Bilingualism, Mind, and Brain

Judith F. Kroll,1 Paola E. Dussias,2 Kinsey Bice,1 and Lauren Perrotti2

1Department of Psychology, 2Department of Spanish, Italian, and Portuguese, Program in Linguistics, Center for Language Science, Pennsylvania State University, University Park, Pennsylvania 16802; email: jfk7@psu.edu

Keywords
bilingual, monolingual, event-related potential, mind, brain

Abstract
The use of two or more languages is common in most of the world. Yet, until recently, bilingualism was considered to be a complicating factor for language processing, cognition, and the brain. The past 20 years have witnessed an upsurge of research on bilingualism to examine language acquisition and processing, their cognitive and neural bases, and the consequences that bilingualism holds for cognition and the brain over the life span. Contrary to the view that bilingualism complicates the language system, this new research demonstrates that all of the languages that are known and used become part of the same language system. The interactions that arise when two languages are in play have consequences for the mind and the brain and, indeed, for language processing itself, but those consequences are not additive. Thus, bilingualism helps reveal the fundamental architecture and mechanisms of language processing that are otherwise hidden in monolingual speakers.
1. INTRODUCTION

Bilinguals are individuals who actively use more than one language, but bilingualism arises in many different ways. Some individuals are exposed to two languages from birth and continue to use both languages throughout their lives. Early bilinguals may live in a bilingual environment where they are the majority-language users or in a context in which only one of the two languages is used by most speakers. Other bilinguals acquire the second language (L2) only after early childhood, once a native language has been firmly established. Like early bilinguals, these late bilinguals may live in a range of different environments in which most individuals become bilingual or in which only some become bilingual. In the research we review, we consider anyone who actively uses two languages to be bilingual, but we also acknowledge that not all bilinguals are the same (e.g., Luk & Bialystok 2013). The differences among bilinguals that result as a function of learning history and the context of language use are important, but so too are differences in the languages themselves and the proficiency and relative dominance with which each language is used.

Three discoveries about bilingualism (Kroll et al. 2014) have shaped the current research agenda. One is that both languages are always active when bilinguals listen to speech, read words in either language, and plan speech in each of the two languages (e.g., Marian & Spivey 2003, Dijkstra 2005, Kroll et al. 2006). The parallel activation of the two languages means that there are influences of the language not in use, even when bilinguals are unaware of those influences. It also means that the two languages potentially compete for cognitive resources and that bilinguals must acquire a way to control or regulate that competition so as not to erroneously use the unintended language or to lose fluency in each language.

The second discovery is that not only does the native language influence the L2, but the L2 comes to influence the L1 once bilinguals are adequately proficient in the L2. If the two languages were represented separately, as suggested by early models (see Kroll & Tokowicz 2005 for a review of the claims about separate versus shared memory representations for the two languages), then bilinguals might be able to function as two monolinguals. Yet all of the evidence suggests otherwise. Indeed, if the two languages could function independently, then late bilinguals who acquired the L2 once the L1 was firmly in place might be expected to reveal fewer cross-language interactions, particularly from the L2 to the L1, and fewer consequences of bilingualism on the native language. Although the age of acquisition may have profound influences on some aspects of processing the L2 and on the structural organization of the neural networks that support language (e.g., Hernández & Li 2007, Wattendorf et al. 2014), recent studies suggest that language proficiency, rather than the age of acquisition, may be the more important factor in shaping cross-language interactions (e.g., Abutalebi & Green 2007, Steinhauer et al. 2009). Furthermore, the neural support for processing the two languages is largely shared; differences in brain activation are more likely to be attributed to the requirement to control the L1 when the L2 is used and to recruit additional cognitive resources than to basic differences in their representation in the brain (e.g., Abutalebi et al. 2005). The fundamental permeability across the two languages makes early and late bilinguals more, rather than less, similar to one another. It also makes cross-language interactions bidirectional, changing the way that the bilinguals process each of the two languages and creating different language profiles for bilingual and monolingual speakers.

The third discovery is that the use of two or more languages has consequences that extend beyond language processing to domain-general cognitive functions. These consequences are particularly beneficial to individuals at points in their lives when they may be vulnerable, especially during early development (e.g., Morales et al. 2013a) and later in life, when ordinary cognitive decline or the presence of pathology places greater demands on the cognitive system and its neural
control (e.g., Bialystok et al. 2007). Life as a bilingual appears to affect the ability to ignore irrelevant information, switch between tasks, and resolve conflicting cognitive alternatives. The initial hypothesis about these cognitive changes was that they were related directly to the first two observations we describe. The parallel activation of the two languages imposes the requirement that bilinguals learn to regulate the resulting cross-language competition. Likewise, coactivation of the bilingual’s two languages implies not only that the L1 affects the L2, but that the L2 comes to affect the L1, with a requirement to regulate the relative activation of each language in a way that is sensitive to the context in which each language is used. The need to control each language, and to potentially reduce the activation of the stronger language to enable processing in the weaker language, has implications not only for language processing itself but also for the way in which the neural networks that accomplish cognitive control are tuned in response to experience. Recent studies make clear that it is not only the general demand to juggle the two languages with skill that may affect the efficiency with which bilinguals resolve competition, but also the ways in which the two languages are used (e.g., Green & Abutalebi 2013, Green & Wei 2014). Bilinguals do many different things with language, and a goal of the recent research is to determine more specifically how the demands of particular language tasks produce changes in the brain and behavior (e.g., Kroll & Bialystok 2013).

In this review, we consider each of these discoveries, focusing primarily on the evidence that we take to be most exciting in the recent literature, and particularly on the recent findings that provide a means to relate cross-language interactions and changes to the native language to the consequences of bilingualism for the mind and the brain. Our review is necessarily selective, with an emphasis on those findings that have not been discussed extensively elsewhere. In particular, the studies conducted over the past two decades that have demonstrated cross-language activation in reading, listening to spoken language, and planning speech have been recently reviewed in some detail (e.g., Kroll et al. 2012, 2013). We briefly discuss the general findings on this topic in Section 2 but point the reader to other sources for additional background.

Critically, we note that much of the research on cross-language interaction has been conducted at the lexical level, in which the properties of words in each of a bilingual’s two languages have been exploited to examine the presence of activation in the language not in use. The findings on bilingual lexical access provide important foundational knowledge about those factors that constrain or encourage influence from one language to the other. But language is about much more than words alone. In Section 3 we address the consequences of bilingualism on the native language, but we focus our review at the sentence level. The grammars of the bilingual’s two languages sometimes converge but often conflict. This research asks how competition across the bilingual’s two grammars is resolved. Traditionally, syntactic processing in the L2 was thought to be guided by transfer from the L1 (e.g., MacWhinney 2005, Pienemann et al. 2005). But the L1 also comes to be affected by the L2. The change in native language processing in response to the L2 is not only a reflection of the fundamental openness of the bilingual’s language system but also a finding that has profound implications for how we understand the status of the native language in language processing and in language learning. If the bilingual’s two languages come to be merged to some degree, then arguments about the separate and special representations for the L2 in late acquirers (e.g., Clahsen & Felser 2006) must be reassessed.

In Section 4, we consider the consequences of bilingualism for cognition and the brain. This area of research is growing at a very rapid pace. Our review is therefore selective but focuses on those studies that enable us to move beyond correlational observations of differences between bilinguals and monolinguals to ask how different aspects of bilingualism draw on cognitive and neural resources. The factors that enable bilinguals to regulate the use of each language have
consequences not only for language but also for domain-general cognitive processes and the neural networks that support them.

2. THE PARALLEL ACTIVATION OF THE BILINGUAL’S TWO LANGUAGES

A widely replicated finding is that a bilingual’s two languages are simultaneously active during language use (e.g., Kroll & Dussias 2013). Intuitively, the native and more dominant language, L1, remains active when using the less dominant second language, L2. However, the converse is also true, such that bilinguals using their L1 also demonstrate concurrent L2 activation. Furthermore, this parallel activation, or cross-language activation, appears to be present in all bilinguals regardless of the languages they speak, including bilinguals whose languages use different written scripts, as with Chinese and English (Thierry & Wu 2007), or different modalities, as with American Sign Language (ASL) and English (Morford et al. 2011).

The evidence revealing parallel activation across the bilingual’s two languages comes from a diverse set of tasks and methods. A common technique for measuring the activity of the nontarget language is to compare the processing of language-ambiguous cognates and homographs with words that unambiguously belong to one of the bilingual’s two languages alone. Cognates are words that have similar form and meaning across a bilingual’s two languages (e.g., piano in English and Spanish). The convergence in lexical form and meaning for cognates tends to facilitate processing on word recognition and word production tasks (e.g., Costa et al. 2000, Dijkstra 2005). In contrast, homographs, which have similar lexical form but conflicting meaning across the bilingual’s two languages (e.g., Spanish carpeta means ‘folder,’ not ‘carpet’), typically produce interference in processing. In theory, monolingual speakers should process cognates and homographs no differently than other words that are matched on lexical properties such as word length and frequency. Therefore, when bilinguals demonstrate sensitivity (i.e., facilitation or interference) to language-ambiguous cognates and homographs compared with matched unambiguous control words, we can infer that the heightened activity of their nontarget language facilitates processing when form and meaning agree (cognates) or that the heightened activity competes with the target language when there is conflict between form and meaning (homographs).

Studies have reported that a bilingual’s nontarget language is active in a range of tasks that require only one of the two languages to be used. Cognate facilitation has been found for isolated word reading in both the L1 (Van Hell & Dijkstra 2002) and the L2 (e.g., Dijkstra 2005), as well as for sentence reading in the L1 and L2 (e.g., Schwartz & Kroll 2006, Libben & Titone 2009). Bilinguals also demonstrate parallel activation of the two languages while listening to speech in either language (e.g., Marian & Spivey 2003, Lagrou et al. 2011). Most surprisingly, both languages appear to be active when bilinguals plan to speak in one language alone, even though initiation of speech planning lies within the control of the speaker (for reviews on bilingual language production, see Kroll et al. 2006, Kroll & Gollan 2014). The pervasive presence of parallel activation at all levels of language processing is further confirmed by the diversity of bilinguals for whom these findings hold true.

Bilingualism in Dutch and English or in Spanish and English is frequently studied. In each case, there are many cognates across both languages and the two languages share the Roman alphabet. But cross-language activation has also been reported for bilinguals whose two languages differ in lexical form. For example, Chinese–English bilinguals may have spoken cognates across languages, with shared phonology, but the script differences prevent orthographic overlap. Despite these structural differences across languages, many past studies have shown that different-script bilinguals experience parallel activation of both languages in much the same way as same-script bilinguals (Thierry & Wu 2007, Hoshino & Kroll 2008). Hoshino & Kroll (2008) demonstrated
that Japanese–English bilinguals were facilitated to the same degree in naming pictures whose names were phonological cognates as Spanish–English bilinguals for whom the cognate names shared both phonological and orthographic information.

An even more dramatic example of cross-language interaction in the face of structurally distinct languages is the case of bimodal bilinguals, individuals who use a spoken or written language and a signed language. Morford et al. (2011) asked deaf ASL–English bilinguals to make semantic relatedness judgments about pairs of words in English. No ASL was presented, but some of the English words had ASL translations that were either similar or different in form. Morford et al. found that the speed of semantic relatedness judgments for the English words was affected by whether the form of the ASL translations converged or conflicted with the semantic relatedness of the English words. This result suggests that there is implicit activation of the translation into the language not in use even when the two languages are strikingly different in lexical form. Likewise, Thierry & Wu (2007) reported evidence using event-related potentials (ERPs) suggesting that early in processing there is neural sensitivity to the translation. This finding is remarkable because the Chinese–English bilinguals in this study performed the semantic relatedness judgment task in English only, without any Chinese actually present and in a context in which the bilinguals were immersed in an English-speaking environment. Such data are remarkable because they demonstrate that conscious awareness of the other language is not required to observe these effects; both languages are momentarily active when bilinguals process words in one language alone, despite the apparent dissimilarity of the bilingual’s two languages that might be used as a means to induce functional separation.

Parallel activation appears to be a product of knowing two languages, independent of what those languages are or how they are used. The fact that cross-language activation is found in language comprehension and production, and for both written and spoken or signed language, suggests that the L1 and L2 are interconnected. The fact that it is found in bilinguals that differ so widely in their two languages suggests that there are at least some common, shared language storage or processing mechanisms that support both languages regardless of what they are. Critically, the parallel activation of the two languages has consequences, both for language use and for cognition in general.

In a well-known paper, Grosjean (1989) claimed that a bilingual is not two monolinguals in one. The context for this claim was a debate about whether language switching is a normal or pathological feature of bilingual language use. We now know that code switching and language mixing are not only normal features of bilingualism but also skills that may effectively enhance cognitive abilities in bilinguals (e.g., Green & Abutalebi 2013, Green & Wei 2014). But it is also important to note that the pervasive parallel activation of the bilingual’s two languages is likely to be a primary mechanism that contributes to language change and that differentiates bilingual and monolingual language use. When a bilingual is using the less dominant L2, the concurrent L1 activation is pervasive. Given the choice between using two words or two structures, one of which shares overlap with the L1, the simultaneous L1 activation will promote that alternative over a word or structure that appears in the L2 only [see, e.g., evidence on cross-language syntactic priming (Hartsuiker et al. 2004)]. Choosing common structures across languages when using the L2 is not that surprising; however, the same process applies to a bilingual using his or her L1, although processing in the L1 may be sufficiently skilled and automatic as to mask the influence of the L2. These small, incremental changes accumulate with increasing L2 proficiency and with time; continuing to select language–similar structures over language-dissimilar structures changes the way both languages are used.

If the bilingual’s two languages are always active, to be able to use the target language there must be regulation to control the influence of the language not in use. The fact that bilinguals
rarely produce erroneous language intrusions (e.g., Gollan et al. 2011) suggests that they develop an effective means to control the language they do not intend to use. Although there is debate about the means by which bilinguals control the use of each language (see Kroll & Gollan 2014 for a recent review of this debate in the realm of speech production), there is evidence in both comprehension and production that the activation of the language not in use must be reduced and that, at least under some circumstances, there is active suppression of the dominant language to enable the processing of the less dominant language (e.g., Kroll et al. 2008).

The practice of resolving competition between both languages has been hypothesized to depend on domain-general cognitive skills that are honed over time with daily language use (see Bialystok et al. 2012 for a review). We discuss this issue in more detail in Section 4, but note that not only does bilingualism come to affect cognition, but a consequence of multiple language use is to produce greater efficiency in language processing itself. Blumenfeld & Marian (2011) tested this hypothesis in a study on spoken word recognition in the L1. They used the visual world paradigm (Allopenna et al. 1998), in which speakers viewed a grid with four objects, one in each of the four corners of the grid. The task is to point or click on an object that is named by a spoken word while eye movements are recorded. Past research has shown both within- and between-language competition when an object shares the phonology of a named object (e.g., Marian & Spivey 2003). The eye-tracking record reveals extended fixations to phonological competitors relative to unrelated controls. Blumenfeld & Marian (2011) compared the performance of bilinguals with that of monolinguals identifying words in their L1. Both groups exhibited effects of the phonological competitor. On a second trial, the grid was presented but without the four objects. Instead, an asterisk appeared in one of the four corners. The task was to indicate the location of the asterisk. Monolinguals were slower to identify the asterisk when it appeared in the position of the phonological competitor, but bilinguals were not, suggesting that the latter were able to more efficiently resolve the competition from the previous trial. As discussed in Section 4, below, the efficiency of resolving conflict, within or across languages or within purely cognitive tasks, is a theme that has emerged in recent work as a hallmark of bilingualism.

3. DYNAMIC CHANGES DURING SYNTACTIC PROCESSING

It is widely known that adult second-language acquisition is marked by variable mastery of the target language, most notably in the acquisition of the sound system and the grammatical structures of the L2 (e.g., Flege 1987, 2007; Hawkins & Chan 1997; Franceschina 2001, 2005). Because L2 acquisition involves more than simply memorizing individual sounds, words, or verbal paradigms, there is an ongoing debate about the causes underlying the observed incomplete acquisition. Some scholars assert that the relevant L2 information—including even subtle aspects of the L2 grammar—can be fully acquired by adult learners (e.g., Schwartz & Sprouse 1994, 1996; Lardiere 1998a,b; Prévost & White 2000; Bruhn de Garavito & White 2002; Keating 2009; Gillon Dowens et al. 2010; Alarcón 2011; Foucart & French-Mestre 2011, 2012; Grüter et al. 2012; Witzel et al. 2012; Hopp 2013) and challenge the existence of fundamental differences (Clahsen & Felser 2006) between native and nonnative speakers. According to this view, errors in L2 performance are modulated by a number of variables, including proficiency and immersion experience in the L2 (e.g., Kotz 2009, Pliatsikas & Marinis 2013), speed of lexical access (Hopp 2013), and availability of cognitive and computational resources (e.g., McDonald 2006, Hopp 2010). Other researchers have argued instead that access to the pool of all potential linguistic structures is not available after the so-called critical period, which coincides roughly with the onset of puberty. In this view, the absence of the relevant structures in the L1 presents an obstacle to achieving complete acquisition of the second language (e.g., Eubank 1993, Hawkins & Chan 1997, Beck
In support of this idea, many studies have highlighted that success in second-language acquisition is strongly influenced by the linguistic properties of the learner’s first language (e.g., Zobl 1982, Andersen 1983, Rutherford 1983) and that the ability of L2 speakers to process knowledge online depends in part on similarities between the L1 and the L2 (Sabourin et al. 2006, Sabourin & Stowe 2008).

It is becoming increasingly clear, however, that transfer phenomena are not restricted to speakers’ incorporation of L1 linguistic elements into their L2 system. Evidence from behavioral and electrophysiological studies has shown not only that the first language affects the second language (e.g., Flege & Davidian 1984, Hancin-Bhatt 1994, Hatzidakis et al. 2011) but also that the second language can come to affect the native language (e.g., Dussias 2003, Dussias & Sagarraga 2007, Valdés Kroff 2012, Dussias et al. 2014). In this section, we discuss some recent evidence examining the interactions between the bilingual’s two languages by focusing on the influence of the L2 on the L1. Research on first-language attrition (e.g., Lambert & Freed 1982, Seliger & Vago 1991, Schmid et al. 2004) has investigated changes in the L1 that come about through lack of exposure, which typically occurs in bilinguals who maintain little if any contact with speakers of the L1 or in bilinguals who have negative attitudes toward their L1 (Cook 2003). Our goal here is to examine the influence of the L2 on the L1 in bilinguals who actively use their two languages daily in a variety of formal and informal contexts and who value maintenance of the L1. Examining this question has important implications not only for our understanding of how bilinguals manage to negotiate their two languages, but more critically for existing assumptions about the plasticity of cognitive and neural representations.

Although evidence of the influence of the L2 on the native language is broad and has been observed at the level of the lexicon (e.g., Ameel et al. 2005, Linck et al. 2009, Baus et al. 2013), the phonology (e.g., Chang 2013), and the syntax (e.g., Dussias & Sagarraga 2007), our discussion focuses on syntactic processing while bilinguals make structural decisions about the words they read or hear. We choose this focus because achieving high levels of L2 syntactic and morpho-syntactic knowledge has been considered to be more constrained by age of acquisition than are other areas of language, such as semantics (e.g., Hahne & Friederici 2001) or pragmatics (e.g., Félix-Brasdefer & Koike 2012). Therefore, evidence suggesting that the purportedly stable L1 syntactic system is subject to influence from the L2 provides a strong piece of evidence for a linguistic system that is far more open and dynamic than previously thought (Hernández et al. 2005).

One significant insight from the L2 acquisition work is that prolonged naturalistic exposure can have profound effects on how a second language is processed, reversing processing strategies that result from transfer of L1 information (Frenck-Mestre 2002) or causing shifts in L2 processing strategies from lexically driven to structurally driven1 (Pliatsikas & Marinis 2013). Given this evidence, an important aspect of the comparison between L2 and L1 speaker performance is to consider how immersion experience might affect L1 processing. Work examining this question shows a great deal of permeability across language boundaries at different levels. For example, in immersion contexts, L2 speakers experience reduced access to the L1 (Linck et al. 2009) and extensive contact with an L2, which can affect both L1 naming performance of common objects (Malt & Sloman 2003) and L1 phonology (e.g., Flege 1987, Flege & Eefting 1987). These findings confirm that a bilingual is not two monolinguals in one brain (Grosjean 1989) and that the

---

1Lexical processing relies on lexical semantic or conceptual information associated with a given lexical target. In contrast, in structurally driven processing structural decisions derive from general principles of configural economy along the lines of those proposed in Kimball (1973) or Frazier & Fodor (1978).
seemingly stable L1 system is open to influence once individuals become proficient in the L2 (e.g., Gollan et al. 2008, Ivanova & Costa 2008, Runnqvist et al. 2013). Given this scenario, one might expect that experience in a second-language environment should also produce changes in syntactic processing in the native language. Findings reported for other linguistic domains (e.g., Bybee 1994) show that the frequency of linguistic items has important consequences for language change. If the frequency of exposure to a second language affects the native language, then bilinguals immersed in the L2 environment, who are matched in proficiency to bilinguals immersed in the L1 environment, should behave differently when processing sentences in their native language.

Dussias & Sagarrà (2007; also see Fernández 2003) tested the exposure hypothesis in a study that investigated the effect of intense contact with English on the resolution of syntactically ambiguous relative clauses in Spanish. Native Spanish and native English speakers differ in how they interpret temporarily ambiguous relative clauses such as Alguien disparó al hijo de la actriz que estaba en el balcón/‘Someone shot the son of the actress who was on the balcony.’ When asked the question ¿Quién estaba en el balcón? (‘Who was on the balcony?’), monolingual Spanish speakers typically respond ‘the son’ (i.e., high attachment preference), whereas monolingual English speakers respond ‘the actress’ (low attachment preference) (Carreiras & Clifton 1999). Using tasks measuring ambiguity resolution during online reading, Dussias & Sagarrà (2007) found that Spanish–English bilinguals immersed in a Spanish-speaking environment processed the ambiguity using a high attachment strategy. This was an expected finding. The interesting result was that bilinguals living in an English-speaking environment strongly favored the low attachment strategy when reading in Spanish, their native language. That is, for these speakers, exposure to a preponderance of English constructions resolved in favor of low attachment may have rendered this interpretation more available, resulting in a low attachment preference when reading in their first language.

One question raised by these results is whether parsing preferences can show short-term changes on the basis of artificial exposure to particular patterns over a predetermined amount of time. In other words, can changes be triggered by overexposing bilinguals to particular structures so that proficient bilinguals who have undergone change in their L1 “move back” and those who have not achieved the desired outcome “move forward”? Past findings from the child processing literature suggest that exposure experience can affect children’s sentence-processing routines (e.g., Cuetos et al. 1996, Trueswell et al. 1999). Similar findings have been reported with adult monolingual speakers, demonstrating powerful implicit learning properties that characterize the human language system. For example, the notorious difficulty that native English speakers experience when processing object-extracted relative clauses such as The reporter that the senator attacked admitted the error relative to subject-extracted relative clauses such as The reporter that attacked the senator admitted the error (Traxler et al. 2002, Staub 2010) disappears with increased exposure to object-relatives (Wells et al. 2009). If the parser’s configuration is related to language exposure (MacDonald 1994, MacDonald & Seidenberg 2006, Gennari & MacDonald 2009) and language contact, then bilinguals’ processing routines are expected to change as a function of the frequency with which the relevant structure appears in an experimental session.

Ongoing experiments on intervention in language exposure provide support for the dynamic nature of parsing. In one study (Dussias et al. 2014), L1 Spanish–L2 English bilinguals identified as being either so-called high attachers or low attachers (via an eye-tracking study) participated in a five-day intervention, during which they read short paragraphs containing relative clauses in which the syntactic ambiguity was resolved opposite to the bilinguals’ natural attachment preference. That is, participants who favored high attachment received a low attachment treatment, and those who favored low attachment received a high attachment treatment. In addition, half the participants received the intervention in Spanish and the other half in English. We know
from monolingual research that experience with L1 structures can guide the interpretation of temporary ambiguity (Wells et al. 2009), but one remaining question is whether exposure to linguistic structures can trigger change when the input is in a different language. Participants returned to the lab after the intervention to participate in two subsequent eye-tracking studies, one that assessed the immediate effect of the intervention and one that assessed the effect of the intervention a week after it was completed. Ongoing analyses show that those participants who originally preferred high attachment switched to a low attachment preference after the intervention and maintained the low attachment preference a week later. Participants who originally showed a low attachment strategy switched to a high attachment preference after the intervention, but eventually reverted back to a low attachment preference. Critically, whether the intervention was in Spanish or in English did not affect the pattern of results. The results indicate that it is not exposure to a language that causes the switch, but rather exposure to the structure itself.

One might wonder whether immersion in the L2 can come to affect any aspect of L1 syntactic processing or whether some structures are more susceptible to intense exposure to a second language. Given that syntactic structures differ along many dimensions, it is possible that different L1 structures are differentially affected by immersion in the second language. For example, effects of naturalistic exposure may become evident in the processing of information that is somewhat “peripheral” to the so-called core syntax, but may be negligible when information has syntactic consequences (e.g., when the selection of verbal mood in a complement clause is a syntactic reflex of the selectional relationship between a matrix verb and its complement). It is also possible that changes are more easily triggered when structures overlap in some way in the bilingual’s two languages (Hartsuiker et al. 2004) than when they do not. An important question, then, is whether L1-specific morphosyntactic information is affected by the influence from the linguistic input available in the L2 environment, such that this information is no longer used during L1 sentence processing in bilinguals in a manner that resembles monolingual use. To test this hypothesis, Valdés Kroff (2012) examined whether intense contact with Spanish–English code-switched speech had consequences in the processing of grammatical gender in Spanish. In many bilingual communities, speakers regularly switch from one language to another, often several times in a single utterance. This phenomenon is called code switching. The ability to engage in fluent code switching is a hallmark of high proficiency in two languages (Miccio et al. 2009), given that successful and fluent code switching requires a high degree of knowledge of and sensitivity to the grammatical constraints of both languages. Grammatical gender was examined because it is a lexically specified feature of nouns that triggers syntactic agreement with other function and open-class elements, such as determiners and adjectives. In the context of code switching, grammatical gender is interesting for yet another reason. In the production literature on Spanish–English code switches, one widely attestable pattern is that when a code switch occurs within a noun phrase composed of a determiner and a noun, the determiner overwhelmingly surfaces in Spanish and the noun in English, such as ‘el building’ and not ‘the edificio.’ Researchers have also documented a production asymmetry in grammatical gender assignment in these mixed noun phrases. The Spanish masculine article el (‘the’) surfaces with English nouns regardless of the grammatical gender of their translation equivalents, for example, ‘el juice’ (Spanish jugo, masculine) or ‘el cookie’ (Spanish galleta, feminine). By contrast, mixed noun phrases involving the Spanish feminine article la (also translated as ‘the’) are rare and occur in restricted environments, such that only English nouns whose Spanish translation equivalents are feminine surface with la in code switching (e.g., ‘la cookie’ but not ‘la juice’; see Poplack 1980, Jake et al. 2002, Otheguy & Lapidus 2003). These production distributions in Spanish–English code switching stand in marked contrast to monolingual Spanish, where the grammatical gender of a noun and its accompanying article must obligatorily match, and where masculine and feminine nouns are evenly distributed (Eddington 2002, Otheguy & Lapidus 2003).
As mentioned above, Valdés Kroff (2012) investigated whether L1 Spanish–L2 English speakers who had been immersed in a code-switching environment used grammatical gender information encoded in Spanish articles to anticipate the gender of an upcoming noun (as Lew-Williams & Fernald 2007 have shown with monolingual speakers of Spanish). Given the asymmetry observed in production data, it seemed plausible that the gender marking of articles would, to a lesser extent, facilitate the processing of code-switched speech. To investigate this hypothesis, Valdés Kroff (2012) recorded the eye movements of Spanish–English bilinguals. Participants saw two pictures that represented objects whose nouns were of either the same or different grammatical genders in Spanish, and listened to code-switched sentences (e.g., Hay un niño que está mirando el candy/'There is a boy looking at the candy'). Words were spoken with a Spanish article that either matched the gender of the word’s Spanish translation equivalent (e.g., el candy_{MASC}, Spanish caramelo) or did not match (target: el candle_{FEM}, Spanish vela). A control group of monolingual speakers listened to sentences entirely in Spanish. Participants were asked to listen to each sentence and to click on a named object. Where the grammatical gender of the Spanish names for each of the pictures was different between the two pictures, the gender information in the article was informative. In that condition, monolingual Spanish speakers showed the expected anticipatory effect on masculine and feminine different-gender trials. Results for the L2 speakers revealed an anticipatory effect, but only on different-gender trials in which the auditory stimulus was feminine. When the auditory stimulus was masculine in different-gender trials, participants did not launch anticipatory looks but rather waited to hear the target noun, meaning that they did not use masculine articles as cues for anticipatory processing. Because the masculine article is overwhelmingly more common than the feminine in Spanish–English code switching but not in monolingual Spanish, these group differences suggest that how grammatical cues are exploited in mixed-language processing is driven by experience with the statistical patterns attested in actual communicative contexts. Critically, they also reveal the key role that linguistic exposure has in language-comprehension processes. Although there surely may be constraints on the sort of cross-language exchanges in bilinguals, the presence of these influences—and especially of those from the L2 to the L1—suggests a dynamic language system that changes in response to contact with other languages. In the final section of this review, we consider the consequences of these dynamics in bilingual language for cognition more broadly.

4. THE COGNITIVE AND NEURAL CONSEQUENCES OF BILINGUALISM

Recently there has been widespread media coverage on the provocative claims about the ability of bilingualism to protect individuals against the deleterious consequences of cognitive aging. Elderly bilinguals have been reported to be better able to switch between tasks, to ignore irrelevant information, and to resolve conflicting cognitive alternatives (see, e.g., Bialystok et al. 2009 for a review). But most striking is the evidence on the consequences of bilingualism when some pathology is present, as in the case of Alzheimer’s-type dementia. Bialystok et al. (2007) reported that individuals with Alzheimer’s disease were diagnosed with the disease an average of four to five years later if they were bilingual rather than monolingual (also see Alladi et al. 2013 for a recent replication in a different language context). Furthermore, brain scans of bilingual and monolingual Alzheimer’s patients who were matched for disease symptoms revealed that the brains of the bilingual patients were more diseased than those of monolingual patients (e.g., Schweizer et al. 2012). This comparison suggests that the active use of two languages protects bilinguals from the symptoms of the disease. The areas of the brain that enable cognitive control are hypothesized to have benefited from bilingual experience. Under conditions of pathology and stress, the skill developed as a function of juggling two languages over the course of a person’s life provides a level
of cognitive reserve to enable individuals to continue to function, despite the presence of disease. Bilingualism itself does not directly affect Alzheimer’s disease but instead enables functional cognition for a longer period of time, relative to monolingualism, following the onset of the disease.

But what aspect of bilingualism is responsible for these benefits? Past research claimed that a life spent resolving cross-language competition, of the sort we have documented at the level of the lexicon and the level of grammar, creates skill more generally in resolving conflict. The hypothesis is that the same cognitive networks that enable selection and decision making in the rest of life outside of language are also engaged when language is used. Bilingualism is then thought to provide a rich foundation for developing these control skills because language is so prevalent in cognitive life. Other activities (e.g., playing video games, becoming a skilled musician, driving a taxi cab through a maze of city streets) will also tune cognitive networks, but important as they are, it is unlikely that these activities will be used as frequently as language.

Until recently, the account for the consequences of bilingualism was that the constant activity of the two languages produces competition that then requires resolution. Bilinguals therefore become expert in resolving competition and then reveal that expertise when a task, linguistic or not, presents similar conflict. The problem with this simple story, elegant as it may seem, is that not all bilinguals reveal precisely the same benefits, nor benefits in every conflict task, nor benefits at every point across the life span (e.g., Kroll & Bialystok 2013; S. Baum & D. Titone, manuscript accepted). As noted in Section 1, there are many different types of bilinguals, and even bilinguals who are similar with respect to the languages they speak may use those languages in different contexts and with different interlocutors. A goal in recent research has been to identify the critical features of bilingual experience, including learning history and the context in which the two languages are used, and the critical aspects of cognitive control that are required to enable proficient language performance under a range of circumstances.

Recent neuroimaging studies have also sought to identify the brain mechanisms that might differentiate monolingual and bilingual experience. For example, Abutalebi et al. (2012) asked young adult bilinguals to perform a variant of the flanker task, a nonlinguistic conflict-resolution task in which individuals must decide whether a symbol, such as an arrow, is facing in one direction or another, in the context of other information that converges or diverges with the solution. They reported greater brain activation in a functional magnetic resonance imaging (fMRI) paradigm for the anterior cingulate cortex (ACC) in monolinguals than in bilinguals when solving the same task, suggesting that bilinguals are more efficient at using the brain areas engaged by conflict resolution than monolinguals. They concluded that bilingualism comes to “tune” the ACC. A similar fMRI result was reported by Gold et al. (2013) using a nonlinguistic switching task with bilinguals and monolinguals who were either younger or older healthy adults. In this study, only the older bilinguals revealed an advantage relative to monolinguals, demonstrating that older bilinguals were also more efficient in switching between conditions than their monolingual counterparts. The absence of an effect for the younger adults, typical in many studies examining the consequences of bilingualism, has been understood as a matter of sensitivity. Young adults are at their peak cognitive performance, and it may be difficult to enhance performance beyond that peak level. Alternatively, it may require more extensive bilingual experience over a longer period of time for these benefits to be revealed. Regardless, in each case these studies demonstrate that under conditions in which cognitive performance does not depend directly on language use, there is a difference between bilinguals and monolinguals that is consistent with the claim that bilingualism affects those brain networks that enable cognitive control.

To begin to explain the range of circumstances over which processing differences are reported for bilinguals and monolinguals, Green & Abutalebi (2013) proposed the adaptive control model (for a recent review of the range of bilingual effects that have been reported, see S. Baum & D. Titone,
manuscript accepted). This model assumes that there are multiple components of cognitive or inhibitory control that are recruited to perform particular language tasks. For example, a different subset of control mechanisms may be engaged when bilinguals code-switch with other similar bilingual speakers and when bilinguals are immersed in an environment in which the L2 is the dominant language. Over time, both under the sort of immediate pressures described in the intervention studies above and over a longer period of time when individuals are immersed in a foreign country and culture, there will be a dynamic set of changes in which there is adaptation to the demands of the environment. The variation in language use and language context, in level of language skill, and in the nature of the language pairings themselves will dictate the ways in which control networks are engaged. Thus, bilinguals will differ not only from monolinguals but also from other bilinguals whose circumstances differ from their own. Green & Abutalebi (2013) use code switching to illustrate the way in which these differences may emerge. As noted above, some bilinguals code-switch intrasententially and others do not. These two types of bilinguals may be otherwise similar in language proficiency, but habitual code switching may engage a subset of control mechanisms on which nonswitchers do not acquire the same level of practice. An area of active research in the current literature seeks to better understand the range of consequences conferred by habitual code switching (e.g., Green & Wei 2014).

In addition to focusing on how different forms of bilingualism affect the engagement of cognitive mechanisms, we can also ask how different language and cognitive tasks themselves engage different components of inhibitory control. Morales et al. (2013b), using different measures of control, examined the performance of bilinguals and monolinguals on proactive and reactive components of inhibition. Proactive inhibition is hypothesized to reflect early attentional mechanisms that enable focus on anticipated conditions. Reactive inhibition is thought to reflect control that occurs only after some set of alternatives are activated, and a selection must occur relatively late in processing. Critically, Morales et al. (2013b) found that a bilingual advantage was not present in general but depended on the interaction between conditions. Under conditions in which there was a demand for adjusting the proactive and reactive components of processing, bilinguals outperformed monolinguals, suggesting not a simple componential advantage but rather a different calibration of cognitive control in processing (also see Costa et al. 2009 for related evidence on how a bilingual advantage depends on the difficulty of monitoring).

Other support for the notion that there is a wide range of cognitive consequences of bilingualism comes from studies of infants who are exposed to more than one language from birth. These “crib bilinguals” are not speaking either of the two languages to which they are exposed, but the multiple exposure has critical consequences for the way in which their representation of speech is tuned (e.g., Sundara et al. 2006) and furthermore imposes a range of consequences for attention and for language discrimination (e.g., Kovács & Mehler 2009, Sebastián-Gallés et al. 2012).

A goal for future research is to begin to identify the constellation of control mechanisms that are engaged by bilingualism. The evidence on language processing suggests a dynamic system in which regulation among competing alternatives is necessary to enable the intended language to be used. Studies of inhibitory control in language processing have focused on conditions that make language selection more difficult, for example, when the use of the two languages is mixed (e.g., Christoffels et al. 2007) or when bilinguals are required to switch from one language to the other (e.g., Meuter & Allport 1999). Under these circumstances, differential costs are typically observed for the L1 relative to the L2, with the interpretation that the more proficient or more dominant language needs to be regulated to enable the use of the weaker or less proficient language. Recent studies on language processing, using the power of converging methods across neuroscience and behavior, have reported findings very similar to those described for cognitive tasks. There appear to be multiple components of inhibitory control, some local and short lived and others more global,
extended in time, and perhaps affecting the entire language at once (e.g., Guo et al. 2011, Misra et al. 2012, Van Assche et al. 2013). For example, Misra et al. (2012) reported evidence for inhibition in L1 speech production in an ERP paradigm when the L1 was spoken following the L2. Of interest is that the inhibitory pattern was long lived in the experiment, with little evidence for immediate recovery. Although the Chinese–English bilinguals who were tested were dominant in Chinese as the L1, they were highly proficient in English as the L2, suggesting that the inhibitory pattern reflected not a lack of proficiency in speaking the L2 but rather a more general requirement to inhibit the L1 when speech was planned in the L2. What is not yet well understood is how these components of inhibitory control in language processing map onto the cognitive differences that have been identified for bilinguals. A goal in ongoing research is to catch language processing on the fly as a means to reveal the causal mechanisms that are responsible for these effects.

5. SUMMARY AND CONCLUSIONS

In the past two decades, there has been an upsurge of research investigating the consequences of bilingualism for language processing, for cognition, and for the brain. This research demonstrates a remarkable level of plasticity across the bilingual’s two languages, with evidence that the two languages engage directly within a single language system that is stretched in different directions by the conflicts and convergences present across each level of language use. Not only is the L2 shaped by the L1, but the L1 comes to be changed by the interactions with the L2. Those interactions, and the constant activation of the two languages, create demands on cognitive systems and the neural mechanisms that control them so that the bilingual is able to function fluently in each language and, at the same time, mix languages when switching is desirable or necessary. These consequences of bilingualism highlight the adaptive nature of language processing and illustrate how bilingual experience becomes a lens for language scientists to understand the relations between language use, cognition, and the brain.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

The writing of this review was supported in part by National Institutes of Health grant HD053146 to J.F.K., National Science Foundation grant BCS-0821924 to P.E.D., National Science Foundation grants BCS-0955090 and OISE-0968369 to J.F.K. and P.E.D., and a Guggenheim Fellowship to J.F.K.

LITERATURE CITED


# Contents

Suppletion: Some Theoretical Implications  
*Jonathan David Bobaljik* ................................................. 1

Ditransitive Constructions  
*Martin Haspelmath* ..................................................... 19

Quotation and Advances in Understanding Syntactic Systems  
*Alexandra D’Arcy* .......................................................... 43

Semantics and Pragmatics of Argument Alternations  
*Beth Levin* ................................................................. 63

Events and Situations  
*Sandro Zucchi* .............................................................. 85

Vagueness and Imprecision: Empirical Foundations  
*Stephanie Solt* ............................................................. 107

Cross-Linguistic Temporal Reference  
*Judith Tonhauser* .......................................................... 129

Variation in Information Structure with Special Reference to Africa  
*Tom Güldemann, Sabine Zerbian, and Malte Zimmermann* .......... 155

Diachronic Semantics  
*Ashwini Deo* ................................................................. 179

The Indo-European Homeland from Linguistic and Archaeological Perspectives  
*David W. Anthony and Don Ringe* ...................................... 199

Correlational Studies in Typological and Historical Linguistics  
*D. Robert Ladd, Seán G. Roberts, and Dan Dediu* .................... 221

Advances in Dialectometry  
*Martijn Wieling and John Nerbonne* ................................... 243
<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sign Language Typology: The Contribution of Rural Sign Languages</td>
<td>265</td>
</tr>
<tr>
<td>Connie de Vos and Roland Pfau</td>
<td></td>
</tr>
<tr>
<td>Genetics and the Language Sciences</td>
<td>289</td>
</tr>
<tr>
<td>Simon E. Fisher and Sonja C. Vernes</td>
<td></td>
</tr>
<tr>
<td>Language Abilities in Neanderthals</td>
<td>311</td>
</tr>
<tr>
<td>Sverker Johansson</td>
<td></td>
</tr>
<tr>
<td>How Nature Meets Nurture: Universal Grammar and Statistical Learning</td>
<td>333</td>
</tr>
<tr>
<td>Jeffrey Lidz and Annie Gagliardi</td>
<td></td>
</tr>
<tr>
<td>Bringing Machine Learning and Compositional Semantics Together</td>
<td>355</td>
</tr>
<tr>
<td>Percy Liang and Christopher Potts</td>
<td></td>
</tr>
<tr>
<td>Bilingualism, Mind, and Brain</td>
<td>377</td>
</tr>
<tr>
<td>Judith F. Kroll, Paola E. Dussias, Kinsey Bice, and Lauren Perrotti</td>
<td></td>
</tr>
<tr>
<td>Taking the Laboratory into the Field</td>
<td>395</td>
</tr>
<tr>
<td>D.H. Whalen and Joyce McDonough</td>
<td></td>
</tr>
<tr>
<td>Errata</td>
<td></td>
</tr>
<tr>
<td>An online log of corrections to Annual Review of Linguistics articles may be found at <a href="http://www.annualreviews.org/errata/linguistics">http://www.annualreviews.org/errata/linguistics</a></td>
<td></td>
</tr>
</tbody>
</table>