calculator (BiLEC)

2.2 How is it measured?
As mentioned above, algorithms are embedded into the spreadsheet to automatically calculate the composite measure. For instance, the current amount of language exposure is derived by first calculating the total number of waking hours, the total number of hours each person spends with the child at home, and the total number of hours' exposure at school, as well as sources outside of home and school. These numbers are then added up which equals to the total number of hours of exposure to the target language per week. Subsequently, this value is further divided by the total number of waking hours per week which then can be used to ultimately derive the current average percentage exposure to the target language per week. In order to understand exactly how the algorithms work for each composite measure, we refer you to the manual available at the above link as well. It takes around 20 to 30 minutes to complete the questionnaire and it is recommended that the researchers administer the questionnaire via Excel with caregivers individually.

2.3 Why is it used?
BiLEC has been widely used among researchers who examine the language development of bilingual children (De Cat, 2018; Garraffa et al., 2015; Hendricks et al., 2018; Kan, 2019; Kubota et al., 2020; Persici et al., 2019; Robinson & Sorace, 2019; Serratrice & De Cat, 2020; Stoehr et al., 2019; Sun et al., 2016; Vender et al., 2018; Verhagen et al., 2020; Wolleb et al., 2018; Andreou et al., 2015). BiLEC differs from other questionnaires (such as LEAP-Q, LSBQ, and BLP) in the sense that it provides separate measures for quantity and quality of exposure (i.e., quantity of exposure is measured in percentages while quality of exposure is measured in rankings (1-5)). However, the majority of studies that used BiLEC to elicit information about the bilinguals' language experience have reported their quantity of exposure, mostly focusing on the following two measures: current quantity of exposure and/or cumulative length of exposure (De Cat, 2018; Garraffa et al., 2015; Hendricks et al., 2018; Kubota et al., 2020; Serratrice & De Cat, 2020; Stoehr et al., 2019; Unsworth et al., 2018; Vender et al., 2016, 2018; Wolleb et al., 2018). The cumulative length of exposure appears to be a better alternative measure to the traditional length of exposure, since it takes into account the exposure to the

target language for each year of a child's life, rather than simply calculating the time elapsed from the date of first exposure to the time of testing. In fact, to our knowledge, Unsworth et al. (2019) is the only study that utilized information about the quality of exposure by differentiating the proportion of exposure at home that was from native versus non-native speakers. Other studies have simply used BiLEC to outline the basic information of the bilinguals such as their age on onset to L2, parental education, and self-reported proficiency of the child and parent (Andreou et al., 2015; Kan, 2019; Sun et al., 2016).

2.4 How has it been used?

BiLEC has been demonstrated to reliably predict linguistic performance in bilingual children. The measures extracted from this questionnaire have been used to predict various dimensions of linguistic knowledge/processing such as gender (Hendricks et al., 2018; Mitrofanova et al., 2018; Rodina & Westergaard, 2017), referential expressions (Serratrice & De Cat, 2020), clitic pronouns (Vender et al., 2016, 2018), classifiers (Kan, 2019), voice onset time (Stoehr et al., 2018, 2019), and vocabulary (Sun et al., 2016; Unsworth et al., 2019). Although limited, some studies (Kubota et al., 2020; Robinson & Sorace, 2019; Verhagen et al., 2020) have used BiLEC measures to predict non-linguistic general cognitive abilities in the bilingual population.

2.5 What are some limitations?

Although the composite measure of BiLEC can be viewed or interpreted to capture some type of bilingual experience, it clearly focuses on quantifying (and qualifying the absolute value of language input and output. Thus, the BiLEC is limited in capturing the unique, variable and especially dynamic "experiences" of bilingual language engagement. Other questionnaires such as LSBQ or LEAP-Q described below that consider *how* bilinguals use the languages and *under* what conditions and/or contexts may be more suitable to operationalize the extent to which speakers actively engage in bilingualism. Furthermore, there are no questions pertaining to language mixing or code-switching in this questionnaire, which is an important facet of bilingualism to consider when correlating it to domain-general cognitive function abilities.

3. The Language Experience and Proficiency Questionnaire (LEAP-Q)

3.1 What is it?

The Language Experience and Proficiency Questionnaire (LEAP-Q; Marian, Blumenfeld & Kaushanskaya, 2007) is arguably the first systematic attempt at providing an instrument to capture some of the most important variables contributing to the wide range of bilingual experiences. It can be used with adolescent and adult bilingual (and multilingual) speakers, sequential or simultaneous, ranging from 14 to 80 years of age. The goal of the LEAP-Q was to establish some reliable correlations between lab based behavioral measures and self-reported variables. As the authors themselves explain in a recent retrospective piece (Kaushanskaya, Blumenfeld & Marian, 2019), the questionnaire was developed to address concerns of replicability across studies in bilingualism that may have to do with a lack of comparability between the different bilingual populations used in these studies.

3.2 How is it measured?

The LEAP-Q can be completed in a pencil-and-paper form, electronically (as a form-based MS Word file) or via interview with a researcher. The questionnaire employs a series of 11-point (range: 0-10) Likert scales to gauge information about language exposure/use and elicit self-rating evaluations of different factors related to language proficiency and language dominance. The first page is designed to gather personal and biographical data (education level, cultural identification, date of arrival in the country if applicable) and to establish the set of languages that the speaker has some knowledge of, which must be listed both in order of dominance and in order of acquisition. Three general questions ask the speaker to indicate their relative use of each language, and their relative preference for each in reading and speaking contexts. The second part of the questionnaire devotes one page to each of the languages, focusing on six main areas: age milestones in language learning, contributors to language acquisition, degree of language immersion, extent of language exposure, self-reported proficiency and self-reported foreign accent, if applicable (Kaushanskaya et al., 2019).

The digital version takes approximately 15 minutes for speakers of two languages. Each additional language covered in the second part adds about 5 minutes to this estimate. Administration through interview is likely to increase time by about 5 to 10 minutes (Kaushanskaya et al., 2019). In the questionnaire's website (https://bilingualism.northwestern.edu/leapq/) the authors provide instructions on how to export the data from the electronic version of LEAP-Q into a spreadsheet for data management. Recently, an implementation of the test in the Qualtrics online platform has been made available on the website, together with data export notes. Note, however, that unlike some of the other questionnaires and instruments reviewed here, the LEAP-Q does not attempt to calculate composite scores (for discussion, see Kaushanskaya et al., 2019; Marian, 2019)—although some studies have aggregated the questionnaire's estimations into single indices (e.g., Dimitropoulou, Duñabeitia & Carreiras, 2011; Krizman et al., 2014).

3.3 Why is it used?

The LEAP-Q was originally meant to serve at least two different purposes: to provide a better motivated and more comparable distribution of discrete categories (e.g., simultaneous vs. early vs. late bilinguals) that were already being used to subdivide bilingual participant samples, and to ensure that certain threshold-based criteria for inclusion were standardized. To do this, the questionnaire focuses on the estimation of two very different sets of magnitudes. The first is the relative weight of each language in different constructs, most notably *dominance* and *proficiency*, that are often claimed to correlate with certain differences in the linguistic behavior for speakers of the same language—even those with roughly similar linguistic profiles (e.g., Birdsong, 2014; Hulstijn, 2012). This is meant to enable researchers to establish fine-grained distinctions in populations of, for example, highly proficient bilingual individuals living in a bilingual society, for whom dominance or relative proficiency may be hard to determine informally or infer from various sociolinguistic proxies. The second aim of the LEAP-Q is to delve deeper into the speaker's experience and relationship with each of their languages, including exposure history, contexts of use and self-perceived proficiency and (foreign) accent.

3.4 How has it been used?

The impact of LEAP-Q has been significant. In 13 years, it has been translated to 24 languages and the original article has been cited well over a thousand times, by studies examining very different types of bi-/multilinguals in very different settings (bilingual societies, immersion settings, foreign language classrooms, to name a few; see Kaushanskaya et al., 2019). Although categorical assignment to different bilingual groups is a common application (e.g., early vs. late, Dimitropoulou et al., 2011) the specific uses that the questionnaire has been given exceed the original scope. Studies employing the LEAP-Q with late sequential bilinguals, for example, have made use of the questionnaire's self-rated proficiency measures to establish inclusion criteria based on minimum proficiency (e.g., Stocco & Prat, 2014). Conversely, some studies have sought to establish a low-pass threshold of L2 proficiency to justify the consideration of a given group as 'monolingual' (e.g., Ansaldo, Ghazi-Saidi & Adrover-Roig, 2015). Oftentimes, however, the questionnaire is simply used to provide a more comprehensive description of bi-/multilingual experimental groups (e.g., González Alonso, Baquero Castellanos & Müller, 2016; Wang et al., 2014), or to ensure a certain L1-status or dominance in societal bilingualism contexts (e.g., Libben & Titone, 2009).

The LEAP-Q has also been used to investigate dominance effects in bilingualism. Keating, VanPatten and Jegerski (2011) compared antecedent preferences for null and overt pronouns in two groups of Spanish heritage speakers in the US, depending on their self-declared dominance (English or Spanish) in the LEAP-Q. Similarly, self-rated proficiency as measured by the LEAP-Q was used by Reichle and Birdsong (2014) in an ERP study investigating proficiency effects in the interpretation of informational vs. contrastive focus in L2 French.

3.5 What are some limitations?

The most apparent limitation of the LEAP-Q is one of scope: the questionnaire is designed to gauge linguistic experience from bi-/multilinguals aged 14 to 80, which leaves out a critical age range for researchers interested in language development. In terms of the kind of factors contributing to the bilingual experience that the questionnaire is designed to capture, language mixing is notably absent. Finally, the self-rated proficiency estimates of the LEAP-Q are potentially limited towards an accurate characterization of bi-/multilingual language ability (e.g., Shi, 2011; see Kaushanskaya et al., 2019, for discussion), despite having been shown to correlate with some behavioral measures of language comprehension in the original validation studies.

4. Entropy- a measure of both social (language) interaction and language maintenance

4.1 What is it?

The concept of entropy is ubiquitous across a number of fields of study but can generally be defined as the degree of randomness or uncertainty within a closed system. Within cognitive science and linguistics, the concept of entropy has been operationalized in several ways including processes related to input in language acquisition (e.g., Jäger, 2007), diversity in interaction related to language use (Gullifer & Titone, 2020), and as a means of explaining processes underlying language maintenance and attrition (Iverson & Miller, 2018). Here, we focus on the latter two of these examples. Differently from the other tools discussed within this chapter, entropy is discussed herein as both (i) a theoretical construct or (ii) a calculation which can be applied to any data collected by a variety of methods related to language use distribution (i.e., it could be calculated over (a subset) of data collected via the other tools described above). In principle, it is thus usable with any age range.

Regarding social interaction, entropy has been operationalized to quantify the nature and extent of the compartmentalization or integration of one’s available languages on a continuum (Gullifer et al., 2018; Gullifer & Titone, 2020). Gullifer and colleagues adapted the principles from Shannon entropy, which, in brief, quantifies the degree of uncertainty or disorder within a system (Shannon, 1948). Applied to language use, low entropy scores (close to 0) represent full separate language uses (e.g., only one in any given social context), which indicate relatively low uncertainty about the nature of linguistic input. By contrast, high scores (up to 1) indicate highly integrated languages (each language used equally) in each social context, which indicate greater uncertainty about the nature of linguistic input.

Entropy has also been proposed to describe the relationship between language exposure and its maintenance/attrition (Iverson & Miller, 2018; Miller & Rothman, 2019). This notion of maintenance, as applied to language attrition is discussed within the Equilibrium Hypothesis (Iverson & Miller, 2018), which predicts that the degree of attrition is tied to the intensity and nature of exposure and use of each language (its relative entropy, which can shift over time). Increased focus on the non-native language (in this case, the competing system) and reduced focus on the native language lead to a situation of increased entropy within the system and thus attrition of the native language.

4.2 How is it measured?

Entropy within social (language) interaction (see Gullifer & Titone, 2020 for full detail) is measured by calculating a proportion of use for each language in given contexts, and then calculating an entropy score by summing the logged proportions of each language and multiplying by -1 to yield a positive entropy value. In an example provided by Gullifer and Titone (2020), a bilingual who speaks English 80% of the time and French 20% at work would have an entropy score of 0.72 for this context, indicating a higher degree of integration of the two languages. A fully documented software package is provided by the authors to calculate entropy scores (https://github.com/jasongullifer/languageEntropy), both within and across contexts (e.g. work, home, study, etc.).

4.3 Why is it used?

With respect to language experience, the aim of entropy is to coalesce degrees of language exposure or use into a single scale representing the diversity of one's language use patterns. The use of language entropy as a measure allows one to quantify the degree of uncertainty that a speaker is faced with regarding which language is needed in any given context. This measure allows for (among other things) comparisons against neurocognitive outcomes related to these experiences (see below for discussion of examples of use of language entropy in relation to neurocognitive outcomes). Different (language) interactional contexts are argued to confer increased demands on several control processes, including conflict monitoring, interference suppression, cue detection, task switching, among others (see e.g. Green & Abutalebi, 2013; Hartanto & Yang, 2016). However, previous work has primarily used dichotomous categorization of these experiences into discrete groups (e.g., into single- or dual language contexts). A major advantage, then, of the use of entropy scores is that it allows for a more nuanced exploration of the overall distribution of bilingual language use both within- and across different social/conversational contexts.

4.4 How has it been used?

Entropy, of language exposure, has been found to correlate with specific neurocognitive adaptations (Gullifer et al., 2018). Specifically, Gullifer and colleagues found increased diversity (entropy) of language use to relate to increased reliance on proactive control strategies while completing an AX-continuous performance task (AX-CPT) which also correlated with increased functional connectivity between the left putamen and anterior cingulate cortex. These results are interpreted as neurocognitive adaptations towards increased reliance on monitoring and selection processes. Entropy of language exposure is now also input as an automated calculation in a new iteration of the language history questionnaire (LHQ3; (Li, Zhang, Yu, & Zhao, 2019), both overall and in several social/conversational settings.

Entropy, as relates to language attrition, was recently applied in a study by Miller & Rothman, (2019) who examined aspects of pragmatic processing in native Spanish-English bilinguals living in an English (L2) immersion context. The authors crucially report that prolonged exposure to the immersion environment lead to an over-acceptance of a specific reading of 'some' (*algunos*) in what would be considered infelicitous contexts to do so. Crucially, the degree to which this occurred was linked to several aspects of language exposure and use, specifically the balance between L1 and L2. These results support predictions of the equilibrium hypothesis, specifically a continuous and nuanced relationship between reduced L1 exposure, entropy, and L1 attrition.

4.5 What are some limitations?

Language entropy was conceived of and thus designed to capture a specific aspect of bilingual experience (diversity of bilingual language use within context). It is thus, by design, unable to quantify other aspects of language experience, including language mixing, duration of use (etc.). It is, then, perhaps most optimally used in complement with other continuous measures of bilingual experience (for an example of this see Gullifer et al., 2018). Furthermore, given the way in which entropy is currently calculated, dominance towards one language or another (e.g, an L1 or L2) cannot be distinguished with this measure. In other words, if someone becomes much more dominant in an L2 and uses that language much more than their L1 (a common occurrence for heritage speaker bilinguals, (Rothman 2009) and/or L1 attriters (Schmid, 2013)), irrespective of age at which the shift in dominance occurs, entropy will not distinguish between these cases from a traditional adult L2 learner whose linguistic distribution significantly favors their L1 as does their dominance. Obviously, there are good reasons to believe that age of acquisition is an important factor worthy of being captured within the construct of entropy.

5. The Language and Social Background Questionnaire (LSBQ)

5.1 What is it?

A first version of the 'Language and Social Background Questionnaire' (LSBQ) was developed by Luk and Bialystok (2013) and later updated by Anderson et al., (2018). The LSBQ focuses on the dynamics and complexities of bilingual language usage patterns, crucially in different contexts related to daily life. The goal of the LSBQ is to quantify the multifaceted bilingual experience, offering several composite scores related to how a bilingual individual might distribute the use of each language in distinct domains of daily life, for example, at home versus at work versus purely social contexts. This means that the LSBQ offers multiple composite scores. Each can be used in statistical modeling to better capture how potential individual differences within bilingual experiences and engagement, specifically in ecologically valid domains of use where they conceivably differ, correlates to any given measurement outcome. The current version of the LSBQ partially overlaps with its predecessor version (Luk & Bialystok, 2013) and shares some commonalities with other similar tools such as the Language Experience and Proficiency Questionnaire (LEAP-Q; Marian, Blumenfeld, & Kaushanskaya, 2007) and the Language History Questionnaire (LHQ 2.0; Li, Sepanski, & Zhao, 2006; Li, Zhang, Tsai, & Puls, 2014). In its latest version, it has been validated across a relatively large sample size. Separate versions are available for three different age groups across the lifespan: young children, young to mid-age adulthood and older adults.

5.2 How is it measured?

The LSBQ is based on self-report and self-assessment with questions about individuals' language use in four main dimensions: (i) time (pre-school, primary school, high school), (ii) setting (home, school, work, religious activities, work, etc.); (iii) interlocutor (friends, parents, partner, etc.); and (iv) activity (social media, praying, watching tv, etc.). The LSBQ's compensation for the reliability of self-reports and its internal validity was established by using multiple questions that are shown to be reliably related in exploratory factor analysis (EFA) with data from 408 adult participants with various non-English languages in a context where English is the majority language. The EFA showed that 74% of the variance was captured by three main factors: (i) non-English home language use and proficiency, (ii) non-English social language use, and (iii) English proficiency. Based on these, the questionnaire is comprised of three sections containing 22 items with sub-items: (i) Social Background serving as a proxy for socioeconomic status (SES) with questions regarding age, country of origin, education, immigration status, etc.; (ii) Language Background serving as a proxy for language proficiency with questions regarding the language(s) the participant is able to understand, speak, read and write, at what age and in what context they learned it; and (iii) Community Language Use Behavior serving as a proxy for language use in different stages of life, different contexts, different situations and different activities. Anderson et al. (2018) tested the internal and construct validity of the LSBQ by correlating the composite factor score as well as the factor scores with previously collected behavioral cognitive data from a subset of participants confirming previous findings: higher composite factor scores (e.g., the more bilingual one is) predict better non-verbal performance and poorer verbal performance.

5.3 Why is it used?

The LSBQ is available free of charge in an online repository (see https://figshare.com/articles/journal_contribution/The_Language_and_Social_Background_Questionnaire_Assessing_Degree_of_Bilingualism_in_a_Diverse_Population_Supplementary_Materials/3972486/5) with an LSBQ Administration/Scoring Manual and an Excel file allowing researchers to calculate factor scores. It yields a composite score serving as a tool for categorical assignment of individuals into groups along a monolingual to bilingual spectrum. Based on their analysis of data from young adults, Anderson et al. (2018) conclude that a composite score of less than −3.13 would categorize one as monolingual while individuals with a score above 1.23 would be regarded as bilinguals. More uniquely, it also produces several factor scores which can be used as continuous measures quantifying the bilingualism experience in various contexts/stages of life. In other words, it goes beyond establishing group membership and offering proxies for proficiency. The LSBQ enables the quantification at an individual level of precisely how any given bilingual with any given linguistic proficiency has experienced and continues to engage with distributing their languages' usage. All researchers studying a myriad of questions related to bilingualism today—from the most linguistic to the most cognitive neuroscience perspectives—are beginning to understand the indispensability of treating individual bilingual outcomes as a consequence of the interplay of a wide spectrum of complex contributory variables. Tools like the LSBQ thus stand out precisely because they not only encourage, but indeed make possible the capturing and quantifying of bilingual experience patterns in a nuanced way.

5.4 How has it been used?

Today, the LSBQ is increasingly becoming one of the most widely used language background questionnaires for linguistic development and outcomes (e.g., Luk and Bialystok, 2013; Peets et al., 2020; Altinkamis & Simon, 2020; Friesen et al., 2015; Mitits et al., 2018) as well as bilingual effects on executive function, working memory, neuroanatomical structure and aging (e.g., DeLuca et al., 2019, 2020; Bialystok et al., 2014; Smith et al., 2019; Sullivan et al., 2017; Barac et al., 2016; Zhang et al., 2020; Anderson et al., 2017). For instance, in a recent MRI study with 65 adult participants, DeLuca et al. (2019) used the LSBQ to calculate factor scores of language experiences in bilingual language use to examine their effects on brain structure and functional connectivity. Their results show that brain adaptations, both structural and functional, are modulated at the individual level by bilingual language experiences of exposure and use. In another study, Peets, Yim and Bialystok (2020) used the LSBQ scores to see how the context of language use and exposure affected bilingual children's performance measured by standardized vocabulary, grammar and reading comprehension tests compared to their monolingual peers. Their results showed that on average there were no group differences; however, for the bilingual group performance in reading comprehension was modulated by oral language and home literacy.

5.5 What are some limitations?

A potential limitation of the LSBQ concerns its focus on bilingualism in the context of only two languages, or at least its lack of overt consideration for multilingualism. That is, in its current form it cannot capture the depth of use and experience in more than two languages. Of course, it is often the case, especially outside of North America, that bilinguals are exposed to and use more than a single additional language. And so, assessing multilingualism per se is not an option via the LSBQ. Underscoring this observation, it is prudent to highlight that Anderson et al.'s (2018) attempt to account for multilingualism in another factor analysis failed to converge. Another limitation, depending on one's interest in using this tool, relates to specific measurements of bilingual engagement that one might need or wish to calculate. For example, language switching is captured only by three questions and one universal type of switching, namely “language-switching”.

6. The Bilingual Language Profile (BLP)

6.1 What is it?

The Bilingual Language Profile (BLP; Birdsong et al., 2012; Gertken et al., 2014) is a tool designed specifically to determine language dominance in bilingual speakers. Even though it is not directly stated in the BLP, the type of questions elicited are best suited for adult bilingual speakers. The BLP assumes that language dominance within each unique bilingual speaker is a construct informed by a combination of different factors influencing the use of each of the languages. Via answers to a series of self-reported questions, the BLP generates a score of language dominance which is both general and context-independent. Under the BLP, the notion of language dominance is treated within a gradient spectrum, quite uniquely for other tools of language dominance assessment (see Solís-Barroso and Stefanich, 2019, for a review). In doing so, it allows for a more fine-grained characterization of a bilingual’s language dominance, and correlated variation, that may otherwise be obscured with measures employing a categorical/dichotomous distinction for language dominance.

6.2 How is it measured?

The BLP calculates a composite score, rendered on the basis of 19 questions distributed across four different modules: (a) *language history*, (b) *language use*, (c) *language proficiency* and (d) *language attitudes*. Unlike other measures of language dominance, the BLP does not weigh the different factors (or modules) differently. Each of the four modules carries the same weighting so that the composite score does not bias results towards any of the factors/contexts included in the questionnaire. The composite score falls within a continuum ranging from -218 to +218, whereby the extreme values represent extreme dominance in one language or the other and the middle points represent more balanced bilingualism.

The BLP can be accessed online (https://sites.la.utexas.edu/bilingual/), where all the testing materials, protocols and instructions to calculate the composite score can be found. The questionnaire can be distributed in either pen-and-paper or web-based format. In total, it takes between 5 and 10 minutes for an adult literate bilingual speaker to complete the questionnaire. The authors of the BLP provide an algorithm in *Google Sheets* that automatically calculates the score for language dominance. The BLP is currently available in 15 different language pairings and for each pairing the BLP is available in both languages.

6.3 Why is it used?

The BLP was originally designed to address the call for standardization of language dominance measures, in consideration of the bilingual profile of speakers of a variety of linguistic combinations and different cultural backgrounds. To this end, the BLP gathers general information that is not specific to any context, type of acquisition or language per se, this allows one to create a general profile of language dominance irrespective of the specific context of acquisition.

As acknowledged by the authors themselves, the BLP can serve two purposes: (i) as a composite score regressor to explain variation and (ii) to provide demographic information of specific populations relative to an individual’s linguistic dominance for various classification purposes.

6.4 How has it been used?

The BLP was originally designed to assess the bilingual dominance in early child bilingualism and L2 acquisition. It has since then been extended to contexts of adult heritage language bilingualism, L1 attrition, L3 acquisition and aphasia in bilingualism (e.g., Cabrelli, Luque, & Finestrat-Martínez, 2019; Faroqi-Shah, Sampspon, Pranger, & Baughman, 2018; Puig-Mayenco & Rothman, 2019; Zyzik & Sanchez, 2019). Much of the research to date using the BLP has examined aspects related to the phonetics and phonology of bilinguals (e.g., Amengual, 2013, 2016; Casillas, 2015; Simonet, 2014; Simonet, Casillas, & Díaz, 2014). More recently, however, it has also been used to examine other domains of language: bilingual morphological processing/priming (Bosch, Veríssimo, & Clahsen, 2019; Veríssimo, Heyer, Jacob, & Clahsen, 2018), lexical recognition (García, Leibold, Buss, Calandruccio, & Rodríguez, 2018), sentence processing (Gertken, 2013), semantic generalizations (Grégoire & Greening, 2020), amongst others. Crucially, the results of these studies point towards the direction that using a measure that conceptualizes language dominance to be a gradient variable has the potential to explain a good portion of variability found within bilingualism that might go unnoticed if language dominance is (a) treated as a categorical variable or (b) confounded with other variables such as order of acquisition or language proficiency.

6.5. What are some limitations?

Although the BLP is a relatively quick and easy questionnaire, its brevity--much appreciated for other reasons in research design— sacrifices detail one would expect, if not require, to capture a finer-grained view of bilingual experience. For example, outside of its remit to quantify and qualify language dominance, the BLP offers no means to collect information regarding language mixing or code-switching. In some ways, then, the BLP focuses on the division of the two languages insofar as they exist independently of each other, whereas additional questions probing language mixing could help to reveal the more ecologically valid reality of many bilinguals where the two languages are overlapping. A further limitation is that the BLP seems to be designed with a specific set of bilinguals in mind, namely those most prolifically studied in university settings (younger adults). One example highlighting this is the BLP's inability to capture language dominance straightforwardly in older populations. Questions quantifying how long a participant has spent using each language in different contexts (e.g., family, workplace) offer a maximum response range of 20 years, thus making it difficult, if not impossible, to quantify variation in experiences that go beyond this.

7. Conclusion

It seems clear that, if we are to understand the impact bilingualism can have on neurocognition (or, for that matter, language processing, rate and outcome of linguistic acquisition, educational outcomes, etc.), we need to employ more precise methods to capture and quantify individual experiences with the activities that render effects more and less likely. Quantitative and qualitative proxies for bilingual language experiences, such as those provided in the five tools reviewed above, are wonderful steps in the right direction. Time will tell how fine-grained and useful they really are. Time will also reveal if what they capture is sufficient and well-motivated, that is, if they focus on enough and/or the right subset of variables for the questions we have. It will also reveal whether each of these tools is equipotential to capture the necessary granularity required to answer the questions of today and how adaptive they will be for the questions of tomorrow. To be clear, this main claim, in our view, follows directly from the underlying mechanisms purported to give rise to neurocognitive adaptations in the first place. The greater or lesser the intensity, the longer or shorter the duration, the closeness or distance in time of a cognitively demanding context, such as bilingualism, should result in differentiation of neurocognitive adaptations (at the synchronic time of testing). After all, the tension between activation and suppression of both languages would differ as a result and likely in highly dynamic ways. How the brain responds to this and becomes more efficient is likely linked to how predictable and intense the need for inhibition is, over various epochs of time. This need will be related to bilingual experiences.

Bilingualism should, in principle, be no different from other activities claimed to give rise to cognitive and neural reserves, such as the navigational skills of taxicab drivers in London (Maguire, Frackowiak & Firth, 1997). However, the ubiquity of bilingualism and its variation of related experiential factors places it in a league of its own. Taxi drivers, at least before the advent of Google maps and similar technology, were forced by their very nature to engage in very highly intense spatial navigation as a function of all hours working, at least in cities like London. When engaged, individuals were completely engaged often for 8 or more hours a day in a very similar way. But how many people are taxi drivers? By comparison, approximately half of the world's population is bilingual. Languages are always simultaneously activated, thus regardless of need or intention, making it a 24-hour context. How different are the daily navigational experiences, in abstract, of driving taxis in major cities (at least before Google maps)? By comparison, there are probably hundreds if not thousands (or more) permutations with respect to the interplay of multiple factors delimiting bilingual experiences any one bilingual has, not to mention that these change (sometimes dramatically) over time for the same individual. The bilingual and how she engages with bilingualism is indeed a moving target. And so, while it is true that the management of more than one system is a 24-hour mental job so to speak, it stands to reason that how the brain adapts to it in terms of efficiency will reflect the variation of its engagement reality.

Our task then is to uncover what explains variation. While mind and brain effects have always been envisaged as adaptive byproducts that would vary, variation should not be random. When it seems random, it likely only means that the patterns of experiences to outcomes as yet eludes us. Understanding that there is not a one-size-fits-all definition of bilingualism has a long history, if only reflected in labels that differentiate subgroups such as L2 learners, simultaneous bilinguals, heritage speaker bilinguals, passive bilinguals, code-switchers, balanced bilinguals, and many more. However, recognizing some differences by means of labels does little to capture, if recognize at all, the deterministic variation that sits under each label and across them. For far too long, studies have been populated by groups of so-called bilinguals that cut across the many labels themselves as if the conditions that separate them would not matter. Equally, for most studies that limit their studies to only one subset (young simultaneous bilinguals, adult L2 learners in immersion), they have functionally ignored vast differences in individual bilingual experiences, as if all have the same proportions or, otherwise, indirectly claiming such should not matter.

As discussed in the introduction, this alone could be one of the primary reasons we see such discrepancies in claims within the relevant literature. Comparing "apple" populations to "orange" populations simply because they all happen to be (bilingual) fruit is not the right way forward to adjudicate between competing theories, hypotheses and claims. Setting new standards for how we document and disclose the backgrounds and relevant experiences of bilinguals so we have "apple" to "apple" comparisons will enable us to do three necessary things simultaneously: (a) better science overall, (b) address the debate on whether bilingualism affects the mind and brain and (c) in light of (b), reveal what the conditions are for bilingual effects. Bilingualism is not black or white. It cannot be reduced to 'yes' or 'no'. It is multicolored and multifaceted, a reality which empirical research needs to better confront.

References

Altinkamis, F., & Simon, E. (2020). Language abilities in bilingual children: The effect of family background and language exposure on the development of Turkish and Dutch. *International Journal of Bilingualism*, 1367006920909889.

Amengual, M. (2013). *An Experimental Approach to Phonetic Transfer in the Production and Perception of Early Catalan-Spanish Bilinguals (PhD Dissertation)*. The University of

Texas at Austin.
Amengual, M. (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*(2), 133–152.
Anderson, J. A., Saleemi, S., & Bialystok, E. (2017). Neuropsychological assessments of cognitive aging in monolingual and bilingual older adults. Journal of neurolinguistics, 43, 17-27.
Anderson, J. A., Mak, L., Chahi, A. K., & Bialystok, E. (2018). The language and social background questionnaire: Assessing degree of bilingualism in a diverse population. Behavior research methods, 50(1), 250-263.
Andreou, M., Knopp, E., Bongartz, C., & Tsimpli, I. M. (2015). Character reference in Greek German bilingual children's narratives. *EuroSLA Yearbook*, *15*(1), 1–40.
Ansaldo, A.I., Ghazi-Saidi, L., & Adrover-Roig, D. (2015). Interference control in elderly bilinguals: Appearances can be misleading. *Journal of Clinical and Experimental Neuropsychology, 37*, 455–470.
Azurmendi, M. J., Larrañaga, N., & Apalategi, J. (2008). Bilingualism, identity, and citizenship in the Basque Country. In Niño-Murcia, M. & Rothman, J. (eds.) *Bilingualism and Identity*, 35-62. Amsterdam. John Benjamins Publishing Company.
Barac, R., Moreno, S., & Bialystok, E. (2016). Behavioral and electrophysiological differences in executive control between monolingual and bilingual children. Child Development, 87(4), 1277-1290.
Bialystok, E. (2017). The bilingual adaptation: How minds accommodate experience. *Psychological Bulletin*, 143, 233.
Bialystok, E. (2016). The signal and the noise: Finding the pattern in human behavior. *Linguistic Approaches to Bilingualism* 6(5): 605-621.
Bialystok, E., Poarch, G., Luo, L., & Craik, F. I. (2014). Effects of bilingualism and aging on executive function and working memory. Psychology and aging, 29(3), 696.
Birdsong, D. (2014). Dominance and age in bilingualism. *Applied Linguistics*, *35*(4), 374-392.
Birdsong, D., Gertken, L., & Amengual, M. (2012). *Bilingual Language Profile: An Easy-to Use Instrument to Assess Bilingualism*. *COERLL*. University of Texas.
Bosch, S., Veríssimo, J., & Clahsen, H. (2019). Inflectional morphology in bilingual language processing: An age-of-acquisition study. *Language Acquisition*, *26*(3), 339–360. Cabrelli, J., Luque, A., & Finestrat-Martínez, I. (2019). Influence of L2 English phonotactics in L1 Brazilian Portuguese illusory vowel perception. *Journal of Phonetics*, *73*(55–69).
Casillas, J. (2015). Production and perception of the /i/-/I/ vowel contrast: The case of L2-dominant early learners of English. *Phonetica*, *72*(2), 182–205.
Dash, T., Berroir, P., Joanette, Y., & Ansaldo, A. I. (2019). Alerting, Orienting, and Executive Control: The Effect of Bilingualism and Age on the Subcomponents of Attention. *Frontiers in Neurology, 10*(October), 1–12. https://doi.org/10.3389/fneur.2019.01122
De Cat, C. (2018). Predicting language proficiency in bilingual children. *Studies in Second Language Acquisition*, 1–47.
DeLuca, V. F., Rothman, J., & Pliatsikas, C. (2018). Linguistic immersion and structural effects on the bilingual brain: a longitudinal study. *Bilingualism: Language and Cognition, 1*, 1–16. https://doi.org/10.1017/S1366728918000883
DeLuca, V., Rothman, J., Bialystok, E., & Pliatsikas, C. (2019). Redefining bilingualism as a spectrum of experiences that differentially affects brain structure and function. *Proceedings of the National Academy of Sciences*, *116*(15), 7565-7574.
DeLuca, V., Rothman, J., Bialystok, E., & Pliatsikas, C. (2020). Duration and extent of bilingual experience modulate neurocognitive outcomes. NeuroImage, 204, 116222.
Dimitropoulou, M., Duñabeitia, J. A., & Carreiras, M. (2011). Masked translation priming effects with low proficient bilinguals. *Memory & Cognition, 39*, 260–275.

Faroqi-Shah, Y., Sampspon, M., Pranger, M., & Baughman, S. (2018). Cognitive control, word retrieval and bilingual aphasia: Is there a relationship? *Journal of Neurolinguistics, 45*, 95–109.

Friesen, D. C., Luo, L., Luk, G., & Bialystok, E. (2015). Proficiency and control in verbal fluency performance across the lifespan for monolinguals and bilinguals. Language, cognition and neuroscience, 30(3), 238-250.

García, P., Leibold, L., Buss, E., Calandruccio, L., & Rodríguez, B. (2018). Code-switching in highly proficient Spanish/English bilingual adults: Impact on masked word recognition. *Journal of Speech, Language, and Hearing Research*, *61*(9), 2353.2363.

Gertken, L., Amengual, M., & Birdsong, D. (2014). Assessing Language Dominance with the Bilingual Language Profile. In A. Lecrecq (Ed.), *Measuring L2 Proficiency: Perspectives from SLA* (pp. 208–225). Bristol, UK: Multilingual Matters.

Gertken, L. M. (2013). *Priming of Relative Clause Attachment During Comprehension in French as a First and Second Language*. The University of Texas at Austin.

Gollan, T.H., Weissberger, G., Runnqvist, E., Montoya, R.I., & Cera, C.M. (2012). Self ratings of spoken language dominance: A multi-lingual naming test (MINT) and preliminary norms for young and aging Spanish-English bilinguals. *Bilingualism: Language and Cognition, 15*(3), 594–615.

González Alonso, J., Baquero Castellanos, S., & Müller, O. (2016). Masked constituent priming of English compounds in native and non-native speakers. *Language, Cognition and Neuroscience*, *31*(8), 1038–1054.

Goodenough FL (1926). Racial differences in the intelligence of school children. *Journal of Experimental Psychology*. 9:388–397.

Green, D. W., & Abutalebi, J. (2013). Language control in bilinguals: The adaptive control hypothesis. *Journal of Cognitive Psychology*, *25*(5), 515–530. https://doi.org/10.1080/20445911.2013.796377

Grégoire, L., & Greening, S. (2020). Fear of the known: semantic generalisation of fear conditioning across languages in bilinguals. *Cognition and Emotion*, *32*.

Grundy, J. G. (2020). The effects of bilingualism on executive functions: An updated quantitative analysis. *Journal of Cultural Cognitive Science*, 4, 177–199 (2020). https://doi.org/10.1007/s41809-020-00062-5

Gullifer, J. W., Chai, X. J., Whitford, V., Pivneva, I., Baum, S., Klein, D., & Titone, D. (2018). Bilingual Experience and Resting-State Brain Connectivity: Impacts of L2 Age of Acquisition and Social Diversity of Language Use on Control Networks. *Neuropsychologia, 117*(May), 123–134. https://doi.org/10.1016/j.neuropsychologia.2018.04.037

Gullifer, J. W., & Titone, D. (2018). *Compute language entropy with {LanguageEntropy}*. https://doi.org/10.5281/zenodo.1406201

Gullifer, J. W., & Titone, D. (2020). Characterizing the social diversity of bilingualism using language entropy. *Bilingualism: Language and Cognition*, *23*(2), 283–294. https://doi.org/10.1017/S1366728919000026

Garraffa, M., Beveridge, M., & Sorace, A. (2015). Linguistic and cognitive skills in Sardinian–Italian bilingual children. *Frontiers in Psychology*, *6*, 1898. Hartanto, A., & Yang, H. (2016). Disparate bilingual experiences modulate task-switching advantages: A diffusion-model analysis of the effects of interactional context on switch costs. *Cognition*, *150*, 10–19. https://doi.org/10.1016/j.cognition.2016.01.01 Hendricks, A. E., Miller, K., & Jackson, C. N. (2018). Regularizing unpredictable variation: Evidence from a natural language setting. *Language Learning and Development*, *14*(1), 42-60. https://doi.org/10.1080/15475441.2017.1340842

Hilchey, M. D., & Klein, R. M. (2011). Are there bilingual advantages on nonlinguistic interference tasks? Implications for the plasticity of executive control processes. *Psychonomic Bulletin and Review, 18*(4), 625–658. https://doi.org/10.3758/s13423-011-0116-7

Hoffmann, C. (2000). Bilingual and trilingual competence: Problems of description and differentiation. Estudios de Sociolinguistica, 1, 83-92

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.

Iverson, M., & Miller, D. (2018). Language attrition and maintenance: Two sides of the same coin? *Linguistic Approaches to Bilingualism*, *7*(6), 704–708. https://doi.org/10.1075/lab.00009.ive

Jäger, G. (2007). Maximum entropy models and stochastic Optimality Theory. In A. Zaenen (Ed.), *Architectures, rules, and preferences: A Festschrift for Joan Bresnan* (pp. 467– 479). Retrieved from https://www.press.uchicago.edu/ucp/books/book/distributed/A/bo5926545.html

Kan, R. T. (2019). Production of Cantonese classifiers in young heritage speakers and majority language speakers. *International Journal of Bilingualism*, *23*(6), 1531–1548.

Kaushanskaya, M., Blumenfeld, H.K., & Marian, V. (2019). The Language Experience and Proficiency Questionnaire (LEAP-Q): Ten years later. *Bilingualism: Language and Cognition*. Epub ahead of print, 2019. DOI: 10.1017/S1366728919000038. Keating, G., VanPatten, B., & Jegerski, J. (2011). Who Was Walking on the Beach?: Anaphora Resolution in Spanish Heritage Speakers and Adult Second Language Learners. *Studies in Second Language Acquisition, 33*(2), 193–221.

Krizman, J., Skoe, E., Marian, V., & Kraus, N. (2014). Bilingualism increases neural response consistency and attentional control: Evidence for sensory and cognitive coupling. *Brain and Language, 128*, 34–40.

Kubota, M., Chevalier, N., & Sorace, A. (2020). How bilingual experience and executive control influence development in language control among bilingual children. *Developmental science*, *23*(1), e12865. https://doi.org/10.1111/desc.12865 Lehtonen, M., Soveri, A., Laine, A., Järvenpää, J., de Bruin, A., & Antfolk, J. (2018). Is bilingualism associated with enhanced executive functioning in adults? A meta-analytic review. *Psychological Bulletin, 144*(4), 394–425. https://doi.org/10.1037/bul0000142 Li, P., Sepanski, S., & Zhao, X. (2006). Language history questionnaire: A web-based interface for bilingual research. Behavior Research Methods, 38(2), 202–210. doi:10.3758/BF03192770

Li, P., Legault, J., & Litcofsky, K. A. (2014). Neuroplasticity as a function of second language learning: Anatomical changes in the human brain. *Cortex, 58*, 301–324. https://doi.org/10.1016/j.cortex.2014.05.001

Li, P., Zhang, F., Tsai, E., & Puls, B. (2014). Language history question- naire (LHQ 2.0): A new dynamic web-based research tool. Bilingualism: Language and Cognition, 17(3), 673–680. doi:10. 1017/S1366728913000606

Li, P., Zhang, F., Yu, A., & Zhao, X. (2019). Language History Questionnaire (LHQ3): An enhanced tool for assessing multilingual experience. *Bilingualism: Language and Cognition*, 1–7. https://doi.org/10.1017/S1366728918001153

Libben, M. R., & Titone, D. A. (2009). Bilingual lexical access in context: evidence from eye movements during reading. *Journal of Experimental Psychology: Learning, memory, and cognition, 35*(2), 381–390.

Loureiro-Rodríguez, V. (2008). Conflicting values at a conflicting age: Linguistic ideologies in Galician adolescents. In Niño-Murcia, M. and Rothman, J. (eds.) *Bilingualism and identity: Spanish at the crossroads with other languages*, (pp. 63-86). Amsterdam, John Benjamins Publishing Company.

Luk, G., & Bialystok, E. (2013). Bilingualism is not a categorical variable: Interaction between language proficiency and usage. Journal of Cognitive Psychology, 25(5), 605–621. doi:10.1080/20445911.2013.795574

Luo, D., Kwok, V. P. Y., Liu, Q., Li, W., Yang, Y., Zhou, K., … Tan, L. H. (2019). Microstructural plasticity in the bilingual brain. *Brain and Language, 196*, 104654. https://doi.org/10.1016/j.bandl.2019.104654

Maguire, E. A., Frackowiak, R. S., & Frith, C. D. (1997). Recalling routes around London: activation of the right hippocampus in taxi drivers. *Journal of Neuroscience, 17*(18), 7103-7110.

Marian, V. (2019). 'Measuring Bilingualism With Self-Reports and Standardized Tests: Is a "Bilingualism Quotient" Realistic?' Keynote address at *Capturing and Quantifying Individual Differences in Bilingualism*. UiT The Arctic University of Norway – Tromsø (Norway), September 3.

Marian, V., Blumenfeld, H.K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research, 50*, 940–967.

Meisel, J.M. (2011). Bilingual language acquisition and theories of diachronic change: Bilingualism as cause and effect of grammatical change. *Bilingualism: Language and Cognition, 14*(2), 121-145.

Miller, D., & Rothman, J. (2019). You win some, you lose some: Comprehension and event related potential evidence for L1 attrition. *Bilingualism: Language and Cognition*, 1–15. https://doi.org/10.1017/S1366728919000737

Mitits, L., Alexiou, T., & Milton, J. (2018). Does the language you speak at home affect the size of your L2 vocabulary?. The Language Learning Journal, 46(5), 569-582. Mitrofanova, N., Rodina, Y., Urek, O., & Westergaard, M. (2018). Bilinguals' sensitivity to grammatical gender cues in Russian: The role of cumulative input, proficiency, and dominance. *Frontiers in Psychology, 9*, 1894. https://doi.org/10.3389/fpsyg.2018.01894

Nichols, E. S., & Joanisse, M. F. (2016). Functional activity and white matter microstructure reveal the independent effects of age of acquisition and proficiency on second-language learning. *NeuroImage, 143*, 15–25. https://doi.org/10.1016/j.neuroimage.2016.08.053 Paap, K. R., Johnson, H. A., & Sawi, O. M. (2015). Bilingual advantages in executive functioning either do not exist or are restricted to very specific and undetermined circumstances. *Cortex, 69*, 265–278. https://doi.org/10.1016/j.cortex.2015.04.014 Peal E, Lambert W (1962). The relation of bilingualism to intelligence. *Psychological Monographs*. 76:1–23

Peets, K. F., Yim, O., & Bialystok, E. (2019). Language proficiency, reading comprehension and home literacy in bilingual children: The impact of context. *International Journal of Bilingual Education and Bilingualism*, 1-15.

Persici, V., Vihman, M., Burro, R., & Majorano, M. (2019). Lexical access and competition in bilingual children: The role of proficiency and the lexical similarity of the two languages. *Journal of Experimental Child Psychology, 179*, 103–125.

Puig-Mayenco, E., & Rothman, J. (2019). Low Proficiency Does Not Mean Ab Initio: A Methodological Footnote for Linguistic Transfer Studies. *Language Acquisition*. Robinson, M. G., & Sorace, A. (2019). The influence of collaborative language learning on cognitive control in unbalanced multilingual migrant children. *European Journal of Psychology of Education, 34*(1), 255–272.

Reichle, R. V., & Birdsong, D. (2014). Processing focus structure in L1 and L2 French: L2 proficiency effects on ERPs. *Studies in Second Language Acquisition, 36*(3), 535–564.

Rodina, Y., & Westergaard, M. (2017). Grammatical gender in bilingual Norwegian–Russian acquisition: The role of input and transparency. *Bilingualism: Language and Cognition, 20*(1), 197-214. https://doi.org/10.1177/1367006916648859

Saer DJ (1923). The effects of bilingualism on intelligence. *British Journal of Psychology*;14:25–38.

Serratrice, L. (2013). Cross-linguistic influence in bilingual development: Determinants and mechanisms. *Linguistic Approaches to Bilingualism, 3*(1), 3-25.

Serratrice, L., & De Cat, C. (2020). Individual differences in the production of referential expressions: The effect of language proficiency, language exposure and executive function in bilingual and monolingual children. *Bilingualism: Language and Cognition*, 23(2), 371–386.

Shi, L-F. (2011). How “proficient” is proficient? Subjective proficiency as a predictor of bilingual listeners’ recognition of English words. *American Journal of Audiology, 20*, 19–32.

Simonet, M. (2014). Phonetic consequences of dynamic cross-linguistic interference in proficient bilinguals. *Journal of Phonetics*, *43*(1), 26–37.

Simonet, M., Casillas, J., & Díaz, Y. (2014). The effects of stress/Accent on VOT depend on language (English, Spanish), consonant (/d/, /t/) and linguistic experience (monolinguals, bilinguals). In *Speech Prosody 7: Proceedings of the 7th International Conference on Speech Prosody: Social and Linguistic Speech Prosody* (pp. 2333–3042). Trinity College.

Shannon, C. E. (1948). A Mathematical Theory of Communication. *The Bell System Technical Journal*, *27*(3), 371–423. https://doi.org/10.1016/S0016-0032(23)90506-5 Smith, S. A., Briggs, J. G., Pothier, H., & Garcia Jr, J. N. (2019). ‘Mental Workouts for Couch Potatoes’: Executive Function Variation among Spanish–English Bilingual Young Adults. Applied Linguistics, 40(3), 413-431.

Solís-Barroso, C., & Stefanich, S. (2019). Measuring Language Dominance in Early Spanish/English Bilinguals. *Languages*, *4*(62).

Stocco, A., & Prat, C.S. (2014). Bilingualism trains specific brain circuits involved in flexible rule selection and application. *Brain and Language, 137*, 50–61.

Stoehr, A., Benders, T., Van Hell, J. G., & Fikkert, P. (2018). Heritage language exposure impacts voice onset time of Dutch–German simultaneous bilingual preschoolers. *Bilingualism: Language and Cognition*, *21*(3), 598–617.

Stoehr, A., Benders, T., Van Hell, J. G., & Fikkert, P. (2019). Bilingual preschoolers’ speech is associated with non-native maternal language input. *Language Learning and Development*, *15*(1), 75–100.

Sullivan, M. D., Prescott, Y., Goldberg, D., & Bialystok, E. (2017). . Executive control processes in verbal and nonverbal working memory. Growing old with two languages: Effects of bilingualism on cognitive aging, 53, 161.

Sulpizio, S., Del Maschio, N., Del Mauro, G., Fedeli, D., & Abutalebi, J. (2019). Bilingualism as a gradient measure modulates functional connectivity of language and control networks. *NeuroImage*, *205*, 116306. https://doi.org/10.1016/j.neuroimage.2019.116306

Sun, H., Steinkrauss, R., Tendeiro, J., & De Bot, K. (2016). Individual differences in very young children’s English acquisition in China: Internal and external factors. *Bilingualism: Language and Cognition*, *19*(3), 550–566.

Unsworth, S. (2013). Assessing the role of current and cumulative exposure in simultaneous bilingual acquisition: The case of Dutch gender. *Bilingualism: Language and Cognition*, *16*(1), 86–110.

Unsworth, S., Brouwer, S., de Bree, E., & Verhagen, J. (2019). Predicting bilingual preschoolers’ patterns of language development: Degree of non-native input matters. *Applied Psycholinguistics*, *40*(5), 1189–1219.

Unsworth, S., Chondrogianni, V., & Skarabela, B. (2018). Experiential measures can be used as a proxy for language dominance in bilingual language acquisition research. *Frontiers in Psychology*, *9*.

Urla, J. (2012). *Reclaiming Basque: Language, nation, and cultural activism.* Reno. University of Nevada Press.

van den Noort, M., Struys, E., Bosch, P., Jaswetz, L., Perriard, B., Yeo, S., … Lim, S. (2019). Does the Bilingual Advantage in Cognitive Control Exist and If So, What Are Its Modulating Factors? A Systematic Review. *Behavioral Sciences*, *9*(3), 27. https://doi.org/10.3390/bs9030027

Vender, M., Garraffa, M., Sorace, A., & Guasti, M. T. (2016). How early L2 children perform on Italian clinical markers of SLI: A study of clitic production and nonword repetition. *Clinical Linguistics & Phonetics*, *30*(2), 150–169.

Vender, M., Hu, S., Mantione, F., Delfitto, D., & Melloni, C. (2018). The production of clitic pronouns: A study on bilingual and monolingual dyslexic children. *Frontiers in Psychology*, *9*, 2301.

Verhagen, J., de Bree, E., & Unsworth, S. (2020). Effects of Bilingual Language Use and Language Proficiency on 24-month-olds' Cognitive Control. *Journal of Cognition and Development*, *21*(1), 46–71.

Veríssimo, J., Heyer, V., Jacob, G., & Clahsen, H. (2018). Selective effects of age of acquisition on morphological priming: Evidence for a sensitive period. *Language Acquisition*, *25*(3), 315–326.

Wang, Y., Kuhl, P. K., Chen, C., & Dong, Q. (2009). Sustained and transient language control in the bilingual brain. *NeuroImage*, *47*(1), 414–422.

Wolleb, A., Sorace, A., & Westergaard, M. (2018). Exploring the role of cognitive control in syntactic processing: Evidence from cross-language priming in bilingual children. *Linguistic Approaches to Bilingualism*, *8*(5), 606–636.

Yamasaki, B. L., Stocco, A., & Prat, C. S. (2018). Relating individual differences in bilingual language experiences to executive attention. *Language, Cognition and Neuroscience*, *33(*9), 1128–1151. https://doi.org/10.1080/23273798.2018.1448092

Zhang, H., Wu, Y. J., & Thierry, G. (2020). Bilingualism and aging: A focused neuroscientific review. *Journal of Neurolinguistics, 54*, 100890.

Zyzik, E., & Sanchez, R. (2019). Beyond accuracy: Heritage speakers' performance on two kinds of acceptability judgement tasks. *Applied Psycholinguistics*, *40*(3), 645–671.