Brain & Language 222 (2021) 105014

Contents lists available at ScienceDirect

Brain and Language

journal homepage: www.elsevier.com/locate/b&l

Research on bilingualism as discovery science

Christian A. Navarro-Torres [a,*], Anne L. Beatty-Martínez [b,*], Judith F. Kroll [c], David W. Green [d]

[a] Department of Language Science, University of California, Irvine, United States
[b] Department of Psychology, McGill University, Canada
[c] School of Education, University of California, Irvine, United States
[d] Department of Experimental Psychology, University College London, United Kingdom

ARTICLE INFO

Keywords:
Statistical power
Replication crisis
Variation
Bilingualism
Bilingual phenotypes

ABSTRACT

An important aim of research on bilingualism is to understand how the brain adapts to the demands of using more than one language. In this paper, we argue that pursuing such an aim entails valuing our research as a discovery process that acts on variety. Prescriptions about sample size and methodology, rightly aimed at establishing a sound basis for generalization, should be understood as being in the service of science as a discovery process. We propose and illustrate by drawing from previous and contemporary examples within brain and cognitive sciences, that this necessitates exploring the neural bases of bilingual phenotypes: the adaptive variety induced through the interplay of biology and culture. We identify the conceptual and methodological prerequisites for such exploration and briefly allude to the publication practices that afford it as a community practice and to the risk of allowing methodological prescriptions, rather than discovery, to dominate the research endeavor.

> *"We have to remember that what we observe is not nature in itself, but nature exposed to our method of questioning."* Werner Heisenberg (1958)

1. Introduction

Research on bilingualism generates debate on the neural bases of language that address fundamental questions about language learning (e.g., the role of critical periods), the specificity of language networks (e.g., the nature of any modularity) and their control (e.g., the domain-generality of such control). More recently, specific aspects of the field, namely the putative cognitive and neural consequences (often framed in the form of advantages) of bilingualism, have become a hotspot for controversy tied to the replication crisis in psychology. The critique of this research appears to be broad, addressing issues of power and sample size (e.g., Brysbaert, 2020; Nichols, Wild, Stojanoski, Battista, & Owen, 2020), failures to replicate (e.g., Paap & Greenberg, 2013), noise in samples and methods (e.g., García-Pentón, Fernández García, Costello, Duñabeitia, & Carreiras, 2016a, 2016b; Valian, 2015), and publication bias (e.g., de Bruin, Treccani, & Della Sala, 2015a; but see Bialystok, Kroll, Green, MacWhinney, & Craik, 2015), suggesting that the effects of bilingualism on cognitive and brain functioning are the result of questionable research practices. Consequently, several prescribed remedies, such as large samples (Brysbaert, 2020) and uniform[1] experimental procedures (García-Pentón et al., 2016a, 2016b), have been marketed as solutions (see also Szucs & Ioannidis, 2020 for an example involving neuroscience more generally). However, such critiques and remedies, though well intended, often fail to place discussions in the broader context of science and its function throughout history. This raises the question of how the implementation of compulsory prescriptions would come to affect research on bilingualism more generally.

In this paper, we argue that the remedies and prescriptions put forward are deceptively simple and place us on a misleading path as they are based on a mischaracterization of the fundamentals of the scientific endeavor. While this paper is geared toward discussing current issues in research on bilingualism, we necessarily draw from the history of science to make the argument self-evident. Our position is that both large

* Corresponding authors at: Department of Language Science, 3151 Social Science Plaza A, University of California, Irvine, Irvine, CA 92697, United States (C.A. Navarro-Torres). Department of Psychology, McGill University, 2001 McGill College Ave., Montréal, QC H3A 1G1, Canada (A.L. Beatty-Martínez).

E-mail addresses: canavar5@uci.edu (C.A. Navarro-Torres), anne.beatty-martinez@mail.mcgill.ca (A.L. Beatty-Martínez).

[1] We use "uniform" and "uniformity" to describe the hypothetical state in which scientific practice would require the application of a single idealized methodology and/or method to assess replicability of an effect.

https://doi.org/10.1016/j.bandl.2021.105014
Received 19 January 2021; Received in revised form 24 August 2021; Accepted 25 August 2021
0093-934X/Published by Elsevier Inc.

samples and conventionalized methods are important, but their role needs to be understood in the context of science as a discovery process, in which research findings are generated through interrelated iterations of exploration and falsification, which in turn lead to new insights and allow for the formulation of new questions. Fundamental to this process is the generation of *variety*[2] that permits incremental advance. The generation of variety serves two purposes: to identify reliable signals in the noise of our observations and to allow the formulation of effective theories and constructs about our world. Hypotheses, for instance, that the shape of the head is correlated with psychological traits (Simpson, 2005) or that bilingualism negatively impacts intelligence (Peal & Lambert, 1962), are discarded along the way. Constraints on the exploration of variety, such as those imposed by prescriptive remedies (e.g., keeping experimental designs as simple as possible), hinder the discovery process and so it is imperative in our view to ensure that methodological injunctions and publication practices are understood within the context of science as a discovery process.

The remainder of the paper is organized as follows: in the next section (*Why a prescriptive science is problematic*), we provide a critique on both practical and conceptual grounds of the rationale for power and uniformity prescriptions. Our line of argument then leads us to consider the implications for research practice in bilingualism (*Articulating the research enterprise of bilingualism*), where we emphasize the value of rich characterization of the sample, practices that enable the assessment of interactions rather than main effects on their own, and the application of sensitive tools. The implication is that without appropriate characterization, and without research practices and tools that lead to effective signal extraction, replication and large samples may be void of scientific interest. In both sections, we illustrate the manifestation of science as a discovery process with a range of past and contemporary examples drawn from research on bilingualism as well as from other fields. We necessarily draw on a range of examples, including those outside bilingualism, because these points are not unique to research on bilingualism; rather, they reflect a healthy and productive scientific enterprise. We do not argue against the importance of replication, the analytic value of Big Data, nor the application of sensitive and conventionalized research tools. Rather, we suggest that the application of method should be grounded in science as a discovery process.

2. Why a prescriptive science is problematic

We proposed above that prescriptions to remedy poor research practice fail to adequately acknowledge science as a discovery process. Curiously, in applying these prescriptions to research on bilingualism, the analogy invoked is bringing an image into focus: just as glasses improve blurry vision, larger samples have been claimed to increase the resolution of data (Brysbaert, 2020). Similarly, methodological uniformity transforms haziness into a well-defined picture (García-Pentón et al., 2016a, 2016b). Such analogies, rhetorically persuasive perhaps, are misleading. There is a sensible motivation to establish the stability of effect sizes for a given class of data (see Lorca-Puls et al., 2018 for actual rather than simulated data in the context of the relation between brain damage and speech articulation), but it is our ability to identify a reliable signal that is key, not sample size or uniformity *per se*. Below we comment on four points to illustrate why power and uniformity prescriptions are insufficient for effective signal extraction.

2.1. Ambiguity is independent of power

Studies of individuals who speak two or more languages have demonstrated a range of consequences for cognition (see Bialystok, 2017 for a review) but controversy surrounds some of these effects. Many large sample studies have yielded null results (Antón et al., 2014; Dick et al., 2019; Duñabeitia et al., 2014; Kałamała, Szewczyk, Chuderski, Senderecka, & Wodniecka, 2020; Nichols et al., 2020; Paap, Anders-Jefferson, Mason, Alvarado, & Zimiga, 2018) and other meta-analyses report inconsistency (Anderson, Hawrylewicz, & Grundy, 2020; Donnelly, Brooks, & Homer, 2019; Grundy & Timmer, 2017; Lehtonen et al., 2018; Mukadam, Sommerlad, & Livingston, 2017; Schroeder, 2018; Sulpizio, Del Maschio, Fedeli, & Abutalebi, 2020). On the face of it, such reports have called putative bilingualism effects into question. However, if statistical power were indeed the solution to ambiguity, then we would expect greater consistency across studies with large samples. Problematically, from a naive prescriptive approach, other large sample studies do report effects of bilingualism (Bak, Nissan, Allerhand, & Deary, 2014; Hartanto, Toh, & Yang, 2018; Santillán & Khurana, 2017). Are these latter studies like “black swans” reducing our belief in the generalization that “all swans are white”?

We need not take the current impasse at face value. Consider a contemporary example. The COVID-19 pandemic made it urgent for scientific communities to address a critical question: does the human body develop long-term immunity to the virus? While some large-scale studies suggest that it does (e.g., Iyer et al., 2020; Wu et al., 2020), other large-scale studies show that the effects are limited (e.g., Liu, Wu, Tao, Zeng, & Zhou, 2020; Pollán et al., 2020). Curiously, it is the collection of single-case patients with reinfection (e.g., Tillett et al., 2020) that initially became more decisive in addressing this question. The point here is that, without an understanding of the boundary conditions of an effect, power, in the form of large sample studies, does not, on its own, improve our ability to extract a signal. In fact, as reinfection cases suggest, and as we further illustrate below, small-n studies that exploit the features of the sample can be more informative than studies with poorly characterized large samples, which are bound to increase noise in our signals[3].

2.2. Discovery and the power of the small

The history of science suggests that we should recognize the value of the small sample to increase signal.[4] Research on bilingualism, as well as cognitive and neuroscience research more generally, also attests to this point. Consider a fundamental question: How does the brain adapt to input deprivation? One claim is that brain specialization is determined by input senses. Vetter et al. (2020) examined brain activity in the primary visual cortex in healthy blindfolded (n = 10) and congenitally blind (n = 5) individuals while listening to natural sounds. Using multivariate pattern analysis, they found that the blindfolded participants activated the primary visual cortex in response to the sounds

[2] We use “variety” to reflect what allows science to act as a discovery process (i.e., diversification in the application of ideas, methods, and scientific practices), as opposed to “variation”, which refers to a number of different referents in the world such as the interactional contexts of language use, within-language variation, typological similarities and differences between languages, individual differences, and an individual’s response to encountered variation. We therefore use a more specific term for clarity’s sake when occasion demands.

[3] As a related point, we note that limitation on inference from poorly-characterized data is not overcome by *meta*-analyses of studies using such data.

[4] We briefly illustrate the case of the discovery of penicillin: Alexander Fleming had searched for antimicrobial agents for years before recognizing the chance finding in a petri dish that led him to examine its anti-microbial properties on mice. But development requires a community of practitioners: it was a decade later before the drug was purified by Florey, Chain, and Heatley in Oxford – their work made urgent by war. Furthermore, the significance of a finding is a community-agreement. The first patient treated (see Barrett, 2018) was Constable Albert Alexander, who had developed sepsis. His immediate recovery was remarkable, but the original penicillin formulation was not optimal, and he died as it was excreted too rapidly. Despite the shortcomings, the constable’s remarkable temporary reprieve was sufficient to convince the team (a community of researchers) that a cure would have been possible if only sufficient drug could have been made.

despite not having access to visual input in the moment. Remarkably, the same pattern of activation was observed in the congenitally blind group, suggesting that it is not sensorial input *per se*, but rather, the tasks performed by a brain region, that shape brain specialization. It is not sample size, but signal quality that is key here.

Neuropsychological data, typically based on a small number of cases, have been instrumental for our understanding of memory systems[5], but have also been critical to the emergence of research on language control in bilinguals, as they establish the face validity of the distinction between language networks and their control (Green & Kroll, 2019). For example, S.J., a Friulian-Italian speaker (Fabbro, Skrap, & Aglioti, 2000), had intact clausal processing for speech comprehension and speech production in both languages, combined with an inability to avoid switching inappropriately in a conversation (e.g., into Friulian when speaking to an Italian-only speaker). More complex control problems reveal dissociations between speech production in one language and translation into it, as exemplified in the alternate antagonism and paradoxical translation of two bilingual patients (Paradis, Goldblum, & Abidi, 1982), modelled narratively in Green (1986) and neurocomputationally in Noor, Friston, Ekert, Price, and Green (2020). Such cases pave the way for neuroimaging research on the nature of recovery in bilingual aphasia in which we can ask, for example, whether recovery depends on perilesional activation or the use of a previously inhibited alternative network.

A final example exploits the presence of bilingualism in two different modalities –speech and sign language. Hearing bimodal bilinguals are a small population of speakers. They are typically either children of deaf adults or sign interpreters. Bimodal bilinguals are able to do something that is impossible to do with two spoken languages, namely speaking one language while simultaneously signing another (i.e., code-blending). Because of this feature, bimodal bilingualism provides a unique opportunity to test claims about how the bilingual's languages are controlled. Initial naturalistic production data from two children (Petitto et al., 2001) and 11 adults (Emmorey, Borinstein, Thompson, & Gollan, 2008) showed that bimodal bilinguals strongly prefer code-blending over switching between sign and speech, suggesting that combining the two languages is relatively free of control demands. More recent work using magnetoencephalography (Blanco-Elorrieta, Emmorey, & Pylkkänen, 2018; see also Emmorey, Li, Petrich, & Gollan, 2020 for converging behavioral evidence) confirmed this finding, but also showed that increased cognitive effort is required when bimodal bilinguals switch out of a code-blend to either language alone, suggesting that it is the disengagement of one language to switch into the other that requires active control.

The point is that, so long the data are adequately characterized (see 'A rich characterization of the sample and an identification of boundary conditions' subsection) and measures are sensitive (see 'Realizing signal extraction' subsection), a small sample, even a single case, that exploits the special properties of a particular population allow for effective signal extraction that can generate new observations and move the field forward. This is not to say that large-sample studies cannot be equally informative, or that these findings should not be replicated, but that the force of the evidence is not based on statistical power alone.

2.3. Discovery acts on variety

Science acts as a discovery process with Darwinian-like properties, except it also possesses a time-binding property in which earlier ideas and methods can be recruited at a later point in time.[6] Just as natural selection depends on biological diversification to ensure the continuation of evolution, science relies on variety to ensure incremental improvements in our signal-extraction abilities. Progress, in the form of new discoveries and insights, is made by the gradual accumulation of patterns that emerge over distinct data and methods, a process that William Whewell referred to as *consilience* (Laudan, 1971).

In some cases, consilience is relatively straightforward: Converging evidence is obtained from variations of a method (see Green & Abutalebi, 2015 for an example of left caudate involvement in language control). For instance, in examining the question of whether bilingualism changes the engagement of control processes, Wu and Thierry (2013) found in a group of Welsh-English bilinguals a modulation of Flanker performance by experimentally inducing a shift in the language context. Using a novel paradigm in which Flanker was interleaved with words from Welsh or English (single-language context) or both languages (dual-language context), they showed that exposure to words in a dual-language context led to greater electrophysiological efficiency in Flanker performance. Since then, several studies have also reported electrophysiological Flanker modulations using variations of the paradigm in different bilingual populations (Bosma & Pablos, 2020; Jiao, Grundy, Liu, & Chen, 2020; Jiao, Liu, de Bruin, & Chen, 2020), suggesting that the effect reflects a more general feature of bilingualism. This is an important discovery not only because it shows how control processes adapt to the language context, but also because it makes a more general point that the relative involvement of control processes on a particular task will depend on the control state of an individual at a particular time (see Hsu, Kuchinsky, & Novick, 2020; Salig, Valdés Kroff, Robert, & Novick, 2021 for an elaboration of this argument).

In other cases, consilience requires us to bring together evidence from different methods and populations. Consider the claim that language processing is determined by factors unique to the language system. Converging evidence from neuroscience suggests that domain-general processes also play a vital role. For instance, in studies examining monolingual brain activity across a variety of linguistic and non-linguistic conflict-related tasks (Hsu, Jaeggi, & Novick, 2017), co-localization and functional connectivity analyses reveal that, although activation of the Multiple Demands system varies across tasks, engagement of the left inferior frontal gyrus is constant across tasks while also co-activating with other task-specific networks. Research on bilingualism, too, attests to this idea (KK Nair, Rayner, Siyambalapitiya, & Biedermann, 2021). In proficient bilinguals, brain potentials reveal that the ability to recover from prediction errors during L2 sentence reading is mediated by individual differences in control ability, but this effect depends on L1 verbal fluency (Zirnstein, van Hell, & Kroll, 2018). The interaction between control and fluency suggests that successful L2 prediction may depend on language-related processes that are partially overlapping with more domain-general control processes. It is the co-ordination, not the presence or absence, of particular processes or brain regions that is relevant (see also Bialystok, 2011; Morales, Gómez-Ariza, & Bajo, 2013).

[5] For instance, once there was a theory that entry to long-term memory required an intact short-term memory. The theory was rendered less tenable by an n = 1 –a patient with a severely damaged short-term memory but an intact long-term memory (Shallice & Warrington, 1970). Conversely, the discovery of patients (n = 6) with damage to long-term memory but relatively intact short-term memory (e.g., Baddeley & Warrington, 1970) undermined proposals that short-term memory is the activation of representations in long-term memory.

[6] Consider the case of Hockett (1985), who hypothesized that hunter-gatherer societies showed a marked lack of labiodentals (e.g., /f/ and /v/) because these incurred greater articulatory effort with their diet-induced edge-to-edge bites. The hypothesis, deemed a *just-so story*, was widely refuted at the time (see Brace, 1986 for a commentary on the matter) based on apparent inconsistencies between the decline of the edge-to-edge bite and the development of agricultural and food processing technologies. Using converging methods from paleoanthropology, linguistics, and evolutionary biology, Blasi et al. (2019) revisited Hockett's conjecture almost three decades later and provided evidence for how changes in fundamental aspects of the ecology (dietary and behavioral practices concerning what food we eat and how we process it) enriched human sound systems by enabling the innovation of a new class of speech sounds.

What makes these ideas (i.e., that control processes are state dependent, or that language draws from both domain-general and language-specific resources) compelling is not the ability to replicate a finding using the same method ad infinitum. Rather, it is the fact that we can identify converging patterns *despite* the use of different tools, procedures, and populations with varying sample sizes, all of which might seem to work against us but may in fact improve identification (i.e., signal quality). Such insights allow us to ask new and more useful questions. That is why deep insights about language and the brain emerge through the application of variety.

2.4. Focus is meaningless without context

Analogies invoking focus of an image via large samples and uniformity minimize the conceptual basis on which an observation is made and are fundamentally misleading for a simple reason. We only know the significance of increased focus because we already know the picture (i.e., the conceptual ground). Experience generates the conceptual ground for our everyday lives: we learn to recognize different objects and entities through our ability to interact with them over time. But the conceptual ground for the processes and causal mechanisms underlying the brain and behavior are typically unknown. Science as a discovery process fundamentally concerns the identification of effective theories and constructs of those unknowns. Such theories are based on our justified true beliefs given the evidence and are an intersubjective agreement about that evidence. Theories contest for that agreement. They are necessarily an intersubjective agreement because our senses and scientific tools do not provide immediate access to the physical world[7], as the introductory quote by German physicist Werner Heisenberg suggests. Thus, we come to know the significance of increased focus not by power or uniformity, but by generating and exploring the conceptual ground, as Fig. 1 gently illustrates. The problem lies not in the pursuit of statistical power or conventionalized tools *per se*, but in assuming that increasing power or achieving uniformity will generate a picture. As the Blind Men and an Elephant parable suggests, a thousand blind people inspecting separate parts of an elephant will yield an enormous effect size for their one bit but yield zilch for totality!

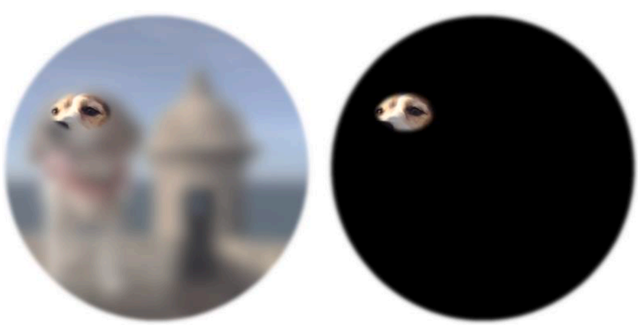

Fig. 1. a (left) and b (right). We formulate our hypotheses in the context of discovery – a possible picture of the world (a). Bringing an event into focus without context (i.e., testing a hypothesis without a conceptual ground) is meaningless (b).

3. Articulating the research enterprise of bilingualism

The key question is how we generate the conceptual ground for effective theories and constructs on research on bilingualism and its consequences. We do so by recognizing that language and the brain are byproducts of evolutionary and ecological processes. Such recognition is a generator of the expertise and intuitions for researchers and can play an important role in recognizing the significance of a chance observation or novel finding just as experience furnishes the hunches of everyday life (Bowers, Regehr, Balthazard, & Parker, 1990). We propose two key factors. First, pre-existing evolutionary older systems are coordinated and the use of these systems exapts[8] mechanisms for the control of language and action more generally (Stout & Chaminade, 2012). Bilingual speakers necessarily must select and control the language of use (e.g., Green, 1986, 1998; see Pliatsikas & Luk, 2016 for a review of data on the overlap). Second, and critical for research on bilingualism, there is a need to characterize the socio-cultural niche in which speakers act. To this end, we emphasize the need for research practices and tools that provide a rich characterization of the participant sample, and ultimately envision a research enterprise focused on the identification of bilingual phenotypes. Under this notion, interactions become of high relevance, and questions eliciting binary outcomes (e.g., is there a bilingual advantage?) become inadequate as they mask the richness of the science. Below we consider each of these points more carefully.

3.1. A rich characterization of the sample and an identification of boundary conditions

The interactional contexts of language use constrain which languages can be used and how they can be used. Characterizing speakers in terms of their habitual community practices, as well as their trajectory through particular contexts, is essential if we are to understand adaptive change[9]. We refer to such characterization as bilingual phenotyping (Adamou, 2010; Beatty-Martínez & Dussias, 2017; Beatty-Martínez et al., 2020; Poplack, 1987). Such phenotyping helps determine the boundary conditions for any adaptive effect because we might predict an

[7] How does the brain reconstruct the physical world? Salzman, Murasugi, Britten, and Newsome (1992) showed that microstimulation of neurons in the middle temporal area selectively distorts motion perception in monkeys. For instance, when applying microstimulation to neurons that selectively respond to objects with upward motion, the direction of motion reported by the monkeys in a direction discrimination task is upwards even when the physical stimuli are projecting downward motion, suggesting that perception is fundamentally abstracting, rather than merely reproducing, the physical world. It is in this sense that our observations are fundamentally an interpretative act (Barrett, 2017).

[8] Critical to evolutionary biology is the distinction between "adaptation", features that are the byproduct of natural selection, and "exaptation", features that attain a new function for their present role regardless of their evolutionary history (Gould & Vrba, 1982). See footnote 6 for an illustration.

[9] Adaptive response must be understood in the context of the variation inherent to evolutionary processes. Biological systems are functionally degenerate: they develop different structural configurations to perform an equivalent function (Deacon, 2010; Edelman & Gally, 2001; Green, Crinion, & Price, 2006). We recognize such degeneracy in our everyday lives: We can wave a greeting with one arm or the other. Likewise, just as we can use different expressions to communicate a particular meaning, proficient bilinguals who habitually codeswitch explore degeneracy cross-linguistically by seeking alternative means to convey their intentions (Beatty-Martínez et al., 2020). Language regions in the brain are asymmetrically organized with a left hemisphere dominance for production, but lateralization can dynamically shift for comprehension in adult L2 learners (Gurunandan, Arnaez-Telleria, Carreiras, & Paz-Alonso, 2020). Degeneracy also enables inter-individual variation in cognitive and brain functioning more generally. For instance, in studies examining proactive vs. reactive control tendencies (i.e., whether goal-relevant information is monitored and maintained before the onset of cognitively challenging tasks or whether it is engaged as needed to changing task demands; Braver, 2012), group comparisons between bilinguals and monolinguals can yield similar behavioral outcomes, but electrophysiological and individual differences analyses reveal different strategies as to how each group coordinates both styles of control (Morales et al., 2013, 2015). The key point is that degeneracy enables recognition of how biological systems adapt to external demands (Edelman & Gally, 2007).

effect for one phenotype but not for another (Bak, 2016). On the grounds of degeneracy (see footnote 9), we can ask what kinds of cognitive and neural changes might be expected given the demands of particular contexts on language and the control processes supporting it (DeLuca, Segaert, Mazaheri, & Krott, 2020; Green & Abutalebi, 2013). We propose that a plausible answer to this question will require the application of ethnographic practices in brain and cognitive sciences (see Billig, 2020; Torres Cacoullos & Travis, 2018 for illustrations in psychology and variationist linguistics, respectively) tied to multi-lab collaborations (Leivada, Westergaard, Duñabeitia, & Rothman, 2020). We illustrate with a comparison of two Spanish-speaking locations, San Juan, Puerto Rico, and Granada, Spain, to highlight three key aspects of rich characterization using ethnography.

First, we need to be aware of the diachronic processes that have shaped the culture and history of a community. Granada is located in the community of Andalusia, officially considered one of several monolingual autonomous communities of Spain. Despite having a longstanding influence of Arabic culture, Andalusia has historically perpetuated ideologies tied to monolingualism, especially throughout the 20th century under the Francoist regime (see Lorenzo & Moore, 2009). As a result, foreign-language prevalence has remained lower in Andalusia relative to the rest of Spain and Europe (de Educación, 2012), creating fewer opportunities for other languages such as English to influence everyday language use. By comparison, Puerto Rico is the byproduct of a rich colonial history spanning across five centuries until the present (see Guzzardo Tamargo, Loureiro-Rodríguez, Acar, & Vélez Avilés, 2018). Although Spanish had been the established language following four centuries of Spanish colonial rule, the island became a US territory after the Spanish-American war at the end of the 19th century and continues to this day to be a non-incorporated territory of the US. Unlike Granada, the history of Puerto Rico created conditions in which American culture would become highly influential for the already established Hispanic culture, especially in the metropolitan area of San Juan.

The observation and description of current community practices in a well-defined speech community[10] (Labov, 2001) leads to the second feature of rich characterization. Determining how bilinguals' languages are habitually used (e.g., whether a speaker has extensive experience codeswitching or not) is important, but doing so requires an understanding of how the global environment of a community supports those practices. For instance, Spanish-English codeswitching is often a prominent form of communication among proficient bilinguals in San Juan whereas speakers from Granada tend to use their languages separately (see Beatty-Martínez & Dussias, 2017; Beatty-Martínez et al., 2020; Beatty-Martínez, Navarro-Torres, & Dussias, 2020 for a more comprehensive characterization). But in terms of interactional demands, the key difference between these two communities is not the frequency of switching behavior *per se*[11]; rather, it is the fact that, given the history and culture of each location, one context (i.e., San Juan) enables speakers to use both languages more openly and opportunistically with little-to-no interactional costs, whereas in the other context (i.e., Granada), there is a strong tendency to expect the use of Spanish (the L1) most of the time, creating constraints as to when speakers expect the use of the L2. This is not to say that codeswitching is not critical to understand how bilinguals control their languages (see Adler et al., 2020; Green, 2018; Hofweber et al., 2019 for how control processes may be engaged during codeswitching), but that the relative involvement of control processes during different kinds of speech acts may also depend on the demands imposed by the global environment. Characterizing speakers in terms of their habitual community practices is vital to understand such dynamics.

The final feature of rich characterization relates to changes in an individual's trajectory of experiences. While some bilinguals live in homogeneous communities where the language dynamics are relatively stable over the lifespan, other bilinguals undergo radical shifts in language use at particular time points (see Kubota, Chevalier, & Sorace, 2020; Pallier et al., 2003 for examples involving international returnees and adoptees). To illustrate, speakers may initially grow up in a home environment where a minority language (e.g., Spanish in the US) is used, but can then become educated and socialized in the majority language of the community (e.g., English in the US) during childhood. As such, some bilinguals (a.k.a., heritage speakers and indigenous-speaking bilinguals) may grow up and become educated in a context where the L1 is the majority language but then shift to an environment where the L2 becomes the dominant language (e.g., Garraffa, Beveridge, & Sorace, 2015; Garraffa, Obregon, & Sorace, 2017; Bonfieni, Branigan, Pickering, & Sorace, 2019; Polinsky & Scontras, 2020). A similar case is observed with young adults seeking higher education who relocate to a new environment (e.g., a foreign country) with a different predominant language (Beatty-Martínez, Bruni, Bajo, & Dussias, 2020). These shifts in language immersion status are likely to generate unique adaptive brain responses. Indeed, emerging evidence suggests that heritage speakers' initial minority-language experience has long-term consequences for language processing in the majority language (Bice & Kroll, 2021) and that contexts with high linguistic diversity (Gullifer et al., 2018) or L1-to-L2-immersion shifts (Beatty-Martínez et al., 2020) may trigger a novel adaptation of control processes in the form of proactive control engagement (see also Blanco-Elorrieta & Pylkkänen, 2018 for a similar observation regarding dual-language contexts).

Although we still know little about the boundary conditions of these effects, the point is that in a main-effect group analysis, different groups of speakers would be assumed to represent the same underlying population of bilinguals (see Weyman, Shake, & Redifer, 2020, for an illustration), despite having remarkably different community practices and/or individual trajectories that become evident through rich characterization. In making such an assumption, we may miss critical information that can change our conclusions. Hence, research on bilingualism[12] is likely to benefit more from small sample studies with rich characterization.

3.2. Beyond main effects and binary oppositions

Under a traditional lens in research on bilingualism, idiosyncratic

[10] Speakers can form part of stable speech communities, such as San Juan and Granada (see Torres Cacoullos & Travis, 2018 for another example in Albuquerque, New Mexico), where most individuals are members of the community, but they can also live in more dynamic and/or cosmopolitan communities, such as many major cities (e.g., London, Montreal) and some countries (e.g., Singapore). The distinction is important because it can help us infer the range of possible phenotypes and interactional demands that are likely to emerge in a given location.

[11] Notably, codeswitching is a relatively infrequent behavior even among habitual codeswitchers (Fricke & Kootstra, 2016), it can also be observed even among non-habitual codeswitchers such as bilinguals from Granada (see Table 9 in Beatty-Martínez & Dussias, 2017), and great discrepancy can exist even among bilingual communities that display habitual codeswitching (see Poplack, 1987 for a contrast between French-English bilinguals in Ottawa vs. Spanish-English Puerto Rican bilinguals in New York). The cognitive consequences of codeswitching in spontaneous discourse remain to be determined, but for now we make the point that an aggregate lump of codeswitchers vs. non-codeswitchers is misleading if rich characterization is not provided.

[12] Although our focus is on bilingualism, this proposal can serve a role in establishing the value of rich characterization and phenotyping procedures more generally. Evidently, the study of variation in language and cognition is central to any population of speakers, as has been established by research examining learning in monolinguals from different linguistic environments (e.g., Bice & Kroll, 2019), individual differences in language processing (e.g., Beatty-Martínez et al., 2020; Pakulak & Neville, 2010; Tanner & van Hell, 2014), as well as the consequences of dialectal experience for lexical and grammatical processing (e.g., Clopper, 2014; Squires, 2014).

patterns are typically discarded as random noise and complexity is equated with complication. Despite several notable critiques to this approach (Baum & Titone, 2014; Fricke, Zirnstein, Navarro-Torres, & Kroll, 2019; Luk & Bialystok, 2013; Tanner, McLaughlin, Herschensohn, & Osterhout, 2013), binary classifications and group comparisons, together with recommendations to keep experimental designs as simple as possible (Brysbaert, 2020), continue to dominate much of research on bilingualism, forcing discussions into a binary opposition not unlike those that have recently characterized the consequences of bilingualism (Nichols et al., 2020; c.f. Leivada et al., 2020), as well as psychological research more generally (Newell, 1973). But given the degenerate nature of biological and cognitive systems, solely main effect 'yes' or 'no' questions are unhelpful. As Bronfenbrenner (1977, p. 518) noted, "in ecological research, the principal main effects are likely to be interactions." To illustrate this point, we return to the study by Zirnstein et al. (2018; see 'Discovery acts on variety' subsection).

Following a main-effect analysis, Zirnstein and colleagues found different electrophysiological responses for L1 and L2 speakers to predictions errors during sentence reading: Only L1 speakers showed reliable electrophysiological costs when encountering semantically unexpected words. At first glance, it would be tempting to conclude that L2 speakers were unable to generate predictions, consistent with previous claims (e.g., Martin et al., 2013; Grüter, Rohde, & Schafer, 2017). But instead of asking whether L2 speakers can generate predictions (a 'yes' or 'no' question), one can ask about the cognitive processes that enable prediction in the first place. Upon examining individual differences in control and verbal fluency in both groups, Zirnstein and colleagues identified a more complex, but also more insightful, picture. First, both L1 and L2 speakers recruited control processes to recover from prediction errors (i.e., increased control ability related to reduced prediction costs). But for L2 speakers, as mentioned previously, there was an interaction between control and L1 fluency, such that increased L1 fluency related to larger prediction costs in the L2. This suggests that L2 speakers had to overcome the challenge of regulating the L1 in order to engage prediction mechanisms in ways comparable to L1 speakers. More critically, the interaction reveals that the absence of an electrophysiological response in the L2 group stemmed from an aggregate of bilingual phenotypes with different configurations of control and regulatory engagement. If we had only asked whether L2 speakers can generate predictions, we might have come to a different conclusion (see also Pulido, 2021; Tanner & van Hell, 2014 for illustrations with adult L2 learners and monolinguals, respectively).

The point is that a simple main-effects approach focused on attaining large samples or replication-via-uniformity would disregard the fact that the form of language and cognitive engagement varies across individuals. Arguably, there may be some important main effects, but the way to identify them is by first seeking out meaningful interactions that are informed by a rich characterization (see Rohrer & Arslan, 2021 for a discussion on the application of interactions). As a discovery process, science benefits from relatively open-ended questions such as "how do these regions in the temporal lobe dissociate during different tasks?" or "what are the possible range of phenotypes that can emerge in this community?". As Calhoun and Bandettini (2020) point out, such questions cast an effective net in making sense of large amounts of data[13].

3.3. Realizing signal extraction

Rich characterization, as well as the framing of our questions, is vital for effective signal extraction, but just as important is determining task and test sensitivity (e.g., for a given sample, to what extent do we expect a non-verbal task to tap executive processes used in language control so that any putative adaptive response of language experience could be realized?). We comment on four aspects. First, task sensitivity might require a revised conception of the task construct that it is designed to tap, such as using within-subject paradigms that allow us to induce different control states and track how they are engaged during language processing (e.g., Adler et al., 2020; Hsu et al., 2020; Navarro-Torres, Garcia, Chidambaram, & Kroll, 2020; Salig et al., 2021), as opposed to exclusively relying on aggregate executive function measures (e.g., Stroop effects) that likely mask degenerate patterns. As such, the insensitivity or the appropriateness of the test to tap into control processes engaged in language use limits their relevance for exploring any putative wider effects on non-verbal control tasks, as acknowledged in recent papers reporting data based on more richly characterized large samples (Gullifer & Titone, 2020a; Kałamała et al., 2020; Kheder & Kaan, 2021).

Second, deepening theoretical understanding also requires that we understand the totality of performance for which we need to consider data from a number of modalities – some of which may be more sensitive to the effects of interest than others. For instance, brain measures may better capture some aspects of early L2 learning than overt behavioral responses (e.g., Bice & Kroll, 2015, 2019; Kurkela, Hämäläinen, Leppänen, Shu, & Astikainen, 2019; McLaughlin, Osterhout, & Kim, 2004), although in other cases, both brain and behavior converge (see Li, Legault, & Litcofsky, 2014 for a review on imaging studies). Multi-lab collaborations can be an effective way to explore the boundary conditions of such issues for a given task and sample. We note, however, that the goal should not be replication *per se* –using paradigms that are simple enough to easily reproduce— but to see collaborations in ways that are designed to exploit variation across different labs and different locations (Leivada et al., 2020).

Third, it is important that we use measures that are reasonably commensurate with the questions being asked. For instance, with respect to determining how bilinguals' languages are habitually used, self-reported data can be informative (e.g., Gullifer & Titone, 2020b) but likely insufficient in the absence of conversational data that correspond to the vernacular of the speech community (Labov, 1984) or that reflect engagement of different attentional/control states when bilinguals shift between different modes of communication (Green, 2019).[14] The application of Network Science (Tiv, Gullifer, Feng, & Titone, 2020) and Information Theory (Gullifer & Titone, 2020b; Feldman, Srinivasan, Fernandes, & Shaikh, 2021) practices can also be of high value regarding

[13] Historically, Leibniz (1690/1951, cited in Gigerenzer, 1991, p. 254) likened scientific enquiry to "an ocean, continuous everywhere and without a break or division". Divided later by Reichenbach (1938) into two seas (the contexts of justification –hypothesis-testing— and the contexts of discovery –the generation of novel ideas): some have argued in favor of a sharp distinction between hypothesis-testing and exploration (Mertzen, Lago, & Vasishth, 2020), while others have argued that the only legitimate scientific practice is hypothesis-driven (Kullmann, 2020). But as we have argued, hypotheses arise in the context of an evolving understanding, and can vary in specificity, which is why there is no sharp division between the two contexts. Science as a discovery process entails the mingling of both creativity and empirical verification.

[14] For example, the conversational topics centered on individuals' personal experiences and that involve in-group members from the same speech community have been shown to increase the likelihood of codeswitching in informal contexts fourfold (Poplack, 1983).

effective phenotyping as they can help us establish correspondence between individual differences in language experience and the extent to which those trajectories reflect (or deviate from) more general community practices.

Finally, more sensitive data analysis practices are likely to be more revealing of individual differences and degenerate patterns. For instance, using ex-Gaussian distributions (Sundh, Collsiöö, Millroth, & Juslin, 2021; von Bastian, Blais, Brewer, Gyurkovics, Hedge, Kałamała, & Wiemers, 2020), delta plots (Morales et al., 2013), or Bayesian mixture models (Ferrigno, Cheyette, Piantadosi, & Cantlon, 2020) to infer the possible range of strategies in a given task, rather than simply averaging effects for a condition. Individual differences also allow us to construct generative models of behavior and neuroplasticity (see Parr, Rees, & Friston, 2018 for an example in neuropsychology) which can be used to computationally model, say, neuroplastic effects of different interactional contexts given a set of behavioral profiles. Further, within a large sample, there may be different phenotypes and we need to be able to explore and characterize these using data-driven techniques such as multivariate statistics (e.g., cluster and/or factor analyses; Hartanto & Yang, 2020; Rodriguez-Fornells, Krämer, Lorenzo-Seva, Festman, & Münte, 2012) if the data are sufficiently rich to detect different profiles. However, although individual differences offer an opportunity for effective phenotyping, they potentially involve the same risks as those observed with main-effect practices in the absence of a rich characterization tied to well-defined speech communities.

We consider it likely that all four aspects are pertinent to advance. Further, in some cases, and to reinforce our earlier point (see 'Discovery and power of the small' subsection), only small samples may be feasible and yield decisive evidence. For example, localization of phonemic restoration effects in the auditory cortices is best achieved through the high signal-to-noise ratio afforded by electrocorticography arrays implants for clinical purposes (see Leonard, Baud, Sjerps, & Chang, 2016), and so establishing convergence in bilingual speakers in two languages may sometimes require small samples with rich characterization (see "A rich characterization of the sample and an identification of boundary conditions" subsection). In short, replication, conventionalized tools, and large samples have value, but their role in the discovery process in research on bilingualism hinges on conceptual, experimental, and analytic advance.

4. Concluding remarks

In this paper, we have emphasized the community-value of incremental contributions via science as a discovery process against the enforcement of prescribed remedies, such as pre-determined sample sizes and/or methods, because we trust in the basic integrity of participants in the enterprise of research on bilingualism and ultimately in the self-correcting dynamic of science itself.

From the point of view of ensuring variety on which the quasi-Darwinian process of science can act, we require the publication of possibilities (e.g., sensitive tasks geared to testing specific processes, rich characterization to identify phenotypic variation, and non-binary questions that enable the exploration of interactions) that may or may not lead to deeper understanding. Replication and reproducibility efforts, important as they are to the scientific enterprise, need to be in service of such aims to advance. As the National Academies of Sciences, Engineering, and Medicine recently acknowledged: "The goal of science is not to compare or replicate [studies], but to understand the overall effect of a group of studies and the body of knowledge that emerges from them" (The National Academies of Sciences, Engineering, and Medicine, 2019, as cited in Miceli, 2019, p. 14). Like the brain, bilingualism is complex, and we are far from having a complete understanding of the boundary conditions of previously reported findings for replicability to be fruitful on its own. And while some have proposed that such understanding lies in the data itself (de Bruin, Treccani, & Della Sala, 2015b), we make the point that the answers ultimately lie in the characterization (i.e., the intersubjective agreement) of the data.

Finally, in establishing the need to view science as a discovery process, we wish to return to the question raised in the introduction of how compulsory prescriptions would come to affect research on bilingualism, as well as psychological and neuroscience research more generally. If we choose to allow prescriptions to dominate the scientific enterprise, then we must ask how they will come to shape not only the environment in which research is currently being conducted, but also how they will shape the minds of young and early-career researchers, and ultimately, whether we are willing to live with the consequences of those choices. Thus, in articulating the research enterprise of bilingualism, we hope to contribute to the establishment of a viable future research enterprise more generally.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The writing of this paper was supported in part by NSF Grant BCS-1946051 to J. F. Kroll and C. A. Navarro-Torres, NIH Fellowship F31-HD098783 to C. A. Navarro-Torres, NIH Fellowship F32-AG064810 to A. L. Beatty-Martínez, and by NSF Grant OISE-1545900 to J.F. Kroll.

References

Adamou, E. (2010). Bilingual speech and language ecology in Greek Thrace: Romani and Pomak in contact with Turkish. *Language in Society, 39*(2), 147–171.

Adler, R. M., Valdés Kroff, J. R., & Novick, J. (2020). Does integrating a code-switch during comprehension engage cognitive control? *Journal of Experimental Psychology: Learning Memory and Cognition., 46*, 741–759. https://doi.org/10.1037/xlm0000755.

Anderson, J. A. E., Hawrylewicz, K., & Grundy, J. G. (2020). Does bilingualism protect against dementia? A meta-analysis. *Psychonomic Bulletin & Review*. https://doi.org/10.3758/s13423-020-01736-5.

Antón, E., Duñabeitia, J. A., Estévez, A., Hernández, J. A., Castillo, A., Fuentes, L. J., … Carreiras, M. (2014). Is there a bilingual advantage in the ANT task? *Evidence from children. Frontiers in Psychology, 5*, 398. https://doi.org/10.3389/fpsyg.2014.00398.

Baddeley, A. D., & Warrington, E. K. (1970). Amnesia and the distinction between long- and short-term memory. *Journal of Verbal Learning and Verbal Behavior, 9*(2), 176–189. https://doi.org/10.1016/S0022-5371(70)80048-2.

Bak, T. H., Nissan, J. J., Allerhand, M. M., & Deary, I. J. (2014). Does bilingualism influence cognitive aging? *Annals of Neurology, 75*(6), 959–963. https://doi.org/10.1002/ana.v75.610.1002/ana.24158.

Bak, T. H. (2016). Cooking pasta in La Paz: Bilingualism, bias and the replication crisis. In Linguistic Approaches to Bilingualism (Vol. 6, Issue 5, pp. 699–717). John Benjamins.

Barrett, L. F. (2017). The theory of constructed emotion: An active inference account of interoception and categorization. *Social Cognitive and Affective Neuroscience, 12*(1), 1–23. https://doi.org/10.1093/scan/nsw154.

Barrett, M. (2018, May 13). Penicillin's first patient. Retrieved from: https://mosaicscience.com/story/penicillin-first-patient-history-albert-alexander-AMR-DRI/.

Baum, S., & Titone, D. (2014). Moving toward a neuroplasticity view of bilingualism, executive control, and aging. *Applied Psycholinguistics, 35*(5), 857–894. https://doi.org/10.1017/S0142716414000174.

Beatty-Martínez, A. L., Bruni, M. R., Bajo, M. T., & Dussias, P. E. (2020). Brain potentials reveal differential processing of masculine and feminine grammatical gender in native Spanish speakers. *Psychophysiology, 58*(3). https://doi:10.1111/psyp.13737.

Beatty-Martínez, A. L., & Dussias, P. E. (2020). Bilingual experience shapes language processing: Evidence from codeswitching. *Journal of Memory and Language, 95*, 173–189. https://doi.org/10.1016/j.jml.2017.04.002.

Beatty-Martínez, A. L., Navarro-Torres, C. A., & Dussias, P. E. (2020). Codeswitching: A bilingual toolkit for opportunistic speech planning. *Frontiers in Psychology, 11*, 1699. https://doi.org/10.3389/fpsyg.2020.01699.

Beatty-Martínez, A. L., Navarro-Torres, C. A., Dussias, P. E., Bajo, M. T., Guzzardo Tamargo, R. E., & Kroll, J. F. (2020). Interactional context mediates the consequences of bilingualism for language and cognition. *Journal of Experimental Psychology-Learning Memory and Cognition, 46*, 1022–1047. https://doi.org/10.1037/xlm0000770.

Bialystok, E. (2011). Coordination of executive functions in monolingual and bilingual children. *Journal of Experimental Child Psychology, 110*(3), 461–468. https://doi.org/10.1016/j.jecp.2011.05.005.

Bialystok, E., Kroll, J. F., Green, D. W., MacWhinney, B., & Craik, F. I. M. (2015). Publication bias and the validity of evidence: What's the connection? *Psychological Science, 26*(6), 944–946. https://doi.org/10.1177/0956797615573759.

Bice, K., & Kroll, J. F. (2021). Grammatical processing in two languages: How individual differences in language experience and cognitive abilities shape comprehension in heritage bilinguals. *Journal of Neurolinguistics, 58*, 100963. https://doi.org/10.1016/j.jneuroling.2020.100963.

Bice, K., & Kroll, J. F. (2019). English only? Monolinguals in linguistically diverse contexts have an edge in language learning. *Brain & Language, 196*, 104644. https://doi.org/10.1016/j.bandl.2019.104644.

Bice, K., & Kroll, J. F. (2015). Native language change during early stages of second language learning. *NeuroReport, 26*, 966–971. https://doi.org/10.1097/WNR.0000000000000453.

Billig, M. (2019). *More examples, less theory*. Cambridge University Press.

Blanco-Elorrieta, E., Emmorey, K., & Pylkkänen, L. (2018). Language switching decomposed through MEG and evidence from bimodal bilinguals. *Proceedings of the National Academy of Sciences, 115*(39), 9708–9713. https://doi.org/10.1073/pnas.1809779115.

Blanco-Elorrieta, E., & Pylkkänen, L. (2018). Ecological validity in bilingualism research and the bilingual advantage. *Trends in Cognitive Sciences, 22*(12), 1117–1126. https://doi.org/10.1016/j.tics.2018.10.001.

Blasi, D. E., Moran, S., Moisik, S. R., Widmer, P., Dediu, D., & Bickel, B. (2019). Human sound systems are shaped by post-Neolithic changes in bite configuration. *Science, 363*(6432), eaav3218. https://doi.org/10.1126/science:aav3218.

Bonfieni, M., Branigan, H. P., Pickering, M. J., & Sorace, A. (2019). Language experience modulates bilingual language control: The effect of proficiency, age of acquisition, and exposure on language switching. *Acta Psychologica, 193*, 160–170. https://doi.org/10.1016/j.actpsy.2018.11.004.

Bosma, E., & Pablos, L. (2020). Switching direction modulates the engagement of cognitive control in bilingual reading comprehension: An ERP study. *Journal of Neurolinguistics, 55*, 100894. https://doi.org/10.1016/j.jneuroling.2020.100894.

Bowers, K. S., Regehr, G., Balthazard, C., & Parker, K. (1990). Intuition in the context of discovery. *Cognitive Psychology, 22*(1), 72–110. https://doi.org/10.1016/0010-0285(90)90004-N.

Brace, C. L. (1986). Egg on the face, f in the mouth, and the overbite. *American Anthropologist, 88*(3), 695–697. https://doi.org/10.1525/aa.1986.88.issue-310.1525/aa.1986.88.3.02a00150.

Braver, T. S. (2012). The variable nature of cognitive control: A dual-mechanisms framework. *Trends in Cognitive Sciences, 16*(2), 106–113. https://doi.org/10.1016/j.tics.2011.12.010.

Bronfenbrenner, U. (1977). Toward an experimental ecology of human development. *American Psychologist, 32*(7), 513–531. https://doi.org/10.1037/0003-066X.32.7.513.

Brysbaert, M. (2020). Power considerations in bilingualism research: Time to step up our game. *Bilingualism: Language and Cognition*, 1–6. https://doi.org/10.1017/S1366728920000437.

Calhoun, V., & Bandettini C. (2020, May 22). Defending fMRI, brain mapping, and discovery science. [Blog post]. Retrieved from: http://www.thebrainblog.org/2020/05/22/defending-fmri-brain-mapping-and-discovery-science/.

Clopper, C. G. (2014). Sound change in the individual: Effects of exposure on cross-dialect speech processing. *Laboratory Phonology, 5*, 69–90. https://doi.org/10.1515/lp-2014-0004.

de Bruin, A., Treccani, B., & Della Sala, S. (2015). Cognitive advantage in bilingualism: An example of publication bias? *Psychological Science, 26*(1), 99–107. https://doi.org/10.1177/0956797614557866.

de Bruin, A., Treccani, B., & Della Sala, S. (2015). The connection is in the data: We should consider them all. *Psychological Science, 26*(6), 947–949. https://doi.org/10.1177/0956797615583443.

Deacon, T. W. (2010). A role for relaxed selection in the evolution of the language capacity. *Proceedings of the National Academy of Sciences of the United States of America, 107*(Supplement_2), 9000–9006. https://doi.org/10.1073/pnas.0914624107.

DeLuca, V., Segaert, K., Mazaheri, A., & Krott, A. (2020, May 29). Understanding bilingual brain function and structure changes? U Bet! A unified bilingual experience trajectory model. DOI: http://dx.doi.10.31234/osf.io/gkysn.

Dick, A. S., Garcia, N. L., Pruden, S. M., Thompson, W. K., Hawes, S. W., Sutherland, M. T., … Gonzalez, R. (2019). No evidence for a bilingual executive function advantage in the ABCD study. *Nature Human Behaviour, 3*(7), 692–701. https://doi.org/10.1038/s41562-019-0609-3.

Donnelly, S., Brooks, P. J., & Homer, B. D. (2019). Is there a bilingual advantage on interference-control tasks? A multiverse meta-analysis of global reaction time and interference cost. *Psychonomic Bulletin & Review, 26*(4), 1122–1147. https://doi.org/10.3758/s13423-019-01567-z.

Duñabeitia, J. A., Hernández, J. A., Antón, E., Macizo, P., Estévez, A., Fuentes, L. J., & Carreiras, M. (2014). The inhibitory advantage in bilingual children revisited: Myth or reality? *Experimental Psychology, 61*(3), 234–251. https://doi.org/10.1027/1618-3169/a000243.

Edelman, G. M., & Gally, J. A. (2001). Degeneracy and complexity in biological systems. *Proceedings of the National Academy of Sciences of the United States of America, 98*(24), 13763–13768. https://doi.org/10.1073/pnas.231499798.

Emmorey, K., Borinstein, H. B., Thompson, R., & Gollan, T. H. (2008). Bimodal bilingualism. *Bilingualism, Language, and Cognition, 11*(1), 43–61. https://doi.org/10.1017/S1366728907003203.

Emmorey, K., Li, C., Petrich, J., & Gollan, T. H. (2020). Turning languages on and off: Switching into and out of code-blends reveals the nature of bilingual language control. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 46*(3), 443–454. https://doi.org/10.1037/xlm0000734.

Fabbro, F., Skrap, M., & Aglioti, S. (2000). Pathological switching between languages following frontal lesions in a bilingual patient. *Journal of Neurology, Neurosurgery and Psychiatry, 68*, 650–652. https://doi.org/10.1136/jnnp.68.5.650.

Feldman, L. B., Srinivasan, V., Fernandes, R. B., & Shaikh, S. (2021). Insights into codeswitching from online communication: Effects of language preference and conditions arising from vocabulary richness. *Bilingualism: Language and Cognition, 24*(4), 791–797. https://doi.org/10.1017/S1366728921000122.

Ferrigno, S., Cheyette, S. J., Piantadosi, S. T., & Cantlon, J. F. (2020). Recursive sequence generation in monkeys, children, U.S. adults, and native Amazonians. *Science. Advances, 6*(26), eaaz1002. https://doi.org/10.1126/sciadv.aaz1002.

Fricke, M., & Kootstra, G. J. (2016). Primed codeswitching in spontaneous bilingual dialogue. *Journal of Memory and Language, 91*, 181–201. https://doi.org/10.1016/j.jml.2016.04.003.

Fricke, M., Zirnstein, M., Navarro-Torres, C. A., & Kroll, J. F. (2019). Bilingualism reveals fundamental variation in language processing. *Bilingualism: Language and Cognition, 22*(1), 200–207. https://doi.org/10.1017/S1366728918000482.

García-Pentón, L., Fernández García, Y., Costello, B., Duñabeitia, J. A., & Carreiras, M. (2016a). The neuroanatomy of bilingualism: How to turn a hazy view into the full picture. *Language, Cognition and Neuroscience, 31*(3), 303–327. https://doi.org/10.1080/23273798.2015.1068944.

García-Pentón, L., Fernández García, Y., Costello, B., Duñabeitia, J. A., & Carreiras, M. (2016b). "Hazy" or "jumbled"? Putting together the pieces of the bilingual puzzle. *Language, Cognition and Neuroscience, 31*(3), 353–360. https://doi.org/10.1080/23273798.2015.1135247.

Garraffa, M., Beveridge, M., & Sorace, A. (2015). Linguistic and Cognitive Skills in Sardinian-Italian Bilingual Children. *Frontiers in Psychology, 6*, 1898. https://doi.org/10.3389/fpsyg.2015.01898.

Garraffa, M., Obregon, M., & Sorace, A. (2017). Linguistic and Cognitive Effects of Bilingualism with Regional Minority Languages: A Study of Sardinian-Italian Adult Speakers. *Frontiers in Psychology, 8*, 1907. https://doi.org/10.3389/fpsyg.2017.01907.

Gigerenzer, G. (1991). From tools to theories: A heuristic of discovery in cognitive psychology. *Psychological Review, 98*, 254–267. https://doi.org/10.1037/0033-295X.98.2.254.

Gould, S. J., & Vrba, E. S. (1982). Exaptation-a missing term in the science of form. *Paleobiology, 8*, 4–15. http://www.jstor.com/stable/2400563.

Green, D. W. (1986). Control, activation, and resource: A framework and a model for the control of speech in bilinguals. *Brain and Language, 27*, 210–223. https://doi.org/10.1016/0093-934X(86)90016-7.

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition, 1*, 67–81. https://doi.org/10.1017/S1366728998000133.

Green, D. W. (2018). Language control and code-switching. *Languages, 3*(2), 8. https://doi.org/10.3390/languages3020008.

Green, D. W. (2019). Language control and attention during conversation: An exploration. In J. W. Schwieter, & M. Paradis (Eds.), *The Handbook of the Neuroscience of Multilingualism* (pp. 427–446). John Wiley & Sons. https://doi.org/10.1002/9781119387725.ch21.

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.

Green, D. W., & Abutalebi, J. (2015). Language control and the neuroanatomy of bilingualism: In praise of variety. *Language, Cognition and Neuroscience, 3798* (October), 1–5. https://doi.org/10.1080/23273798.2015.1084428.

Green, D. W., Crinion, J., & Price, C. J. (2006). Convergence, degeneracy, and control. *Language Learning, 56*, 99–125. https://doi.org/10.1111/j.1467-9922.2006.00357.x.

Green, D. W., & Kroll, J. F. (2019). The neurolinguistics of bilingualism: Plasticity and control. In G. I. de Zubicaray, & N. O. Schiller (Eds.), *The Oxford Handbook of Neurolinguistics*. https://doi.org/10.13153/diam.8.2006.233.

Grundy, J. G., & Timmer, K. (2017). Bilingualism and working memory capacity: A comprehensive meta-analysis. *Second Language Research, 33*(3), 325–340. https://doi.org/10.1177/0267658316678286.

Grüter, T., Rohde, H., & Schafer, A. J. (2017). Coreference and discourse coherence in L2. *Linguistic Approaches to Bilingualism, 7*(2), 199–229. https://doi.org/10.1075/lab10.1075/lab.7.210.1075/lab.15011.gru.

Gullifer, J. W., & Titone, D. (2020a). Engaging proactive control: Influences of diverse language experiences using insights from machine learning. *Journal of Experimental Psychology: General. Advance online publication., 150*(3), 414–430. https://doi.org/10.1037/xge0000933.

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

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*, 123–134. https://doi.org/10.1016/j.neuropsychologia.2018.04.037.

Gurunandan, K., Arnaez-Telleria, J., Carreiras, M., & Paz-Alonso, P. M. (2020). Converging evidence for differential specialisation and plasticity of language systems. *Journal of Neuroscience, 40*, 9715–9724. http://www.jneurosci.org/lookup/doi/10.1523/JNEUROSCI.0851-20.2020.

Guzzardo Tamargo, R. E., Loureiro-Rodríguez, V., Acar, E. F., & Vélez Avilés, J. (2019). Attitudes in progress: Puerto Rican youth's opinions on monolingual and code-switched language varieties. *Journal of Multilingual and Multicultural Development, 40*(4), 304–321. https://doi.org/10.1080/01434632.2018.1515951.

Hartanto, A., Toh, W. X., & Yang, H. (2018). Context counts: The different implications of weekday and weekend video gaming for academic performance in mathematics, reading, and science. *Computers & Education, 120*, 51–63. https://doi.org/10.1016/j.compedu.2017.12.007.
Hartanto, A., & Yang, H. (2020). The role of bilingual interactional contexts in predicting interindividual variability in executive functions: A latent variable analysis. *Journal of Experimental Psychology: General, 149*, 609–633. https://doi.org/10.1037/xge0000672.
Heisenberg, W. (1958). *Physics and Philosophy*. New York: Harper.
Hockett, C. F. (1985). Distinguished lecture: F. *American Anthropologist, 87*(2), 263–281. https://doi.org/10.1525/aa.1985.87.issue-210.1525/aa.1985.87.2.02a00020.
Hofweber, J., Marinis, T., Treffers-Daller, J., Sekerina, I. A., Spradlin, L., & Valian, V. (2019). Predicting executive functions in bilinguals using ecologically valid measures of code-switching behaviour. *Bilingualism, Executive Functions, and Beyond. Questions and Insights*, 181–205.
Hsu, N. S., Jaeggi, S. M., & Novick, J. M. (2017). A common neural hub resolves syntactic and non-syntactic conflict through cooperation with task-specific networks. *Brain and Language, 166*, 63–77. https://doi.org/10.1016/j.bandl.2016.12.006.
Hsu, N. S., Kuchinsky, S. E., & Novick, J. M. (2021). Direct impact of cognitive control on sentence processing and comprehension. *Language, Cognition and Neuroscience, 36*(2), 211–239. https://doi.org/10.1080/23273798.2020.1836379.
Iyer, A. S., Jones, F. K., Nodoushani, A., Kelly, M., Becker, M., Slater, D., … Charles, R. C. (2020). Dynamics and significance of the antibody response to SARS-CoV-2 infection. *MedRxiv: The Preprint Server for Health Sciences*, 1–31. https://doi.org/10.1101/2020.07.18.20155374.
Jiao, L.u., Grundy, J. G., Liu, C., & Chen, B. (2020). Language context modulates executive control in bilinguals: Evidence from language production. *Neuropsychologia, 142*, 107441. https://doi.org/10.1016/j.neuropsychologia.2020.107441.
Jiao, L., Liu, C., de Bruin, A., & Chen, B. (2020). Effects of language context on executive control in unbalanced bilinguals: An ERPs study. *Psychophysiology, 57*, 1–16. https://doi.org/10.1111/psyp.13653.
Kałamała, P., Szewczyk, J., Chuderski, A., Senderecka, M., & Wodniecka, Z. (2020). Patterns of bilingual language use and response inhibition: A test of the adaptive control hypothesis. *Cognition, 204*, 104373. https://doi.org/10.1016/j.cognition.2020.104373.
Kheder, S., & Kaan, E. (2021). Cognitive control in bilinguals: Proficiency and code-switching both matter. *Cognition, 209*, 104575. https://doi.org/10.1016/j.cognition.2020.104575.
Kubota, M., Chevalier, N., & Sorace, A. (2020). Losing access to the second language and its effect on executive function development in childhood: The case of "returnees". *Journal of Neurolinguistics, 55*, 100906. https://doi.org/10.1016/j.jneuroling.2020.100906.
Kullmann, D. M. (2020). Editorial. *Brain, 143*, 1045. https://doi.org/10.1093/brain/awaa082.
Kurkela, J. L. O., Hämäläinen, J. A., Leppänen, P. H. T., Shu, H., & Astikainen, P. (2019). Passive exposure to speech sounds modifies change detection brain responses in adults. *NeuroImage, 188*, 208–216. https://doi.org/10.1016/j.neuroimage.2018.12.010.
Labov, W. (1984). Field methods of the project in linguistic change and variation. In J. Baugh, & J. Scherzer (Eds.), *Language in Use* (pp. 28–52). Upper Saddle River, New Jersey: Prentice-Hall.
Labov, W. (2001). Principles of linguistic change. Social factors (Vol. II). Oxford: Blackwell.
Laudan, L. (1971). William Whewell on the consilience of inductions. *The Monist*, 368–391. http://www.jstor.org/stable/27902225.
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*, 394–425. https://doi.org/10.1037/bul0000142.
Leibniz, G. W. (1951). The horizon of human doctrine. In P. P. Wiener (Ed.), *Selections* (pp. 73–77). New York: Scribner's Sons (Original work published 1690).
Leivada, E., Westergaard, M., Duñabeitia, J. A., & Rothman, J. (2021). On the phantom-like appearance of bilingualism effects on neurocognition: (How) should we proceed? *Bilingualism: Language and Cognition, 24*(1), 197–210. https://doi.org/10.1017/S1366728920000358.
Leonard, M. K., Baud, M. O., Sjerps, M. J., & Chang, E. F. (2016). Perceptual restoration of masked speech in human cortex. *Nature Communication, 7*, 13619. https://doi.org/10.1038/ncomms13619.
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.
Liu, T., Wu, S., Tao, H., Zeng, G., & Zhou, F. (2020). Prevalence of IgG antibodies to SARS-CoV-2 in Wuhan: Implications for the ability to produce long-lasting protective antibodies against SARS-CoV-2. *MedRxiv, 169*, 1–30. https://doi.org/10.1101/2020.06.13.20130252.
Lorca-Puls, D. L., Gajardo-Vidal, A., White, J., Seghier, M. L., Leff, A. P., Green, D. W., … Price, C. J. (2018). The impact of sample size on the reproducibility of voxel-based lesion-deficit mappings. *Neuropsychologia, 115*, 101–111. https://doi.org/10.1016/j.neuropsychologia.2018.03.014.
Lorenzo, F., & Moore, P. (2009). European language policies in monolingual southern Europe: Implementation and outcomes. *European Journal of Language Policy, 1*(2), 121–135. https://doi.org/10.3828/ejlp.2009.3.
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. https://doi.org/10.1080/20445911.2013.795574.
McLaughlin, J., Osterhout, L., & Kim, A. (2004). Neural correlates of second-language word learning: Minimal instruction produces rapid change. *Nat Neurosci, 7*(7), 703–704. https://doi.org/10.1038/nn1264.
Martin, C. D., Thierry, G., Kuipers, J.-R., Boutonnet, B., Foucart, A., & Costa, A. (2013). Bilinguals reading in their second language do not predict upcoming words as native readers do. *Journal of Memory and Language, 69*(4), 574–588. https://doi.org/10.1016/j.jml.2013.08.001.
Miceli, S. (2019, Sept 3). Reproducibility and replicability in research. Retrieved from: https://www.nationalacademies.org/news/2019/09/reproducibility-and-replicability-in-research.
Mertzen, D., Lago, S., & Vasishth, S. (2020, February 28). The benefits of preregistration for hypothesis-driven bilingualism research. http://dx.doi.10.31234/osf.io/nm3eg.
Ministerio de Educación, Cultura y Deporte. (2012). Estudio europeo de competencia lingüística EECL: Informe español. Madrid: Secretaría General Técnica. https://www.comunidad.madrid/sites/default/files/doc/educacion/sgea_eval_eecl_nacional_vol1_2011.pdf (accessed 5 January 2020).
Morales, J., Gómez-Ariza, C. J., & Bajo, M. T. (2013). Dual mechanisms of cognitive control in bilinguals and monolinguals. *Journal of Cognitive Psychology, 25*(5), 531–546. https://doi.org/10.1080/20445911.2013.807812.
Morales, J., Yudes, C., Gómez-Ariza, C. J., & Bajo, M. T. (2015). Bilingualism modulates dual mechanisms of cognitive control: Evidence from ERPs. *Neuropsychologia, 66*, 157–169. https://doi.org/10.1016/j.neuropsychologia.2014.11.014.
Mukadam, N., Sommerlad, A., & Livingston, G. (2017). The relationship of bilingualism compared to monolingualism to the risk of cognitive decline or dementia: A systematic review and meta-analysis. *Journal of Alzheimer's Disease, 58*(1), 45–54. https://doi.org/10.3233/JAD-170131.
KK Nair, V., Rayner, T., Siyambalapitiya, S., & Biedermann, B. (2021). Domain-general cognitive control and domain-specific language control in bilingual aphasia: A systematic quantitative literature review. *Journal of Neurolinguistics, 60*, 101021. https://doi.org/10.1016/j.jneuroling.2021.101021.
The National Academies of Sciences, Engineering, and Medicine (2019, September 11). Reproducibility and Replicability - Harvey Fineberg [Video]. YoutTube. https://www.youtube.com/watch?v=lMD9TM1HAcc.
Navarro-Torres, C. A., Garcia, D. L., Chidambaram, V., & Kroll, J. F. (2019). Cognitive control facilitates attentional disengagement during second language comprehension. *Brain Sciences, 9*(5), 95. https://doi.org/10.3390/brainsci9050095.
Newell, A. (1973). You can't play 20 questions with nature and win: Projective comments on the papers of this symposium. In W. G. Chase (Ed.), *Visual information processing* (pp. 283–308). New York, NY: Academic Press.
Nichols, E. S., Wild, C. J., Stojanoski, B., Battista, M. E., & Owen, A. M. (2020). Bilingualism affords no general cognitive advantages: A population study of executive function in 11,000 people. *Psychological Science, 31*(5), 548–567. https://doi.org/10.1177/0956797620903113.
Noor, S., Friston, K. J., Ekert, J. O., Price, C. J., & Green, D. W. (2020). Neuromodulatory control and language recovery in bilingual aphasia: An active inference approach. *Behavioral Sciences, 10*, 161. https://doi.org/10.3390/bs10100161.
Paap, K. R., Anders-Jefferson, R., Mason, L., Alvarado, K., & Zimiga, B. (2018). Bilingual advantages in inhibition or attentional control: More challenges. *Frontiers in Psychology, 9*, 1409. https://doi.org/10.3389/FPSYG.2018.01409.
Paap, K. R., & Greenberg, Z. I. (2013). There is no coherent evidence for a bilingual advantage in executive processing. *Cognitive Psychology, 66*(2), 232–258. https://doi.org/10.1016/j.cogpsych.2012.12.002.
Pallier, C., Dehaene, S., Poline, J.-B., LeBihan, D., Argenti, A.-M., Dupoux, E., & Mehler, J. (2003). Brain imaging of language plasticity in adopted adults: Can a second language replace the first? *Cerebral Cortex, 13*, 1047–3211. https://doi.org/10.1093/cercor/13.2.155.
Pakulak, E., & Neville, H. J. (2010). Proficiency differences in syntactic processing of monolingual native speakers indexed by event-related potentials. *Journal of Cognitive Neuroscience, 12*, 2728–2744. https://doi.org/10.1162/jocn.2009.21393.
Paradis, M., Goldblum, M.-C., & Abidi, R. (1982). Alternate antagonism with paradoxical translation in two bilingual aphasic patients. *Brain and Language, 15*, 55–69. https://doi.org/10.1016/0093-934X(82)90046-3.
Parr, T., Rees, G., & Friston, K. J. (2018). Computational Neuropsychology and Bayesian Inference. *Frontiers in Human Neuroscience, 12*, 61. https://doi.org/10.3389/fnhum.2018.00061.
Peal, E., & Lambert, W. E. (1962). The relation of bilingualism to intelligence. *Psychological Monographs: General and Applied, 76*, 1–23. https://doi.org/10.1037/h0093840.
Petitto, L. A., Katerelos, M., Levy, B. G., Gauna, K., Tétreault, K., & Ferraro, V. (2001). Bilingual signed and spoken language acquisition from birth: Implications for the mechanisms underlying early bilingual language acquisition. *Journal of Child Language, 28*(2), 453–496. https://doi.org/10.1017/S0305000901004718.
Pliatsikas, C., & Luk, G. (2016). Executive control in bilinguals: A concise review on fMRI studies. *Bilingualism, Language, and Cognition, 19*(4), 699–705. https://doi.org/10.1017/S1366728916000249.
Polinsky, M., & Scontras, G. (2020). Understanding heritage languages. *Bilingualism: Language and Cognition, 23*(1), 4–20. https://doi.org/10.1017/S1366728919000245. Cambridge Core.
Pollán, M., Pérez-Gómez, B., Pastor-Barriuso, R., Oteo, J., Hernán, M. A., Pérez-Olmeda, M., … Vázquez de la Villa, A. (2020). Prevalence of SARS-CoV-2 in Spain (ENE-COVID): A nationwide, population-based seroepidemiological study. *The Lancet, 396*(10250), 535–544. https://doi.org/10.1016/S0140-6736(20)31483-5.
Poplack, S. (1983). Bilingual competence: Linguistic interference or grammatical integrity? In E. Olivares (Ed.), *Spanish in the US Setting: Beyond the Southwest* (pp. 107–131). Arlington: National Clearinghouse for Bilingual Education.

Poplack, S. (1987). In *Contrasting patterns of code-switching in two communities" in aspects of multilingualism* (pp. 51–77). Uppsala: Borgström.

Pulido, M. F. (2021). Individual chunking ability predicts efficient or shallow L2 processing: Eye-tracking evidence from multiword units in relative clauses. *Frontiers in Psychology, 11*. https://doi.org/10.3389/fpsyg.2020.60762110.3389/fpsyg.2020.607621.s001.

Reichenbach, H. (1938). *Experience and prediction*. Chicago: University of Chicago Press.

Rodriguez-Fornells, A., Krämer, U. M., Lorenzo-Seva, U., Festman, J., & Münte, T. F. (2012). Self-assessment of individual differences in language switching. *Frontiers in Psychology, 2*. https://doi.org/10.3389/fpsyg.2011.00388.

Rohrer, J. M., & Arslan, R. C. (2021). Precise answers to vague questions: Issues with interactions. *Advances in Methods and Practices in Psychological Science*. https://doi.org/10.1177/25152459211007368.

Salig, L. K., Valdés Kroff, J. R., Robert, S. L., & Novick, J. M. (2021). Moving from bilingual traits to states: Understanding cognition and language processing through moment-to-moment variation. *Neurobiology of Language*, 1–62. https://doi.org/10.1162/nol_a_00046. Advanced publication.

Salzman, C. D., Murasugi, C. M., Britten, K. H., & Newsome, W. T. (1992). Microstimulation in visual area MT: Effects on direction discrimination performance. *Journal of Neuroscience, 12*(6), 2331–2355.

Santillán, J., & Khurana, A. (2018). Developmental associations between bilingual experience and inhibitory control trajectories in Head Start children. *Developmental Science, 21*(4), e12624. https://doi.org/10.1111/desc.2018.21.issue-410.1111/desc.12624.

Schroeder, S. R. (2018). Do bilinguals have an advantage in theory of mind? *A meta-analysis. Frontiers in Communication, 3*, 36. https://doi.org/10.3389/fcomm.2018.00036.

Shallice, T., & Warrington, E. K. (1970). Independent functioning of verbal memory stores: A neuropsychological study. *Quarterly Journal of Experimental Psychology, 22* (2), 261–273. https://doi.org/10.1080/00335557043000203.

Simpson, D. (2005). Phrenology and the neurosciences: Contributions of F. J. Gall and. *J. G. Spurzheim. ANZ Journal of Surgery, 75*(6), 475–482. https://doi.org/10.1111/ans.2005.75.issue-610.1111/j.1445-2197.2005.03426.x.

Squires, L. (2014). Social differences in the processing of grammatical variation. University of Pennsylvania Working Papers. *Linguistics, 20*, 178–188.

Stout, D., & Chaminade, T. (2012). Stone tools, language, and the brain in human evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences, 367* (1585), 75–87. https://doi.org/10.1098/rstb.2011.0099.

Sundh, J., Collsiöö, A., Millroth, P., & Juslin, P. (2021). Precise/not precise (PNP): A Brunswikian model that uses judgment error distributions to identify cognitive processes. *Psychonomic Bulletin & Review, 28*(2), 351–373. https://doi.org/10.3758/s13423-020-01805-9.

Sulpizio, S., Del Maschio, N., Fedeli, D., & Abutalebi, J. (2020). Bilingual language processing: A meta-analysis of functional neuroimaging studies. *Neuroscience & Biobehavioral Reviews, 108*, 834–853. https://doi.org/10.1016/j.neubiorev.2019.12.014.

Szucs, D., & Ioannidis, J. PA. (2020). Sample size evolution in neuroimaging research: An evaluation of highly-cited studies (1990–2012) and of latest practices (2017–2018) in high-impact journals. *NeuroImage, 221*, 117164. https://doi.org/10.1016/j.neuroimage.2020.117164.

Tanner, D., McLaughlin, J., Herschensohn, J., & Osterhout, L. (2013). Individual differences reveal stages of L2 grammatical acquisition: ERP evidence. *Bilingualism: Language and Cognition, 16*, 367–382. https://doi.org/10.1017/S1366728912000302.

Tanner, D., & van Hell, J. G. (2014). ERPs reveal individual differences in morphosyntactic processing. *Neuropsychologia, 56*, 289–301. https://doi.org/10.1016/j.neuro psych ologia.2014.02.002.

Tillett, R. L., Sevinsky, J. R., Hartley, P. D., Kerwin, H., Crawford, N., Gorzalski, A., … Pandori, M. (2021). Genomic evidence for reinfection with SARS-CoV-2: A case study. *The Lancet Infectious Diseases, 21*(1), 52–58. https://doi.org/10.1016/S1473-3099(20)30764-7.

Tiv, M., Gullifer, J. W., Feng, R. Y., & Titone, D. (2020). Using network science to map what Montréal bilinguals talk about across languages and communicative contexts. *Journal of Neurolinguistics, 56*, 100913. https://doi.org/10.1016/j.jneuroling.2020.100913.

Torres Cacoullos, R., & Travis, C. E. (2018). *Bilingualism in the Community: Code-switching and Grammars in Contact*. Cambridge University Press. Cambridge Core http://dx.doi.10.1017/9781108235259.

Valian, V. (2015). Bilingualism and cognition. *Bilingualism: Language and Cognition, 18* (1), 3–24. https://doi.org/10.1017/S1366728914000522.

Vetter, P., Bola, Ł., Reich, L., Bennett, M., Muckli, L., & Amedi, A. (2020). Decoding natural sounds in early "visual" cortex of congenitally blind individuals. *Current Biology, 30*(15), 3039–3044.e2. https://doi.org/10.1016/j.cub.2020.05.071.

von Bastian, C. C., Blais, C., Brewer, G. A., Gyurkovics, M., Hedge, C., Kałamała, P., … Wiemers, E. A. (2020, July 27). Advancing the understanding of individual differences in attentional control: Theoretical, methodological, and analytical considerations. http://dx.doi.10.31234/osf.io/x3b9k.

Weyman, K. M., Shake, M., & Redifer, J. L. (2020). Extensive experience with multiple languages may not buffer age-related declines in executive function. *Experimental Aging Research, 46*(4), 291–310. https://doi.org/10.1080/0361073X.2020.1753402.

Wu, J., Liang, B., Chen, C., Wang, H., Fang, Y., Shen, S., … Lu, S. (2020). SARS-CoV-2 infection induces sustained humoral immune responses in convalescent patients following symptomatic COVID-19 Correspondence. *MedRxiv*. https://doi.org/10.1101/2020.07.21.20159178.

Wu, Y. J., & Thierry, G. (2013). Fast modulation of executive function by language context in bilinguals. *Journal of Neuroscience, 33*, 13533–13537. https://doi.org/10.1523/JNEUROSCI.4760-12.2013.

Zirnstein, M., van Hell, J. G., & Kroll, J. F. (2018). Cognitive control ability mediates prediction costs in monolinguals and bilinguals. *Cognition, 176*, 87–106. https://doi.org/10.1016/j.cognition.2018.03.001.