Understanding bilingual brain function and structure changes? U Bet!

A Unified Bilingual Experience Trajectory model

Vincent DeLuca, Katrien Segaert, Ali Mazaheri, & Andrea Krott
Center for Human Brain Health, School of Psychology, University of Birmingham, United Kingdom

Address for correspondence:
Vincent DeLuca
52 Pritchatts Road, Birmingham, B15 2TT, UK
v.deluca@bham.ac.uk
Abstract

A growing body of research shows that the brain adapts functionally and structurally to specific bilingual experiences. These brain adaptations seem related to modulations in cognitive processes (specifically the executive functions). However, the trajectory of these adaptations is varied and seems at least partially dependent on different aspects of language exposure and use. Here we provide a review of the existing theoretical models covering bilingualism-induced neuroplasticity. Moreover, we propose a unifying framework (Unifying the Bilingual Experience Trajectories, UBET) to more comprehensively map the relationship between the various neurocognitive adaptations and different aspects of bilingual experience trajectories, focusing on intensity and diversity of language use, language switching, relative proficiency, and duration of bilingual experience. Crucially, we also outline predictions regarding both relationships between different bilingual experience factors and relationships between the measurable neurocognitive adaptations. Our framework offers a theoretical backdrop and clear testable predictions for future large-scale empirical studies on individual differences in bilingual trajectories and their effects on neurocognitive adaptations.

Keywords

bilingualism; neuroplasticity; individual differences; executive functions
1 Introduction

In the past decade, there has been a stark increase in research examining the effects of bilingualism on the mind and brain. This research suggests that bilingualism modulates domain-general cognitive processes, specifically executive control (Bialystok, 2017; Bialystok, Craik, & Luk, 2012), likely through changes in brain structure and brain function (Li, Legault, & Litcofsky, 2014; Pliatsikas & Luk, 2016). However, inconsistencies exist between studies as to whether and how these effects manifest both at behavioral (Antoniou, 2019; Valian, 2015) and neural levels (García-Pentón, Fernández García, Costello, Duñabeitia, & Carreiras, 2016).

The arguments for no bilingual effects on neurocognition are predicated on the notion that if any adaptations (or ‘advantages’) do stem from bilingualism, then they should be robust across all groups and types of bilinguals when compared to monolingual controls (see e.g. Nichols, Wild, Stojanoski, Battista, & Owen, 2020). However, this line of argument is flawed. Given the complexity of the bilingual experience, it is likely that a range of cognitive and neural adaptations are variably required to adequately handle the demands associated with engaging with more than one language (Bialystok, 2017; Li et al., 2014). Examining bilinguals as a homogeneous group would, then, potentially conflate these experiences and modulate the measurable neurocognitive outcomes (see e.g. Bak, 2016; Bialystok, 2016; Poarch & Krott, 2019). Indeed, several models have been proposed, which attempt to both explain the existing variance within the data and offer predictions for future work examining individual differences in bilingualism and their neurocognitive effects (Abutalebi & Green, 2016; Calabria, Costa, Green, & Abutalebi, 2018; Grundy, Anderson, & Bialystok, 2017; Pliatsikas, 2019b; Stocco, Yamasaki, Natalenko, & Prat, 2014). These models have helped to delineate some of the variance across results in different studies and provide useful frameworks for the interpretation of future results. However, each of these models only explains certain aspects of both the bilingual experience and the range of associated neural and cognitive adaptations. In this paper, we propose a comprehensive framework to assess and unify the diverse neurocognitive outcomes caused by individual differences in bilingual experience trajectory and as such account for the divergent findings previously observed in the field. Table 1 clarifies some terms as we use them throughout this review.

1.1 Where the field stands: neurocognitive adaptation to bilingualism

The mechanisms underlying neurocognitive adaptation to bilingualism are thought to be a result of the way the brain handles and uses the languages at one’s disposal. These languages are thought to be constantly and jointly active, which in turn creates a state of competition for both language comprehension and production (Dijkstra & van Heuven, 2002; Green & Eckhardt, 1998; Marian & Spivey, 2003). For successful communication to occur, then, this conflict must be resolved (Abutalebi & Green, 2007; Bialystok, Craik, Green, & Gollan, 2009; Bialystok et al., 2012; Kroll, Dussias, Bogulski, & Kroff, 2012). The act of resolving the competition confers increased demands on several neurocognitive systems that are involved in language processing and control. The brain is thought to adapt both functionally and structurally to more effectively handle the several demands associated with bilingual language control (Li et al., 2014; Pliatsikas, 2019a). Crucially, the cognitive and neural processes underlying language control have been found to be at least partially shared with domain-general executive functions (Anderson, Chung-Fat-Yim, Bellana, Luk, & Bialystok, 2018; De Baene, Duyck, Brass, & Carreiras, 2015; Garbin et al., 2010). Adaptations to language control demands are thus thought to affect non-linguistic control processes.
Evidence for the notion of adaptation to language control demands has been found in structural plasticity of the brain (Burgaleta, Sanjuán, Ventura-Campos, Sebastián-Gallés, & Ávila, 2016; Mamiya, Richards, Coe, Eichler, & Kuhl, 2016; Mårtensson et al., 2012; Mechelli et al., 2004), neural activity modulation (Abutalebi et al., 2012; Anderson, Chung-Fat-Yim, et al., 2018; Ansaldo, Ghazi-Saidi, & Adrover-Roig, 2015; Bialystok et al., 2005; Dash, Berroir, Joanette, & Ansaldo, 2019; Kousaie & Phillips, 2017; Luk, Anderson, Craik, Grady, & Bialystok, 2010; Morales, Yudes, Gómez-Ariz, & Bajo, 2015; Mechelli et al., 2004), and behavioral task performance (Costa, Hernández, & Sebastián-Gallés, 2008; Morales, Gómez-Ariz, & Bajo, 2013; Stasenko, Matt, & Gollan, 2017). However, the results of these studies are variable pertaining to the nature and extent of neurocognitive plasticity. While the literature on brain structure and functional neuroimaging often reports apparent changes in the brain’s architecture brought about by bilingualism, the nature and location of adaptation and/or brain recruitment differs between studies (see for discussion García-Pentón et al., 2016).

Similarly, the literature on behavioral differences between monolingual and bilingual speakers does not consistently find effects. While many studies report differences (e.g. Costa et al., 2008; Prior & Macwhinney, 2010; Wiseheart, Viswanathan, & Bialystok, 2016; Zhou & Krott, 2018), others find partial or mixed effects (Coderre, van Heuven, & Conklin, 2013; Costa, Hernández, Costa-Fuidella, & Sebastián-Gallés, 2009; Hernández, Martin, Barceló, & Costa, 2013) or no significant effects at all (Antón, Carreiras, & Duñabeitia, 2019; Antón et al., 2014; Paap, Anders-Jefferson, Mason, Alvarado, & Zimiga, 2018; Paap, Johnson, & Sawi, 2015a). This has led to claims that there are no consistent neurocognitive effects of bilingualism at all (Paap et al., 2015a; Paap, Johnson, & Sawi, 2015b). However, there are several potential explanations for the inconsistency of results. First, the present trend of early studies finding predominantly positive (significant) results for bilingual experience with later work containing more negative effects is not uncommon in scientific discoveries (Leivada, Westergaard, Duñabeitia, & Rothman, in press.). Second, bilingualism is a very complex and dynamic experience with a range of potential factors that likely condition the nature of any neurocognitive adaptations (Bak, 2016; Luk & Bialystok, 2013).

A number of calls have been made to re-examine bilingualism, not as a categorical distinction but as a continuum of experiences that will variably affect adaptation (Bak, 2016; Bialystok, 2016; de Bruin, 2019; Li et al., 2014; Luk & Bialystok, 2013; Luk & Pliatsikas, 2016; Zirnstein, Bice, & Kroll, 2019). The bilingual experience trajectory comprises a range of factors relating to exposure to and use of one’s languages (Anderson, Mak, Keyvani Chahi, & Bialystok, 2018; Green & Abutalebi, 2015; Surrain & Luk, 2017). Individual experience-based factors require distinct specific neurocognitive processes to adequately handle their associated control and processing demands. The degree to which adaptations occur to these demands is thus governed by the nature and degree of experience (Bialystok, 2016; Luk & Bialystok, 2013). Consequently, examining individual differences in language experiences, and how these modulate neurocognitive adaptations, should help to explain some of the variability in neural and cognitive adaptation to bilingualism.

A growing body of evidence does show specificity of adaptations to individual bilingual experiences both from a cross-sectional (Beatty-Martínez et al., 2019; Bonfieni, Branigan, Pickering, & Sorace, 2019; DeLuca, Rothman, Bialystok, & Pliatsikas, 2019, 2020; Gullifer et al., 2018; Luo et al., 2019; Mamiya et al., 2016; Nichols & Joanisse, 2016; Ooi, Goh, Sorace, & Bak, 2018; Pliatsikas, DeLuca, Moschopoulou,
and longitudinal perspective (DeLuca, Rothman, & Pliatsikas, 2018; Grant, Fang, & Li, 2015; Hervais-Adelman, Moser-Mercer, & Golestani, 2015; Hofstetter, Friedmann, & Assaf, 2016; Hosoda, Tanaka, Nariai, Honda, & Hanakawa, 2013; Legault, Fang, Lan, & Li, 2019; Mårtensson et al., 2012; Mohades et al., 2015; Schlegel, Rudelson, & Tse, 2012). Findings between studies here are more consistent. For example, research examining prolonged duration of bilingual experience shows neural and cognitive adaptations towards increased efficiency in language control. This has been found to manifest as a general shift from reliance on task-relevant cortical regions towards subcortical structures or posterior regions such as the cerebellum, as well as decreased white matter structure in frontal regions (e.g. DeLuca et al., 2018; Nichols & Joanisse, 2016; Pliatsikas et al., 2017).

1.2 Measuring neurocognitive adaptations in bilinguals

Neuroimaging allows researchers the opportunity to investigate a range of potential neurocognitive effects of the bilingual experience, both at the structural and functional level (Luk, Pliatsikas, & Rossi, 2020). Tables 2 clarifies some brain structure abbreviations as we use them throughout this review.

Structural magnetic resonance imaging (sMRI) affords the opportunity to quantify patterns in regional grey matter- and white matter structure. Patterns of grey matter structure can be quantified in a variety of ways including grey matter volume or density (Ashburner & Friston, 2000), cortical thickness (Ad-Dab’bagh et al., 2005), and surface area as measured by vertex analyses (Patenaude, Smith, Kennedy, & Jenkinson, 2011). Bilingual experience has previously been found to be associated with changes in patterns of grey matter structure across various cortical and subcortical brain regions implicated in language and cognitive control processes (see for review Li et al., 2014; Pliatsikas, 2019a).

Diffusion tensor imaging (DTI) captures the diffusivity of water molecules in tissue, which serves as a powerful way to qualitatively and quantitatively investigate microstructural characteristics of the white-matter tracts in the brain (Le Bihan, 2003). Some of the more common measures stemming from this imaging technique relate to the degree of diffusivity in general (mean diffusivity; MD), diffusivity along a tract (axial diffusivity; AD), perpendicular to a tract (radial diffusivity; RD), and fractional anisotropy (FA) which is a scalar value computed from the ratio of AD and RD values. Each of these values is thought to relate to specific microstructural properties of white matter including fiber density, axonal diameter, and degree of myelination (Mori & Zhang, 2006; Smith et al., 2006). Bilinguals have been found to diverge from monolinguals in terms of diffusivity patterns across a range of tracts (e.g. Anderson, Grundy, et al., 2018; Hämäläinen, Sairanen, Leminen, & Lehtonen, 2017; Luk, Bialystok, Craik, & Grady, 2011; Mohades et al., 2012; Pliatsikas, Moschopoulou, & Saddy, 2015; Rossi et al., 2017; Singh et al., 2017).

In addition to brain structure, patterns in brain activation and functional connectivity, both at rest and in conjunction with specific stimuli or responses, can be inferred using functional magnetic resonance imaging (fMRI). fMRI measures where changes in oxygenated blood flow (the BOLD signal) are occurring in the brain, both at rest and during a task. fMRI is highly useful for assessing the location(s) of neural activity; however, its temporal resolution is limited, with the BOLD response to simple experimental stimuli occurring seconds after its onset. Using fMRI, studies have found monolinguals and bilinguals to exhibit diverging patterns of brain recruitment to handle given executive control demands (see for review Pliatsikas
 Specifically, bilinguals have either been found to recruit alternative or additional regions and networks (Anderson, Chung-Fat-Yim, et al., 2018; Ansaldo et al., 2015; Luk et al., 2010) or vary in terms of the degree of activation in regions or networks (e.g. Abutalebi et al., 2012; Costumero, Rodríguez-Pujadas, Fuentes-Claramonte, & Ávila, 2015) to handle the same cognitive demands.

A way of examining the time course of brain activity with the millisecond temporal resolution at which neurons operate is through the electroencephalogram (EEG), examining either evoked components or induced changes in oscillatory activity. With the evoked component approach, an averaging procedure is needed to extract the “event” related signal called the event related potential (ERP) which represents the neural activity precisely time- and phase-locked in response an experimental event. The peaks and troughs in an ERP waveform often follow a stereotypical temporal pattern of positive and negative voltage deflections, classified as ‘components’ (Luck, 2005) which are assumed to map onto various cognitive constructs. One component often modulated by the experience of bilingualism is the N2 component, a negative peak occurring around 200 ms after stimulus onset, with a maximal topography over frontal-central electrode sites (when using an average reference montage). The N2 is involved in the detection of potentially conflicting information (e.g. is the stimulus an infrequently presented target or commonly occurring non-target) and inhibition of prepotent responses. The earlier the latency of the N2 and the larger its amplitude, the more readily the conflict detection process is thought to have occurred. The P300 being most prominent over parietal electrodes is a positive potential occurring 300-500 ms after the occurrence of a novel/rare/conflicting stimulus (e.g. the letter O occurring after a series of X’s). Previous research has suggested that the peak latency of the P300 reflects the time it took for the participant to automatically classify the stimulus as novel, and its amplitude reflecting the confidence of the classification. Finally, the error related negativity (ERN), a negative response with a scalp topography over frontal central electrodes occurring 50 to 100 ms after an erroneous response, even without the explicit awareness of the participant. It has been speculated to reflect an online conflict monitoring system. Previous research has found bilingual experience to interact with the latency, amplitude and scalp distributions of waveforms during executive function tasks (see for review Grundy et al., 2017). Bilinguals have been found to exhibit larger and/or earlier N2 and P300 components than monolinguals in inhibition-related tasks such as the flanker or Simon task (e.g. Kousaie & Phillips, 2012, 2017; Sullivan, Janus, Moreno, Astheimer, & Bialystok, 2014). Other work has found scalp distribution of these components to become more broad in bilinguals (Timmer et al., 2017). In tasks tapping into aspects of proactive and reactive control (such as the AX-CPT), bilinguals have been found to exhibit ERP patterns associated with greater reliance on reactive control processes than monolinguals (greater N2/P300 amplitudes but only for reactive control condition) (e.g. Morales et al., 2015).

In contrast to the event-related averaging approach, the oscillatory changes in the EEG reflect activity that while time-locked to an experimental event, is not phase-locked. EEG activity contains rhythmic activity at different characteristic frequencies called frequency bands, which are believed to be produced by specific areas of the brain and related to specific cognitive processes (Mazaheri, Slagter, Thut, & Foxe, 2018; Siegel, Donner, & Engel, 2012). One important rhythm, particularly relevant to control processes, is the theta rhythm (3-7), with a scalp topography over frontal-midline sites. The theta rhythm has primarily been observed to increase in amplitude during the implementation of executive processes (Nigbur, Ivanova, & Stürmer, 2011). Activity at the frequency range of 8-14 Hz, referred to as the alpha rhythm, has primarily been observed over sensory regions, such as the visual, sensory-motor, and auditory cortices (see recent review by Van Diepen, Foxe, & Mazaheri, 2019). The increase in the amplitude alpha activity over a brain
area is believed to signal its inhibition or idling. The prevalent view on the functional role of the alpha rhythm during cognition is that it is involved in gating information processing away from task-irrelevant regions, to the task relevant (Van Diepen et al., 2019). This technique has been used less frequently within the bilingualism literature to measure brain activity (but see Rossi & Prystauka, 2020), but we will elaborate on the potential of EEG for future research later in this review.

1.3 Models of neurocognitive adaptation to bilingualism

Several models have been proposed which outline different facets of bilingual experience and the related neural and cognitive adaptations to them. We focus here on recent models (Fig 1) that examine different aspects of the bilingual experience (e.g. duration of L2 use, degree/nature of switching, etc.) and that provide predictions about how the brain adapts to these experiences. The models discussed herein are, in turn, based on earlier models that also covered language-related neuroplasticity, such as the bilingual interactive activation model (van Heuven, Dijkstra, & Grainger, 1998), inhibitory control model (Green & Eckhardt, 1998), and declarative procedural model (Paradis, 2009; Ullman, 2004). Thus, while these earlier models are not directly reviewed, we do build on their predictions indirectly. It is also worth noting that the models reviewed here primarily discuss language production as a primary driver of adaptation, but do not explicitly distinguish this from comprehension. Adaptations to experience-based factors involve both the variable enhancement of existing structures and networks with changing environmental demands, and repurposing of existing modules/regions to new processes with continued bilingual language exposure (Hernandez et al., 2019; Rodriguez, Archila-Suerte, Vaughn, Chiarello, & Hernandez, 2018).
The Adaptive Control Hypothesis (ACH) (Abutalebi & Green, 2016; Calabria et al., 2018; Green & Abutalebi, 2013), discusses adaptation to bilingual experience trajectory from a perspective of intensity of engagement with one’s languages. Specifically, the ACH proposes that several cognitive processes are required for successful language control, including goal maintenance, response inhibition, conflict monitoring, interference suppression, salient cue detection, task engagement and disengagement, and opportunistic planning. These processes are variably required, depending on the conversational context.

There are three general categories of communicative context. First, the single language context captures the use of only one of the available languages in a specific environment. This context is assumed to only engage goal maintenance and is therefore thought to require only increased input from the inferior frontal gyrus (IFG). Second, the dual language context captures situations in which both languages are used, but with different interlocutors. This context is argued to require several control processes, namely goal maintenance, interference control, salient cue detection, selective response inhibition, and task engagement/disengagement. Due to this complexity, it engages a wide network of brain regions (IFG (bilateral), anterior cingulate cortex (ACC), inferior parietal lobule (IPL), caudate, putamen, thalamus, and cerebellum). Third, the dense code-switching context captures situations in which both languages are used with the same interlocutor, with frequent language switches, including those within utterances. This type of language usage is less taxing on control processes. It requires only opportunistic planning and relatively less inhibition. It is thought to use a control network of the left IFG and cerebellum. Individuals can engage several (or all) of these interactional contexts and may even shift between them on a regular basis. However, it is argued that increased engagement with a specific context reinforces (by necessity) the cognitive processes required by it and their underlying networks. These network reinforcements, then, are argued to result in increased performance on tasks that tap into the associated cognitive processes and increased
plasticity and connectivity in the relevant brain regions associated with them.

The **Conditional Routing Model (CRM)** (Stocco, Lebiere, & Anderson, 2010; Stocco et al., 2014) proposes a more functional series of adaptations to bilingual language control and processing, with a special focus on the basal ganglia (in particular the striatum). This model is based on the notion of the basal ganglia functioning as a gate of neural signals to prefrontal and other cortical regions (Stocco et al., 2010). It is argued that the basal ganglia (specifically, the caudate and putamen) has the capacity to select rules in response to specific, prerequisite conditions and to override habitual responses encoded in cortico-cortical connections. Bilingualism effectively trains this capacity because the acquisition and use of an additional language means frequent selection and switches between rules and representations. Training the system gives increased efficiency in rule acquisition, selection, and application. Since all these processes are implicated in general executive function processes, bilingualism affects performance in non-linguistic executive control tasks that rely on rule selection, switching and top-down control to maintain a rule behavior in the presence of distracting information.

The **Bilingual Anterior to Posterior and Subcortical Shift (BAPSS)** framework (Grundy et al., 2017) proposes a series of neural adaptations to prolonged experience with the additional language. The BAPSS framework specifically proposes that early stages of L2 (second language) use require increased reliance on frontal cortical regions (e.g. ACC and IFG), due to increased demands on both language and executive control processes, which are predominantly served by these regions. With prolonged L2 use, neural reliance shifts from frontal regions to subcortical and posterior regions (e.g. basal ganglia, thalamus, occipital lobes) commensurate with more automated and efficient language control and processing. The BAPSS framework also makes predictions regarding neural activity. With respect to electrophysiology, in relation to stimuli probing or requiring switching and inhibitory control prolonged duration of L2 use will correlate with both earlier latencies and increased amplitudes in the N2/P300 components reflecting earlier and more automatic detection of conflicting input, and more confidence in its classification. Decreased error-related negativity (ERN) amplitudes, are also predicted with prolonged experience, indicative of lower resources being allocated at this point, given greater resources being allocated temporally earlier (i.e. at the N2). The framework also discusses these neural adaptations to bilingualism as a mechanism for counteracting the anterior shift in neural reliance associated with ageing (Grant, Dennis, & Li, 2014).

Finally, the **Dynamic Restructuring Model (DRM)** (Pliatsikas, 2019b) proposes a system of stages in neuroanatomical adaptations to the duration of L2 exposure and usage and the varying cognitive demands resulting from it. Three stages of adaptation are proposed. The first stage, *initial exposure*, confers early adaptations geared towards language learning and early control demands. Adaptations to these new language control demands result in increased reliance on cortical structures implicated in cognitive control and short-term memory, including the IFG, ACC, inferior parietal lobule/superior parietal lobule (IPL/SPL), and hippocampus. The second stage, *consolidation*, occurs with increased L2 experience and is characterized by shifts towards more efficient control of the two languages. This results in increased structural connectivity between regions, impacting on several white matter tracts including the inferior fronto-occipital fasciculus (IFOF), anterior thalamic radiation (ATR) and superior longitudinal fasciculus (SLF), increases in subcortical grey matter (e.g. caudate, putamen, and thalamus), and a ‘renormalization’ or return to baseline levels of cortical grey matter. The final stage, *peak efficiency*, deals with adaptation towards automation in L2 processing and control. This results in increases in cerebellar grey matter and decreases in frontal connectivity (e.g. within forceps minor), as reliance shifts to posterior regions.
1.4 Connecting predictions of different existing models

What the reviewed models have in common is that they are based on the notion of neurocognitive adaptations to be maximally effective in language control, either by adapting to increased specific cognitive demands or by adapting towards increased efficiency of handling existing demands. The models also largely converge on the regions and networks implicated within these notions (Fig 2). The subcortical structures (specifically the basal ganglia & thalamus) form a crucial part of the language control network within each model. They are most often implicated across each model as relating to language selection and control. Fronto-parietal regions such as the (bilateral) IFG, ACC, and supramarginal gyrus/inferior parietal lobule (SMG/IPL) are related to increased language control and language processing demands. Posterior regions such as the cerebellum (DRM) and occipital lobes (BAPSS) are often implicated in processes related to increased efficiency in language control. The nature and extent of the reliance on and subsequent adaptations in these regions and pathways differ in latency and intensity, depending on the model.

As summarized above, the models examine different aspects of bilingual experience trajectory, and thus varying neural outcomes are predicted by each model, either in terms of latency or distribution. But none of the models appear to capture all the factors that have been proposed to modulate neural adaptations. The ACH makes predictions regarding adaptations towards the intensity of engagement with the L2 in different conversational settings, or at least to the extent that use of more than one language is supported in that setting. However, the model does not, for instance, discuss whether or how these networks and regions are modulated by overall duration of bilingual experience. It is possible that prolonged engagement in a specific language setting may mean that the regions of the associated network might not be equally affected, but that certain regions within the network are more strongly affected. This is particularly the case in a dual-
language setting where the widest control network is engaged. Specifically, with prolonged exposure to a
dual language setting the brain becomes more efficient in handling the associated cognitive demands.
Neurobiologically, this would manifest as increased reliance on the cerebellum and subcortical structures
(specifically caudate/putamen) within the control network, with less reliance on cortico-frontal structures
(ACC/IFG). In contrast to the ACH, the BAPSS framework and the DRM make predictions regarding the
adaptations over time, but do not specify how other experience-based factors, particularly the intensity of
L2 engagement, might affect the rate at which these adaptations occur. Finally, the CRM makes reference
to increased proficiency in L2 use (increased role of the basal ganglia in routing of information and
selection), but the latency by which this occurs is not explicitly predicted, nor the degree to which intensity
of L2 exposure might affect this latency.

2 UBET: A broader view on adaptations to the bilingual experience

In what follows, we will put forward a unified comprehensive framework, which we refer to as
Unifying the Bilingual Experience Trajectories (UBET), on the relationship between the bilingual
experience trajectory and neurocognitive adaptations. In this, we are guided by the overlap of the existing
models but also make some new, testable predictions. As illustrated in Figure 3, a variety of neurocognitive
adaptations are predicted depending on a range of individual experience-based factors. More specifically,
some experience-based factors are predicted to primarily lead to individual differences in control demands
while others to individual differences in automaticity/efficiency (Fig 3). Crucially, though, we propose that
these experiences will modulate the effects of one another in terms of how the brain adapts. The specific
experience-based factors, their neurocognitive effects, and (modulatory) relationships are discussed in the
following section. Regarding the experience-based factors that give rise to these neurocognitive
adaptations, there is a range of specific language experiences that might variably translate to different neural
adaptations. Herein, we grouped these under four general categories: 1) intensity and diversity of L2 use,
2) language switching, 3) relative proficiency, and 4) duration of bilingual use (Fig 3).
Intensity and diversity of language use, taken together, refer to the extent and variety of situations in which one is exposed to and uses an additional language (Grosjean, 2016). Diversity of language use here stems from the notion of entropy as operationalized by Gullifer and Titone (2020), specifically the degree of compartmentalization or integration of one’s languages within and across specific contexts. Intensity of use refers to the extent to which an additional language is at least available or used in daily life, overall. Although intensity and diversity of language use are different aspects, they are combined here as they will likely converge with respect to the nature of the demands placed on language control and processing faculties. Specifically, both greater intensity and diversity of language use are predicted to be associated with increased control demands (Beatty-Martínez et al., 2019; Green & Abutalebi, 2013).

Language Switching—herein is defined as a general term encompassing the extent to which one switches between- and/or mixes their languages on a regular basis. We discuss language switching using the typology proposed by Muysken (2000), and the control demands associated (Green & Wei, 2014; Hofweber, Marinis, & Treffers-Daller, 2016). The different types of switching are argued to exist on a spectrum including single language use, alternation between languages, insertion, and dense code switching (Hofweber et al., 2016). The associated control demands range from maximal interference suppression in alternation or inter-utterance switches to minimal requirements of interference suppression in dense code-switching (Hofweber et al., 2016; Treffers-Daller, 2009). Engagement with a specific type of switching (as defined above) will necessitate commensurate neurocognitive adaptations (Adler, Valdés Kroff, & Novick, 2020; Hofweber et al., 2016). Cognitive demands associated with monitoring changing conflict, however, follow the reverse pattern: dense code switching requiring the most monitoring and single language use requiring the least

![Diagram](image-url)
Conflict monitoring (Hofweber et al., 2016). Despite requiring potentially differential engagement of control processes, neurocognitive adaptations to language switching are herein proposed to fall under control demands (discussed further in section 2.1).

Relative proficiency refers generally to the balance of fluency and comfort/confidence in using the languages at one’s disposal. Relative proficiency herein is operationalized as being formed of two subcomponents: language proficiency and dominance. Proficiency herein is defined as the extent to which the representation fits a target (native-like) state, typically as measured by a standardized test. Of the two subcomponents, it is possible that proficiency (past a certain point of attainment) does not necessarily affect control processes and indeed might be considered an outcome measure in its own right (see for discussion DeLuca et al., 2018). However, within early stages of acquisition, previous work has shown higher proficiency levels to relate to increased degrees of neural plasticity (Mamiya et al., 2016; Mårtensson et al., 2012). Although to some degree related to proficiency, here we define language dominance as the comfort with using and/or ease of access to a given representation (Bedore et al., 2012). Dominance is a subject-internal relative measure, whereas proficiency (as defined above) in each language is not necessarily dependent on the other. In terms of neurocognitive adaptation, more balanced relative proficiency should lead to increased efficiency of language control (Stocco et al., 2014; Yamasaki et al., 2018), but this is likely dependent on several of the other experience-based factors (see section 2.3) (Beatty-Martínez et al., 2019; Grundy et al., 2017; Gullifer et al., 2018). While fully balanced relative proficiency is rare, a minimum threshold should likely be met for both maintenance of the linguistic representations and control demands to be continuously necessitated (Iverson & Miller, 2017; Miller & Rothman, 2019).

Finally, duration of L2 use refers to the overall length of time one engages with more than one language. Neurocognitive adaptations to longer duration of experience is predicted to lead to increased efficiency of language control and processing (Grundy et al., 2017; Pliatsikas, 2019b).

2.1 Neurocognitive adaptations

The neurocognitive adaptations to the experience-based factors described above can be grouped into two general categories: 1) adaptations due to changes in executive control demands, and 2) adaptations due to changes in efficiency.

As described above, changes in executive control demands arise from various demands on language use and control, such as increased intensity and diversity of language use and increased language switching. A remarkable change in control demands is expected when L2 is used in novel situations or patterns, e.g. from initial exposure to an L2, or when the degree of intensity of L2 usage dramatically increases or decreases. Cognitive processes relevant to language control become increasingly stressed with higher language control demands, and the brain adapts accordingly. Given previous findings, we predict the following neurocognitive adaptations to occur as a result of changes in control demands. Neural correlates of control demands include increased recruitment of cortical regions involved in control processes (e.g. the IFG, ACC, and IPL). Structurally, these regions are predicted to exhibit greater grey matter volume to handle the additional control demand (Calabria et al., 2018; Pliatsikas, 2019b). Increased control demands have also been shown to manifest as a general reliance on proactive control strategies (Gullifer et al., 2018), which might be evident in the performance in non-verbal behavioral tasks, such as smaller mixing/switching costs in switching tasks and superior suppression of interfering information/stimuli in selective attention tasks. Increased language control demands should reinforce the use of these networks in non-linguistic contexts.
Consequently, the increased control demands should likely also manifest as increased stimulus-related theta band power over mid-frontal scalp regions, increased alpha band power over task-irrelevant sensory cortices and suppression of alpha band power over task-relevant regions. Theta band activation has previously been linked with several processes including conflict monitoring and interference suppression (Janssens, De Loof, Boehler, Pourtois, & Verguts, 2018; Nigbur et al., 2011). The increase in alpha activity in task-irrelevant region has been associated with functional inhibition of distracting and efficient allocation of resources (see Van Diepen et al., 2019).

**Efficiency** arises from becoming accustomed to existing language control demands. In this instance, cognitive economy dictates adaptations that allow for more efficient and automated cognitive processes to handle these demands. These adaptations likely only occur once the language user is maximally effective at handling the cognitive load associated with the control demands of the (overall) language environment. Thus, prolonged duration of L2 use (a scale of years) and increased L2 proficiency will confer adaptations that group under efficiency. We predict the following neurocognitive adaptations to occur as a result of changes in efficiency. Neural correlates of efficiency (as discussed by the BAPSS framework, CRM and DRM) include preferential recruitment of subcortical structures (specifically caudate, putamen, and thalamus) (Stocco et al., 2014) and posterior regions such as the occipital lobes and cerebellum (Grundy et al., 2017; Pliatsikas, 2019b). Increasing automaticity will be associated with prior cortical grey matter increases reverting towards near-baseline levels (Grundy et al., 2017; Lövdén, Wenger, Mårtensson, Lindenberger, & Bäckman, 2013; Pliatsikas, 2019b), with increases then occurring in subcortical and posterior structures (Grundy et al., 2017; Pliatsikas, 2019b). Regarding ERP components related to attentional/executive control, prolonged L2 exposure should result in earlier latency and increased amplitude for components such as the N2 and P300 (Grundy et al., 2017). Here as well, it is likely that increased L2 exposure during this timeframe would shorten the latency by which this transition occurs. Increases in stimulus-related alpha power in task-irrelevant cortical regions would also indicate a shift towards efficiency, specifically in the allocation of attentional resources. As explained above, alpha is argued to be an index of the gating of activity in non-relevant brain regions for specific tasks, including interference suppression and the biasing of attentional states (Foxe, Simpson, & Ahlfors, 1998). Indeed, previous work has shown increased alpha band activation to be linked to inhibitory processes in tasks requiring conflict monitoring and interference suppression (Janssens et al., 2018). Finally, cognitively, increased efficiency is predicted to manifest as a reliance on reactive control processes (Gullifer et al., 2018).

### 2.2 Relationships between different neurocognitive outcomes

Patterns of neurocognitive adaptation are likely interrelated, although this has not been explicitly discussed in previous models. Here we discuss several predictions of what relationships between brain structure, function, and cognition exist, as modulated by bilingual experience-based factors.
We predict that brain function patterns will follow similar patterns of adaptation to structural adaptations with initial exposure to an L2. In other words, the initial stages of contact with the L2 will result in increased functional recruitment of mainly fronto-cortical regions to handle the increased load. We also predict this increase in functional recruitment will be associated with increases in grey matter in these regions. Increased efficiency and automation lead to a reduction in executive control demands, meaning that functional recruitment patterns in fronto-cortical regions will begin to revert towards a baseline level (Grundy et al., 2017). In line with predictions of the DRM, we predict that this reversal will be associated with decreased grey matter within these regions (Pliatsikas, 2019b) and as result with a decreased theta power induced by executive control demands (Fig 4).

Regarding the relationship between adaptations in stimulus-related oscillatory dynamics and brain structure, an increase in alpha power over sensory regions in situations requiring executive control could indicate a reliance on subcortical structures and the cerebellum for control processes (see Mazaheri et al., 2009 for discussion) (Fig 4). Thus, we expect an increase in alpha power over cortical regions in tasks requiring interference suppression as task demands become more reliant on subcortical structures (particularly the basal ganglia) (Mazzetti et al., 2019; Talakoub et al., 2016) and reduced cortical requirements to maintain successful language control. Interestingly, a recent study observed a relationship between the inter-individual variability among participants’ abilities to modulate alpha activity over the visual cortex, and volume of the globus pallidus, a basal ganglia structure (Helfrich et al., 2017). Previous research has suggested that alpha activity over the sensory-cortices to be partially controlled by sources in the prefrontal cortex (Helfrich, Huang, Wilson, & Knight, 2017; Mazaheri et al., 2010, 2014; Sauseng, Feldheim, Freunberger, & Hummel, 2011). We speculate that with prolonged bilingual experience, a shift would occur from frontal cortical structures to more subcortical regions and tracts connecting these, modulating alpha activity (Fig 4).
What do these changes mean for cognitive (behavioral) outcomes pertaining to executive control (reaction times and accuracy)? We argue that those will improve in early stages of bilingualism and then plateau at the level of peak efficiency. Specifically, upon peak performance being reached (in terms of behavioral responses), this performance would be maintained with an increase in alpha over task-relevant and irrelevant cortices, indicating a further shift towards automated and more efficient processing. Regardless, maintenance of peak efficiency will drive the specific trajectory of adaptations. That is, cognitive adaptations are likely to follow neural adaptations—only arriving at what is required to handle the competing systems. After this, the adaptation trajectory shifts to decreasing (neuro)cognitive requirements for maintaining this level of performance. Therefore, it is unlikely that, once peak performance has been reached, changes in (neural) processing speed would manifest any differently on behavioral measures in executive function tasks such as reaction times. However, as has been argued previously, the degree to which task performance is affected likely also depends on the nature of the task used (see for discussion Bialystok, 2016; Valian, 2015). Employing tasks which require a greater degree of attentional control, such as the attentional network task (ANT), or manipulating (typically increasing) the ratio of congruent to incongruent trials (e.g. Costa et al., 2009) is more likely to elicit behavioral effects stemming from bilingual experience (Bialystok, 2017; Zhou & Krott, 2018). Similarly, tasks which isolate the contributory aspects of executive control, such as the AX-continuous performance task (AX-CPT) which is thought to isolate proactive and reactive control, will likely better capture the behavioral correlates of neurocognitive adaptation to specific experience-based factors (see e.g. Gullifer et al., 2018). For tasks better suited to directly measure differences between proactive and reactive control, we predict increased proactive control to be associated with greater frontal activation and theta band power (Fig 4). Shifts towards more reliance on reactive control for control demands would be associated with a suppression of alpha power over task relevant cortical regions, and recruitment of subcortical/posterior structures (Grundy et al., 2017) (Fig 4).

2.3 Relationships between experience-based factors making up the bilingual experience trajectory

Crucially, while the above experience-based factors are expected to relate to specific neural outcomes, they do not occur in isolation. We therefore need to consider interactions of the factors and what these interactions mean with regards to neurocognitive adaptations (Sulpizio, Del Maschio, Del Mauro, et al., 2020). While it is possible to formulate predictions for the combinations of some of our experience-based factors, others have either not been sufficiently studied or we do not have sufficient theoretical reason to relate them. Furthermore, the socio-linguistic environment in which a bilingual speaker is situated should determine the combination and subsequent neurocognitive effects of these experience-based factors. In what follows, we will describe two likely interactions between our experience-based factors, specifically how they might modulate their respective neurocognitive adaptations.

First, we predict that increased intensity and/or diversity of language use will shorten the latency by which adaptations towards efficiency and automation occur with duration of use (Fig 3). Prolonged L2 immersion as an experience-based factor provides a key example of this. Regarding the role of language use, immersion in non-native language settings can reliably be assumed to at least intensify exposure to the L2 and increase the opportunities for use within a given timeframe (Linck, Kroll, & Sunderman, 2009). As such, the language control system optimizes to handle the additional language control load, in the above cases towards maximizing efficiency of language processing and control. Previous research has found that highly immersed late-acquired L2 bilinguals and early-L2 acquired bilinguals exhibit very similar brain structure adaptations towards increased efficiency and automation (as predicted by BAPSS and DRM), relative to...
functional monolingual control groups. Specifically, in the subcortical structures, immersed bilingual groups showed increased surface displacement in the putamen, thalamus and globus pallidus (Pliatsikas et al., 2017) relative to monolingual controls, a restructuring pattern that highly overlaps with early simultaneous bilinguals not experiencing L2 immersion (Burgaleta et al., 2016). A similar pattern has been found for white matter tracts. A study by Pliatsikas and colleagues found increased white matter myelination in the corpus callosum, IFOF, and SLF for immersed bilinguals (Pliatsikas et al., 2015). Here again, this pattern overlaps with early-acquired bilinguals who were not living in L2 immersion settings (García-Pentón, Perez Fernandez, Iturria-Medina, Gillon-Dowens, & Carreiras, 2014). The tracts and structures reported in the above studies have been implicated in duration-based models (BAPSS and DRM) as being increasingly used as language control becomes more automated. Such effects, then, can be interpreted as increased intensity of L2 use shortening the latency by which adaptations towards increased efficiency/automation of language control occur. Given the current evidence, the UBET framework would predict a shift towards reliance on reactive control processes (e.g. Gullifer et al., 2018) with prolonged L2 immersion.

Neurocognitive adaptations to degree of relative proficiency are likely to be conditioned by both the opportunities for use (e.g. compartmentalization of languages, Beatty-Martínez et al., 2019; Gullifer et al., 2018) and the duration of engagement (Abutalebi & Green, 2016; Grundy et al., 2017; Pliatsikas, 2019b). As specified earlier, L2 proficiency is likely conditioned to both duration of L2 use and intensity/diversity of language use (Fig 3), particularly past a specific point of attainment. Furthermore, maintenance of the L1 during this timeframe will alter the latency and nature of adaptation towards efficiency. Based on existing theory (Calabria et al., 2018; Pliatsikas, 2019b) we argue that prolonged L1 maintenance would maintain the neural adaptations in language efficiency (increased subcortical and cerebellar use). Decreased L1 maintenance would require renewed executive control demands (fronto-cortical recruitment), specifically in situations where the L1 is required or used (Iverson & Miller, 2017).

3 Effects of the bilingual experience trajectory on adaptations in language functioning

Although a full discussion of it is beyond the scope of this paper, it is worth noting that bilingualism and its component experiences also affect language (Fricke, Zirnstein, Navarro-Torres, & Kroll, 2019; Kroll, Dussias, & Bajo, 2018). Some examples of this relationship are briefly discussed in this section. Regarding the representation of syntax in bilinguals, different theories have proposed that syntactic representations are separate (Amaral & Roeper, 2014; de Bot, 1992) or shared (Hartsuiker, Pickering, & Veltkamp, 2004). Evidence suggests that bilingual experience trajectories may affect the degree to which syntactic representations are shared between two languages, with a specific role of proficiency (e.g. Bernolet, Hartsuiker, & Pickering, 2013; Hartsuiker & Bernolet, 2017; van Heuven & Dijkstra, 2010). A more recent meta-analysis shows overlapping neural networks and structures are used for processing the L1 and L2, although this depends on the level examined (e.g. phonology, lexico-semantic, syntax) (Sulpizio, Del Maschio, Fedeli, & Abutalebi, 2020). Sulpizio and colleagues (2020) also note that few fMRI studies to date have examined whether a modulatory role exists for bilingual experience trajectories in language representation and processing. In the ERP literature, some evidence exists for a specificity of effects of such experience trajectories on language functioning. For example, increased prior code-switching experience has also been to impact the neural correlates of grammatical gender and decreased switch costs in production (Beatty-Martínez & Dussias, 2017). Regarding processes underlying language acquisition, increased passive exposure to an additional language has been found to increase sensitivity (at a neural level) to a vowel harmony rule in a novel language in monolingual adults (Bice & Kroll, 2019). Regarding
the connection between language control and domain-general cognitive control, research suggests that overlapping regions and networks are recruited to handle language- and cognitive control demands, respectively, in bilinguals (e.g. Anderson, Chung-Fat-Yim, et al., 2018; De Baene et al., 2015; Sulpizio, Del Maschio, Fedeli, et al., 2020). However, our understanding of a potential modulatory role of bilingual experience trajectories on the relationship between language- and domain-general cognitive control remains limited. Future work might also incorporate more combinations of linguistic- and nonlinguistic measures to comprehensively assess neurocognitive impacts of various bilingual experience trajectories in bilingualism.

4 Concluding remarks

Bilingualism is a complex experience which comprises a range of experience-based factors. Equally, there are several ways in which the brain adapts to these experiences. However, both the adaptations and experience-based factors which give rise to them do not occur independently, and thus likely modulate their respective neural effects. Herein, we have outlined a framework mapping these relationships between the bilingual experience trajectory and neurocognitive outcomes (UBET). As outlined, the component experience-based factors can be broadly summarized into four categories: intensity & diversity of use, language switching, relative proficiency, and duration. Adaptations to these experiences can be broadly summarized into two general domains: efficiency and control demands. The primary effect of longer duration of use and more balanced language proficiency will be adaptations towards increased efficiency, whereas an increase in diversity/intensity of use and frequent controlled language switching will primarily increase control demands. Depending on the nature of the socio-linguistic environment in which a bilingual speaker is situated, these experiences will occur in varying proportions and will modulate their respective effects. Increased intensity of use should shorten the latency by which duration of use causes adaptations towards increased efficiency. Similarly, language switching, particularly inter-sentential and inter-personal, would enhance the adaptations towards increased control demands brought on by diversity and/or intensity of L2 use.

The field of bilingualism research has taken the direction toward examining individual differences in language experience and this is encouraging. However, we believe that the field now needs to come together to relate the range of neurocognitive adaptations and experience-based factors which give rise to them. Herein we have proposed a framework of bilingual adaptation in which multiple (possibly interrelated) neurocognitive outcomes occur to ensure maximal efficacy, and then efficiency, in language control and processing. By considering individual language experience and their related adaptations together, we can more comprehensively map the dynamic nature of bilingualism-induced neurocognitive adaptations.
References


Cognition.


No General Cognitive Advantages: A Population Study of Executive Function in 11,000 People. 


Table 1: Operationalization of key terms used within this review.

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>bilingualism</td>
<td>the use of more than one language</td>
</tr>
<tr>
<td>L2</td>
<td>one's second language. Herein, this is interpreted more generally as the</td>
</tr>
<tr>
<td></td>
<td>'additional' language, to include simultaneous bilingualism.</td>
</tr>
<tr>
<td>bilingual experience</td>
<td>the combination of all relevant factors pertaining to bilingual language</td>
</tr>
<tr>
<td>trajectory</td>
<td>exposure and use</td>
</tr>
<tr>
<td>experience-based factor</td>
<td>the combination and timecourse of several (relevant) experience-based</td>
</tr>
<tr>
<td></td>
<td>factors pertaining to the use of more than one language</td>
</tr>
<tr>
<td>duration of use</td>
<td>a general aspect of bilingual experience (e.g. duration of use)</td>
</tr>
<tr>
<td>intensity/diversity of use</td>
<td>overall length of time one engages with more than one language</td>
</tr>
<tr>
<td>language switching</td>
<td>extent and variety of situations in which one is exposed to and uses an</td>
</tr>
<tr>
<td></td>
<td>additional language</td>
</tr>
<tr>
<td>relative proficiency</td>
<td>extent to which one switches between- and/or mixes their languages on a</td>
</tr>
<tr>
<td></td>
<td>regular basis</td>
</tr>
<tr>
<td></td>
<td>balance of fluency and comfort/confidence in using the languages at one’s</td>
</tr>
<tr>
<td></td>
<td>disposal</td>
</tr>
</tbody>
</table>

Table 2: Abbreviations of brain regions and tracts

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Structure/tract</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACC</td>
<td>Anterior cingulate cortex</td>
</tr>
<tr>
<td>ATR</td>
<td>Anterior thalamic radiation</td>
</tr>
<tr>
<td>BG</td>
<td>Basal Ganglia</td>
</tr>
<tr>
<td>CC</td>
<td>Corpus callosum</td>
</tr>
<tr>
<td>Cereb</td>
<td>Cerebellum</td>
</tr>
<tr>
<td>Hipp</td>
<td>Hippocampus</td>
</tr>
<tr>
<td>IFG</td>
<td>Inferior frontal gyrus</td>
</tr>
<tr>
<td>IFOF</td>
<td>Inferior fronto-occipital fasciculus</td>
</tr>
<tr>
<td>IPL</td>
<td>Inferior parietal lobule</td>
</tr>
<tr>
<td>MFG</td>
<td>Middle frontal gyrus</td>
</tr>
<tr>
<td>OL</td>
<td>Occipital lobes</td>
</tr>
<tr>
<td>SLF</td>
<td>Superior longitudinal fasciculus</td>
</tr>
<tr>
<td>SMG</td>
<td>Supramarginal gyrus</td>
</tr>
<tr>
<td>SPL</td>
<td>Superior parietal lobule</td>
</tr>
<tr>
<td>Thal</td>
<td>Thalamus</td>
</tr>
</tbody>
</table>