The Pennsylvania State University

The Graduate School

The College of the Liberal Arts

EXAMINING STRUCTURAL CONSTRAINTS IN BILINGUAL LANGUAGE PROCESSING AND SPEECH PRODUCTION

A Thesis in

Psychology

by

Christian A. Navarro-Torres

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Master of Science

May 2016

The thesis of Christian A. Navarro-Torres was reviewed and approved* by the following:

Judith F. Kroll
Distinguished Professor of Psychology, Linguistics, and Women's Studies
Thesis Co-Adviser

Paola E. Dussias
Professor of Spanish, Linguistics and Psychology
Thesis Co-Adviser

Daniel Weiss
Associate Professor of Psychology and Linguistics

Jonathan Cook
Assistant Professor of Psychology

Rena Torres Cacoullos
Professor of Spanish and Linguistics

Melvin Mark
Professor of Psychology
Head of the Department of Psychology

*Signatures are on file in the Graduate School

ABSTRACT

One of the most important findings in studies of bilingual language processing is that bilinguals co-activate information from both languages across a variety of contexts and domains. While previous studies have reported syntactic co-activation/interaction, very few have examined co-activation of lexical-syntactic information in production and its relation to domain-general cognition. Here we investigate whether bilingualism modulates the use of subtle syntactic constraints in the L2 via a-adjectives (e.g., afraid, alive), which have been shown to resist prenominal use despite their adjectival status (Boyd & Goldberg, 2011). English-monolinguals, Spanish-English bilinguals (immersed in their L2), and German-English bilinguals (immersed in their L1) performed a sentence production task that elicited the use of prenominal and relative clause constructions. The task yielded three main findings. First, while all groups exhibited the a-adjective constraint, bilinguals were overall more stable in their preference. A follow-up analysis revealed that this pattern was related to errors on filler trials: those who prenominalized 3rd person singular verbs (*the sings* boy…)* were more likely to also prenominalize a-adjectives, and this effect was found to be strongest in a subgroup of monolinguals. Crucially, bilinguals as a whole were less prone to generating these errors. Given the task's level of difficulty, we argue that bilinguals may have performed the task differentially using a proactive strategy. Second, Spanish-English bilinguals were slower overall producing correct sentences, potentially reflecting inhibition of L1 syntactic information, but German-English bilinguals showed similar latencies to monolinguals, possibly indicating less cross-language interference. Finally, sentences containing a-adjectives were overall produced slower, even when compared to filler trials involving relative clauses, suggesting activation that both lexical and syntactic information are activated at the onset of planning.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGMENTS

This research was funded in part by a National Science Foundation (NSF) Graduate Research Fellowship Grant DGE1255832 to C. Navarro-Torres, and by an NSF Partnership for International Research and Education (PIRE) Grant OISE-0968369 to J.F. Kroll and P.E. Dussias.

I would also like to thank the members of the committee for their invaluable input and guidance. This project could not have been possible without Judy Kroll and Paola (Giuli) Dussias, whom were incredibly supportive from the beginning and encouraged me to follow this line of research against all odds. I would also like to thank Holger Hopp for helping me carry out the experiments in his lab in Germany, and for his guidance in pushing me to further pursue this line of research. I also thank Adele Goldberg and Jeremy Boyd, who were kind enough to share and discuss their materials for the original study in which this thesis is based on.

Lastly, I owe many thanks to several current and former members of the Center for Language Science, including Jason Gullifer and Anne Beatty-Martinez, who gave me a lot of valuable feedback throughout the development of the project, and my research assistants: Nick Pollio, Ashley Wagner, Michele Roques-Brador, and Maria Fabrizio for their help.

1. Introduction

A pervasive finding in studies of bilingual language processing is that, even in a unilingual context, both languages are constantly active and competing for selection (e.g., Costa, 2005; Dijkstra, 2005; Kroll, Dussias, Bice & Perrotti, 2015). This kind of language non-selectivity (or parallel activation) has consistently been found at the single-word level across different language pairs (e.g., Hoshino & Kroll, 2008; Thierry & Wu, 2007), across different stages of second language (L2) proficiency (Christoffels, De Groot, & Kroll, 2006; Van Hell & Dijkstra, 2002), and even across language modalities (i.e., from spoken to signed language; Emmorey, Borinstein, Thompson, & Gollan, 2008). Evidence of parallel activation has also been observed with respect to syntactic co-activation/interaction in both comprehension (Dussias & Sagarra, 2007) and production (Pickering & Ferreira, 2008). The presence of parallel activation in the absence of speech errors suggests that bilinguals must have a mechanism of language control that is either not present or systematically exploited in monolinguals. Both quantitative and qualitative aspects of such control mechanisms have been hypothesized to have consequences for other domains of cognition throughout a bilingual's lifespan (Bialystok, Craik, Green, & Gollan, 2009), as evidenced by how bilinguals perform in tasks involving executive functions (Carlson & Meltzoff, 2008), metalinguistic (Bialystok, 2001) and attentional skills (e.g., Costa, Hernández, and Sebastián-Gallés, 2008). Despite this, very few studies have investigated how these domain-general consequences operate when bilinguals have to produce grammatical structures in their L2, and how particular kinds of cognitive demands are imposed by specific (grammatical) features of the two languages. To our knowledge, this is one of the very first studies investigating these issues in the context of a usage-based linguistic theory (i.e., Construction Grammar; Bybee, 2010, 2013; Goldberg, 1995, 2013).

1.1. Parallel activation at the lexical level in comprehension

Evidence for parallel activation has been primarily shown in word recognition tasks. Such tasks typically ask bilinguals to read out-of-context single words that are ambiguous with respect to form and/or meaning. To illustrate, many language pairs contain translation equivalents that share similar form and/or phonology (cognates; e.g., *piano* in English is written exactly the same way in Spanish, while the pronunciation is also very similar), but the same languages can also contain words that share similar form but differ in meaning (homographs; e.g., the word *pan* means *'bread'* in Spanish). This makes it possible to compare word-recognition performance using these language-ambiguous words relative to non-ambiguous words. For example, Dijkstra, Van Jaarsveld, and Ten Brinke (1998) had Dutch-English bilinguals perform a visual lexical decision task in their L2 (i.e., determining whether the word was a real word in English or not) that included Dutch-English ambiguous words, as well as English control words. The authors found that bilinguals were faster at detecting cognates relative to controls (hence, the cognate facilitation effect). On the other hand, De Groot, Delmaar, and Lupker (2000) carried out a similar study using a combination of lexical decision and translation recognition tasks. In the latter, pairs of words were presented simultaneously (one word in the L1 Dutch, and the other in the L2 English), and participants were asked to determine whether these pairs were translation equivalents. Crucially, critical trials contained homographs in one of the two words (e.g., *glad-slippery* are translation equivalents, even though the first word interferes with its orthographical English equivalent). Overall, the results showed that bilinguals were slower when responding to trials containing homographs relative to control trials (hence, the homograph interference effect).

While cognate facilitation has been shown consistently in many studies, the effect of homograph interference has been found to vary depending on the context in which homographs are embedded and/or specific task demands (e.g., Brenders, Van Hell & Dijkstra, 2011; Von Studnitz & Green, 2002). This has been taken to suggest that, at least with respect to out-of-context words, the role of shared semantics may be critical in determining the strength of co-activation (Schwartz & Van Hell, 2012). However, since the experiments carried out by Dijkstra et al. (1998), many studies have reported similar effects of language co-activation in a variety of contexts (for a review, see Dijkstra, 2005; Schwartz & Van Hell, 2012). These are not only seen at the behavioral level, but also using neurophysiological techniques (i.e., event-related-potentials; ERPs) that can map out the time course of these processes. For example, Midgley, Holcomb, and Grainger (2011) used a semantic categorization task with English native speakers who were learners of French. Participants were instructed to read a word on the screen and press a button whenever they saw a word referring to an animal. Signature brainwaves revealed a cognate facilitation effect (as evidenced by a differential negativity around 300-400 ms following stimulus onset) not only in the L1, but also in the weaker L2, suggesting that language co-activation can take place at lower levels of proficiency. Similar findings have also been found in highly proficient bilinguals whose languages do not share the same writing system (Thierry & Wu, 2007), and even in cases where the language modalities are different (Morford et al., 2011).

Despite the robust evidence of parallel activation found with word-recognition tasks, a question that remains is whether non-selectivity is also manifested when words are read in a sentential context. One could argue that the out-of-context effects are a result of increased lexical ambiguity when presenting words in isolation. But as it turns out, sentence context does not seem to eliminate the effect of lexical co-activation for bilinguals (Schwartz & Van Hell, 2012). In a study by Duyck, Van Assche, Drieghe, and Hartsuiker (2007), Dutch-English bilinguals performed a simple sentence reading task, in addition to a series of lexical decision tasks, in their L2. For the sentence reading task, critical words (i.e., cognates and non-cognates) were embedded in a sentence context (e.g., I would like you to repeat that *dance* (cognate) /*smile* (non-cognate)). As in previous word-recognition studies, there was a cognate facilitation effect (i.e., faster reading times in the critical region) relative to non-cognates. Later studies also replicated these effects by having bilinguals read in their L1, which presumably is more resistant to influences from the non-dominant language (Van Assche, Duyck, Hartsuiker, & Van Diependaele, 2009). Overall, language-specific cues (e.g., phonology) do not seem to be sufficient to eliminate the effect of language co-activation (Chambers & Cooke, 2009). It is only when sentences are highly constrained semantically (i.e., when the target final word becomes highly predictable given the semantic characteristics of the sentence) that both cognate effects and homograph effects seem to disappear (Elston-Güttler, Gunter, & Kotz, 2005; Libben & Titone, 2009; Schwartz & Kroll, 2006).

1.2. Parallel activation at the lexical level in production and the selection problem.

Despite the fact that speech planning is presumably initiated by the speaker, studies of bilingual word access in production also show effects of cross-language activation (Kroll, Bobb, & Wodniecka, 2006; Kroll, Bobb, Misra, & Guo, 2008). Many of these studies have adopted similar strategies to those used in word comprehension studies, particularly with regards to the presence of language-ambiguous words. Costa, Caramazza and Sebastián-Gallés (2000), for example, had Catalan-Spanish bilinguals and Spanish monolinguals perform a simple picture naming task in which critical trials contained pictures corresponding to cognates. Like in the word comprehension studies, the results showed a cognate facilitation effect in both of bilingual's languages (i.e., faster naming in cognate trials relative to non-cognate trials), although the effect was strongest when naming in the L2. Crucially, monolinguals did not show this effect. A similar study by Hoshino and Kroll (2008) also found cognate facilitation for Spanish-English and Japanese-English bilinguals (whose writing systems are different) while naming pictures in their L2. Other production studies using variants of the Stroop task have also found effects of parallel activation in the form of interference (i.e., slower naming) when distractor words are semantically related to the picture, and in the form of facilitation when distractor words have phonological properties related to the word to be produced (e.g., Costa, Miozzo, & Caramazza, 1999; Hermans, Bongaerts, De Bot, & Schreuder, 1998). Similar effects are seen when bilinguals are asked to produce the translation of a word (e.g., La Heij, De Bruyn, Elens, Hartsuiker, & Helaha, 1990; Miller & Kroll, 2002). Like the word recognition data, the evidence on bilingual speech production tasks suggests that the two languages are activated and available, and that the spread of co-activation can extend all the way to the phonological level, even in the absence of orthographic overlap.

The co-activation of both languages during speech planning presents a problem for bilinguals, since they rarely make the error of speaking the unintended language (e.g., Gollan et al., 2011), and yet they are able to switch languages (i.e., codeswitch) on the fly within a conversation (e.g., Muysken, 2000; Myers-Scotton, 2002). This suggests that bilinguals must develop a unique ability to control the selection of each language in accordance to the interactional context (Green & Abutalebi, 2013). By and large, most studies suggest that the solution to the selection problem must come from a mechanism that engages active inhibitory control (e.g., Kroll, Bobb, Misra, & Guo, 2008). Green (1998) proposed the Inhibitory Control (IC) model to account for the nonselective nature of the language production system in bilinguals, while at the same time ensuring proper language selection via the orchestrated use of language tagging, task-schemas, and reactive inhibition. In this sense, the IC model is unique in that it proposes a domain-general mechanism that is responsible for language selection. While there is strong neurophysiological evidence supporting the IC model (Abutalebi & Green 2007, 2008), there is also an emerging literature supporting the idea that there are multiple independent components of inhibitory control that may be differentially exploited by bilinguals depending on the language context in which the task is performed (Abutalebi et al., 2008), the degree of task difficulty (Costa et al., 2009), and the degree to which participants are better able to adjust the use of these inhibitory components in order to maximize monitoring performance (Morales, Gómez-Ariza, & Bajo, 2013). Taken together, these studies suggest that there may be unique cognitive demands imposed when bilinguals have to produce or use a single word as a consequence of parallel activation, and that the way these demands are used may depend on the task demands. However, it is still not clear whether co-activation is present at higher levels of language activation (i.e., at the level of grammar), and whether it has any consequences for processing grammatical structures.

1.3. Parallel activation at the level of grammar and its relation to lexical activation.

Studies that show evidence for grammatical cross-language co-activation/interaction come primarily from two sources. One has been to examine the global consequences of parallel activation in sentence processing. For example, Dussias and Sagarra (2007) examined the resolution of structural conflicts across the two languages of bilingual speakers using syntactically ambiguous relative clauses. They had Spanish monolinguals and Spanish-English bilinguals (with little or extended English immersion) read sentences that were syntactically ambiguous with regards to relative clauses constructions. These relative clauses contained two potential attachment sites (e.g., *arrestaron a la hermana del carnicero que estaba divorciada*/'Someone arrested the sister of the butcher$_{MASC}$ who had been divorced$_{FEM}$'). With respect to these attachments sites, previous research has shown that Spanish speakers prefer to attach the ambiguous relative clause (que estaba divorciada/who was divorced$_{FEM}$) to the first noun within the Noun Phrase (e.g., hermana/sister), which is also known as high attachment. In contrast, English speakers typically prefer to attach the relative clause to the second noun of the Noun Phrase (carnicero/butcher), or low attachment. For monolinguals and bilinguals with little immersion experience, results showed a preference for high attachment (the first noun). Bilinguals with extensive immersion in an English-speaking environment showed a preference for low-attachment (the second noun), suggesting that exposure to English may have altered their parsing preference. These results show that bilingual's grammatical representations are not fixed and can be highly interactive across languages. This suggests a dynamic language system that can change in response to language exposure (e.g., MacDonald & Seidenberg, 2006; Gennari & MacDonald, 2009).

Another approach that has been used to address the issue of grammatical cross-language interactions is syntactic priming, a phenomenon in which the appearance of a specific syntactic structure facilitates the subsequent production or processing of that structure (e.g., Bock, 1986). Research on syntactic priming in bilinguals has shown that the syntactic representations of two languages may utilize partially overlapping representations (for a review, see Hartsuiker & Ferreira, 2008; Schoonbaert et al., 2007). Cross-language priming has been found for structures that overlap in linear word order. For example, a seminal study by Hartsuiker, Pickering and Veltkamp (2004) had Spanish-English bilinguals perform a picture description task with another participant who was also a Spanish-English bilingual. In this task, participants took turns describing pictures using a sentence. Crucially, one of the participants (a confederate) was scripted to read and use sentences in Spanish using either an active (e.g., *el taxi persigue al camión*/the taxi chases the truck), passive (*El camión es perseguido por el taxi*/The truck is chased by the taxi), intransitive (e.g., *El taxi acelera*/The taxi accelerates), or OVS (*El camión lo persigue un taxi*/ The truck (chasee) it chases a taxi (chaser)) construction. The actual participant was instructed to describe the pictures in English. Hartsuiker et al. (2004) found that, immediately after hearing a passive in Spanish, participants used more passive constructions than in conditions where the passive was not used by the confederate. They took this as evidence that syntactic representations can also be integrated across languages. This was also taken to propose a model that accounts for the integrated nature of bilingual language representations, including syntactic information. In this model, the construction of a sentence frame is lexically driven. Therefore, grammatical information associated with the lexical representation will guide the construction of the frame. For bilingual speakers, grammatical (i.e., lexico-syntactic) representations are connected via the same categorical and combinatorial nodes. Activating a given lemma leads to the activation of a grammatical structure, which is unspecified for language. This model highlights the interactive nature of bilinguals' grammatical constructions during speech production while also accounting for the interconnection between lexical units and syntactic units. However, it may also be possible that the nature of verbs is not as lexically based as other kinds of

One limitation of the syntactic priming studies is that it makes it challenging for characterizing parallel activation in contexts where bilinguals experience both facilitation and conflict during online speech planning. After all, bilinguals use structures that sometimes converge and sometimes compete. A recent study by Runnqvist, Gollan, Costa, and Ferreira

(2013) addressed this issue by asking highly-proficient Spanish-English bilinguals and Mandarin-English bilinguals to perform a sentence production task in English (their dominant language at the time of testing). Participants were probed to generate sentences using specific syntactic structures (i.e., active/passive structures, or, pre vs post-modified noun phrase alternations). They found that bilinguals were overall slower in generating sentences, but that this slowdown was modulated by whether the structures were more or less frequent in the other language, and by whether there was linear word order overlap across the two languages. The authors concluded that these findings supported the weaker links hypothesis (Gollan, Cera, & Sandoval, 2008). Because bilinguals speak each of their languages less often than monolinguals, they become disadvantaged in speaking, resulting in slower access to lexical (Gollan et al., 2011) and syntactic (Runnqvist et al., 2013) information. However, it is possible that these effects also depend on how cognitive (inhibitory) demands are used (Kroll & Gollan, 2014), on which bilingual population is studied, and on how the task demands in a given situation interact with the demands imposed by language-specific features. Following the latter point, it is possible that lexical information imposes particular demands on how syntactic information is accessed. A recent study by Hopp (2015) suggests that lexical processing can induce non-native syntactic processing during sentence reading in L2 learners. However, this is a domain that has not been systematically studied in highly proficient bilinguals in production, with the exception of a study by Gullifer (2015) using syntactic priming. One way of addressing this issue is by investigating how bilinguals negotiate the use subtle lexical-syntactic constraints in their L2. To motivate this issue, we adopt a usage-based linguistic theory known as Construction Grammar, which is discussed below.

1.4. A usage-based model of language constraints.

Usage-based approaches argue that language is an emergent property of cognition and experience. In this sense, grammatical constructions should not be taken as abstract rule-based features. Instead, they should be taken as form-meaning mappings, which are employed through language use and become conventionalized in the speech community (Bybee, 2010, 2013; Goldberg, 2013). They are language units relating the defining properties of their morphological, syntactic, and lexical form with particular semantic, pragmatic, and discourse functions (Bates & MacWhinney 1987; Bybee, 2008, 2010; Goldberg 1995, 2006, 2012; Lakoff 1987; Langacker 1987). Because of this, constructions are presumed to arise from cognitive domain-general mechanisms, which can in turn impact our linguistic representations. Some of these domain-general processes include memory storage, chunking, analogy, categorization (Bybee, 2010), and Statistical Preemption (Boyd & Goldberg, 2011).

For the most part, spoken language is constrained by the meaning we want to convey or by grammatical features. In certain cases, however, a given formulation can be 'legal' on semantic and grammatical grounds, and yet strongly dispreferred. For example, even though the vast majority of English adjectives appear in prenominal positions (e.g., the *sleepy* boy), there are some adjectives that disprefer prenominal use (e.g., 'the *asleep* boy' is dispreferred over 'the *sleepy* boy'). These adjectives, known as *a*-adjectives, begin with a syllabic schwa ('a'), are morphologically segmentable (*a* + stem), and are no longer than two syllables (e.g., a/sleep, a/live, a/float; Goldberg, 2011). How do speakers know to avoid using these adjectives in prenominal position? The statistical preemption hypothesis proposes that the consistent witnessing of a form (the boy that's *asleep*/the boy is *asleep*) in the input blocks the use of a novel form (the *asleep* boy) when both alternatives share similar meaning and/or function (Goldberg, 1995).

Evidence for the statistical preemption hypothesis has been found in corpora with argument structure constructions (i.e., with verbs; Goldberg, 2011). A recent study by Boyd and Goldberg (2011) sought to investigate the role of statistical preemption and categorization in native English speakers within an experimental setting. To do this, they conducted three speech production experiments (only two are discussed here) in which the use of a-adjectives was manipulated. In the three experiments, participants were shown pairs of pictures that had a label underneath, some of which were real a-adjectives (e.g. asleep) or non a-adjectives (e.g., sleepy), and some of which were novel labels that resembled real a-adjectives (e.g., ablim) or non a-adjectives (e.g., chammy). A target picture performed an action onscreen (i.e., moved towards an object) and participants were asked to produce a sentence using the target's label. Before performing this part of the experiment, participants were introduced to an exposure block where they saw sentences containing novel adjectives (e.g., the *tooky* frog moved to the star). In the first experiment, Boyd and Goldberg investigated the extent to which the a-adjective constraint exists, and whether it can induce generalization of the constraint into novel words. To do this, they manipulated the exposure block so that participants only saw novel non-a-adjectives (*tooky*). The results showed that speakers did in fact avoid using real a-adjectives prenominally, and that novel a-adjectives followed the trend of avoiding prenominal use as well, although the latter effect was much weaker. The authors concluded from this experiment that a-adjective category is in fact productive by exhibiting effects of categorization (i.e., applying the constraint to new words). For the second experiment, the authors wanted to address whether the results in the first experiment could be moderated by statistical preemption. To do this, a preemptive context was introduced in the exposure block. In other words, prior to performing the production task, participants witnessed some instances of novel a-adjectives being used in a relative clause construction (e.g., the fox that's *ablim* moved to the star). As a result, the use of relative clause constructions for real and novel a-adjectives dramatically increased (i.e., prenominal uses went down), suggesting speakers can fully generalize the constraint via SP and Categorization.

1.5. The Present Study

In the present study, we adopt the paradigm by Boyd and Goldberg (2011) to examine English sentence-production abilities in two bilingual groups for whom English is the L2: Spanish-English bilinguals (immersed in their L2) and German-English bilinguals (immersed in their L1). First, we investigate whether language experience (i.e., bilingualism) modulates the use of subtle syntactic constraints in the L2 (English) via a-adjectives. Following the weaker links hypothesis (Gollan et al., 2008), it is possible that bilinguals would produce a less stable preference pattern for a-adjectives and therefore would be less able to generalize the constraint into novel words, given that they have less experience in each language. On the other hand, previous research has shown that bilinguals outperform monolinguals in a wide variety of tasks involving executive functions (e.g., Costa, Hernández, and Sebastián-Gallés, 2008) and metalinguistic skills (e.g., Bialystok, 2001; Bruck & Genesee, 1995; Galambos & Goldin-Meadow, 1990), presumably because the constant use of two languages imposes additional demands on these aspects of cognition. Following the latter rationale, we hypothesized that, regardless of the language pair, bilinguals as a whole would exhibit at least comparable patterns to monolinguals (as shown by Boyd and Goldberg, 2011), and that if task difficulty was increased, bilinguals would remain more sensitive to these constraints, produce a more stable pattern relative to monolinguals, and exhibit greater generalization patterns. To address this, we adjusted the format of Experiment 1 in Boyd & Goldberg (2011) by imposing time constraints, by making the trial sequences in the production block less predictable, and by eliminating specific cues in the exposure block (see the materials and procedure sections below).

Second, we investigate how the demands imposed by language-specific features affect processing speed of grammatical information. This yields two predictions: If bilinguals experience a frequency-lag in their L2 (Gollan, et al., 2008), then both groups should be slower overall than monolinguals in producing correct sentences. If the effects of facilitation and interference reflect inhibitory control processes that resolve cross-language competition (Kroll, Bobb, & Wodniecka, 2006), then differences in naming speed should emerge between the two bilingual groups in a systematic way. In one scenario, we predicted that German-English bilinguals would be the slowest of the groups tested because of their immersion status in their L1, which would yield global costs for suppressing the L1 in order to speak in the L2. Conversely, Spanish-English bilinguals who have been immersed in the L2 context for several years may show facilitation in the task. We refer to this as the *global* inhibitory account. Such an account is motivated by previous studies showing that L2 language immersion has consequences for how the L1 is used and processed (e.g., Linck, Kroll, & Sunderman, 2008). In a second scenario, we predicted Spanish-English bilinguals to be the slowest group because the task requires the use of prenominal and relative clause constructions and because Spanish also makes use of postnominal constructions (e.g., *el carro verde* literally translates to *the car green*), Spanish-English bilinguals may experience syntactic interference at the onset of planning. German-English bilinguals, on the other hand, may experience less interference, because there is more overlap between English and German in the sense of using structures (prenominal and relative clauses) that are highly salient in both languages.

A third and final issue that we investigate in the present study is whether lexical information influences access to syntactic information during speech planning. A-adjectives are an ideal case for testing this because: 1) they are more representative of typical lexical items (unlike verbs, which can be conceived as being more 'grammatical'); and 2) they also give rise to a constraint that is ultimately syntactic (avoiding prenominal use vs. preferring predicative constructions). We predicted, then, that sentences containing a-adjectives would generate slower responses, even when compared against sentences containing the same syntactic information (i.e., a relative clause), but with a different lexical item (i.e., verbs).

2. Methods

2.1. Participants

Thirty-two native English monolinguals participated in the study. Two participants were excluded due to technical problems with the equipment during testing. All participants were functionally monolingual in English and were recruited from the Psychology subject pool at The Pennsylvania State University.

Fourteen Spanish-English bilinguals were recruited at the Pennsylvania State University to participate in the study. One participant was excluded from the analyses given that she reported English as her native language and showed low proficiency in Spanish, and a second participant was also excluded due to technical problems with the equipment during the main task. The remaining twelve participants reported having Spanish as their first or native language, and had moved to the mainland US within the last five years. Ten out of the twelve participants were born and primarily raised in Puerto Rico, while the other two were born and raised in Mexico and Colombia. All participants began formal English training in primary school up until the end of secondary school, although they also reported frequent exposure to English during their first six years of age. Participants received $10 for their participation in the session.

Sixteen German-English participants were recruited at the University of Mannheim in Germany to also participate in the study. All participants reported German as their first or native language, and began learning and using English formally between the ages of six and thirteen. Participants received the equivalent of €10 for their participation in the session.

All participants gave informed consent, and the procedures had the approval of the Institutional Review Board of the Pennsylvania State University. Participants completed a language history questionnaire to assess self-rated levels of language proficiency. Self-ratings were made on a scale of 0 to 5, with 0 indicating very limited proficiency and 5 indicating very high proficiency. Additionally, a verbal fluency task was used to assess production abilities. In this task, participants were presented with a series of categories (e.g., animals, fruits, colors) one at a time, and are prompted to generate as many category exemplars as possible within 30 seconds. Monolinguals performed the task in English only, while Spanish-English bilinguals and German-English bilinguals performed blocks of Spanish/English or German/English, respectively. The order of the language blocks was counterbalanced across participants. Finally, we used an Operation-Span task (La Pointe & Engle, 1990) to assess working memory ability.

Table 1 shows the descriptive characteristics of each group on the basis of the tasks described above. Some participants contain missing data for the O-span due to equipment malfunction during the task (see Table 1 for valid *N*). A series of one-way ANOVAs was performed to determine whether the groups were matched on each these characteristics, with group as the between-subjects factor. Any significant main effects, where appropriate, were followed-up with post-hoc analyses using the Bonferroni correction.

There were overall differences in the mean age of the three groups, F (2, 55) = 11.506, $p < .001$, such that monolinguals were significantly younger than both Spanish-English bilinguals (p = .026) and German-English bilinguals ($p < .001$). The bilingual groups did not differ in their mean age (p = .594). Self-ratings indicated no differences between Spanish and German proficiency status (ps > .309), but there was an overall differences in English proficiency in production, F (2, 55) = 20.782, $p < .001$, and in comprehension, F (2, 55) = 7.736, p = .001, such that German-English bilinguals rated themselves lower than the other two groups (ps < .01). However, the Verbal Fluency scores (average number of exemplars across categories) suggest a slightly different picture: All three groups performed similarly in English (p = .201), but German-English bilinguals scored higher in their L1 (German) relative to Spanish-English bilingual L1 (Spanish) scores, F (1, 26) = 16.676, $p < .001$.

Table 1. Participant characteristics

	Monolinguals		Spanish-English Bilinguals		German-English Bilinguals	
	Valid *N*	***M*(SD)**	**Valid *N***	***M*(SD)**	**Valid *N***	***M*(SD)**
Age (in years)	30	18.57 (1.22)	12	21.25 (3.49)	16	22.69 (4.32)
L1 Self-Ratings[1]						
Production	30	5 (0)	12	4.92 (.29)	16	4.56 (1.21)
Comprehension	30	5 (0)	12	4.92 (.29)	16	4.50 (1.37)
L2 Self-Ratings						
Production	NA	NA	12	4.66 (.49)	16	3.94 (.93)
Comprehension	NA	NA	12	4.92 (.29)	16	4.31 (1.08)
Verbal Fluency[2]						
English	30	47.70 (8.37)	12	48.17 (9.52)	16	43.38 (7.45)
Spanish or German	NA	NA	12	41.58 (7.81)	16	53.69 (7.73)
O-span score (out of 60)	28	44.82 (9.15)	12	41.08 (5.25)	16	43.44 (8.71)

[1]Self ratings were made on a scale of 0 to 5, with 0 indicating very limited proficiency and 5 indicating very high proficiency.

[2]Verbal Fluency scores reflect the average number of exemplars generated across categories.

2.2. Materials

The materials used for the main task were taken from Boyd and Goldberg (2011) and adapted into E-Prime. The experiment consisted of an exposure block and a production block. In the exposure block, participants watched the experimenter go over several trials to ensure understanding of the task. In the production block, participants were asked to complete the trials themselves. For every trial within both blocks, pairs of animal images were presented onscreen, each of which contained a label underneath. One of these labels was a target label, and the other was a foil label. The image with the target label always performed an action on the screen (e.g., a fox moves towards an apple). After each trial, a question appeared onscreen (e.g., *which fox moved to the apple?*), which the experimenter (in the exposure block) and participant (in the production block) had to answer out loud using the target label within a sentence.

There were a total of 12 trials for the exposure block, and a total of 32 trials for the production block. All of the trials in the exposure block contained novel labels only. Half of the trials in the production block (16) contained a target label corresponding to an experimental condition. These target labels were manipulated for Adjective Novelty (real vs. novel) and Adjective Class (a-adjective vs. non a-adjective). This resulted in a 2 (Adjective Novelty) by 2 (Adjective Class) design, meaning that each of these experimental conditions had four target labels total (see Table 2). The remaining half of trials (16) consisted of fillers (see Table 3). Eight of these filler trials contained third-person singular verbs (e.g., smokes, runs) to encourage use of relative clause construction, and to guard against other spurious effects such as syntactic priming. The remaining filler trials contained high frequency adjectives that are typically found in prenominal/attributive constructions (e.g., slow, fast).

Participants were randomly assigned to view one of four counterbalanced versions of the task. Both the exposure and production block trials were pseudo-randomized across each counterbalanced version. For the exposure block, all of the trials consisted of novel labels. Crucially, target labels never appeared in the production block (e.g., gorpy, floggy) and always occurred in prenominal position when presented using a sentence onscreen (e.g., *the floggy bird moved to the apple*). Note that the latter manipulation was not implemented in the original study by Boyd and Goldberg (2011), in which half of the novel non a-adjectives occurred in a relative clause (e.g., *the bird that's floggy moved to the apple*). This manipulation was included to further bias participants to use prenominal constructions with adjectives. Thus, if participants used relative clause constructions, then it would make for a stronger case that they were employing knowledge of the a-adjective constraint. On the other hand, the first trial in the production block was always a filler trial. Afterwards, critical and filler trials always alternated between one another. There were no more than two fillers of the same type (adjective labels vs. verb labels) that occurred consecutively. There were no more than two critical trials from the same condition that occurred in a row. Each trial was counterbalanced in such a way that half of the target labels appeared on the left, and half on the right.

Table 2. List of target labels for critical trials by condition. Foil labels are in parentheses.

Real		**Novel**	
A-adjectives	**Non A-adjectives**	**A-adjectives**	**Non A-adjectives**
afloat (sinking)	*floating (sinking)*	*ablim (zecky)*	*chammy (zoopy)*
afraid (brave)	*frightened (brave)*	*adax (zedgy)*	*flitzy (zappy)*
alive (dead)	*living (dead)*	*afraz (zibby)*	*gecky (zunderful)*
asleep (vigilant)	*sleepy (vigilant)*	*agask (zintesting)*	*slooky (zinky)*

Table 3. List of target labels for filler trials. Foil labels are in parentheses.

3rd-person-singular verbs	**High frequency adjectives**
bites (votes)	*bad (good)*
cheats (lies)	*fast (slow)*
gambles (smokes)	*good (bad)*
golfs (snowboards)	*old (young)*
lies (cheats)	*slow (fast)*
smokes (gambles)	*strong (weak)*
snowboards (golfs)	*weak (strong)*
votes (bites)	*young (old)*

2.3. Procedure

Each participant was tested in a sound-attenuated room individually. They were seated in front of a computer monitor that was connected to a button box and digital recorder. Participants first completed the LHQ, followed by the main task, the verbal fluency, and operation span task. At the beginning of the main task, participants were carefully briefed on the experimental procedure, and were told to only use labels exactly as they appeared onscreen. Participants were also encouraged to use full sentences (e.g., *the fox that's asleep moved to the apple*) and avoid saying sentence fragments only (e.g., *the fox that's asleep*). These points were emphasized in six practice trials in which participants received direct feedback from the experimenter. The first three practice trials were performed by the experimenter, while participants were encouraged to perform the last three trials. Participants were then told that some trials would contain words that they had never seen before, and were encouraged to use them as if they were real words.

The experimenter then proceeded to the exposure block, during which he performed the task as if he was the actual participant. At the beginning of each trial, the first pair of animal images appeared on the screen for two seconds (see Figure 1). Here participants were told that they would have to say both of the labels (target and foil) out loud before disappearing. The experimenter made sure to do this himself during the exposure block to ensure that participants understood this component of the task. The next sequence showed the target image performing an action within one second. Actions performed by the target image were constantly changing for every trial. That is, sometimes the target image moved towards an object (or simply towards a location without reference); other times the image shrunk or enlarged in size; and at other times it disappeared. This was done to make the task less predictable, thus forcing participants to remain attentive throughout the task. Finally, a question appeared in the middle of the screen for five seconds. For the exposure block only, answers were provided underneath the question and the experimenter read them out loud. After the question disappeared, a fixation sign appeared at the center of the computer screen indefinitely. Participants were told that this was the only time where they would have to press a key to proceed to the next trial. After the exposure block, the experimenter left the room to begin the production block. Here participants had to formulate their own descriptions on each trial. Two microphones were used to record both accuracy and reaction times, respectively.

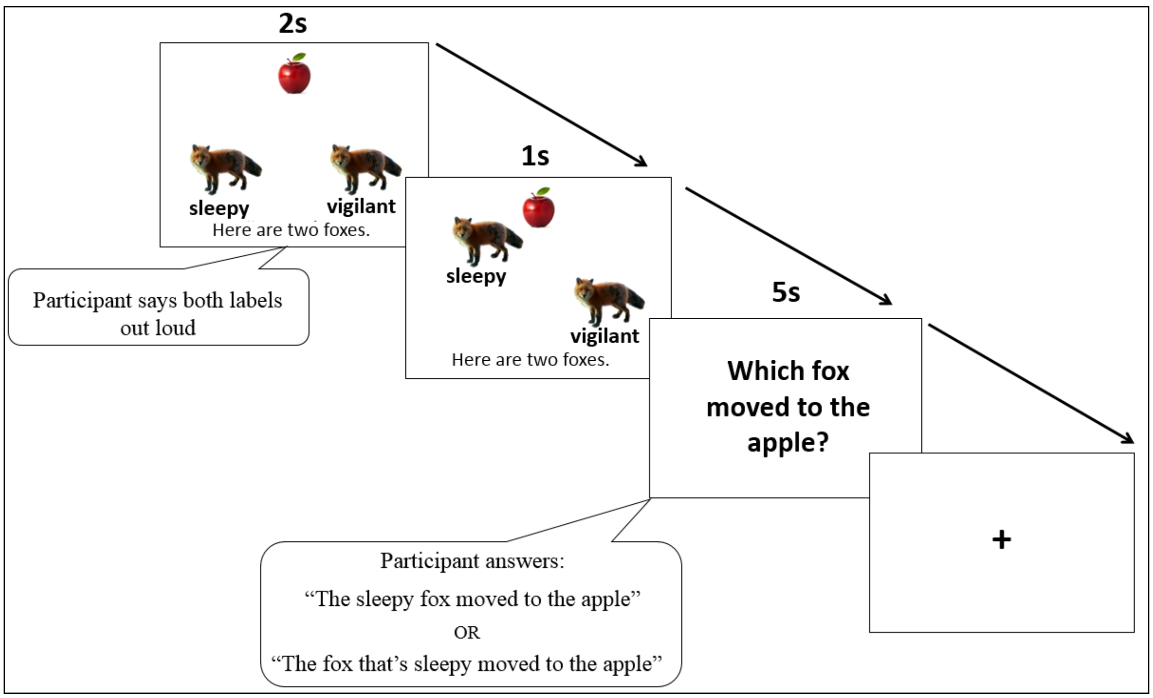

Figure 1. Illustration of a sequence of events in a single trial with the expected response from the participant (shown in the callout box).

3. Results

3.1 Analysis of proportion (accuracy) data

Sentence responses were coded for accuracy and for whether the target label was used prenominally or in a relative clause. Responses were considered incorrect if the participant modified, and/or used the wrong the label construction (e.g., *The fox that's cool and asleep moved to the apple*), if the formulation was incomplete (e.g., *The fox that's asleep…*), or if there was no response. False starts, reformulations, and hesitations were included in this part of the analysis as long as the response was not incorrect. For the monolingual group, there were a total of 480 critical trial responses. Eleven critical trials that contained incorrect responses (i.e., using foil labels when producing the formulation, incomplete formulations, or no response) were excluded from the analysis. For the bilingual group, there were a total of 448 critical trial responses, and nine were excluded due to incorrect responses.

This dataset was analyzed using logistic mixed effects models in the R computing language and environment (R Development Core Team, 2010). Logistic regression analyses are ideal for categorical dependent variables that have a binary outcome (Jaeger, 2008), and the mixed effects component properly accounts for within-subject variance, even in the absence of data points. In our model, the dependent variable was binomial (prenominal vs. relative clause placement), and responses that contained prenominal constructions were set as the reference level. Comparison of different models indicated that the best model fit was achieved when Adjective Novelty (real vs. novel), Adjective Class (a-adj vs. non a-adj), Group (monolinguals vs. bilinguals), and an interaction between Group and Adjective Novelty, were specified as fixed effects. The model also allowed random intercepts to vary by participant, and random slopes for

Adjective Class to vary by participant. Other random effects did not improve the model fit significantly. Because performance in this portion of the task was virtually identical across the two bilingual groups, they were collapsed as one group to gain more statistical power in making the group comparison against monolinguals[1].

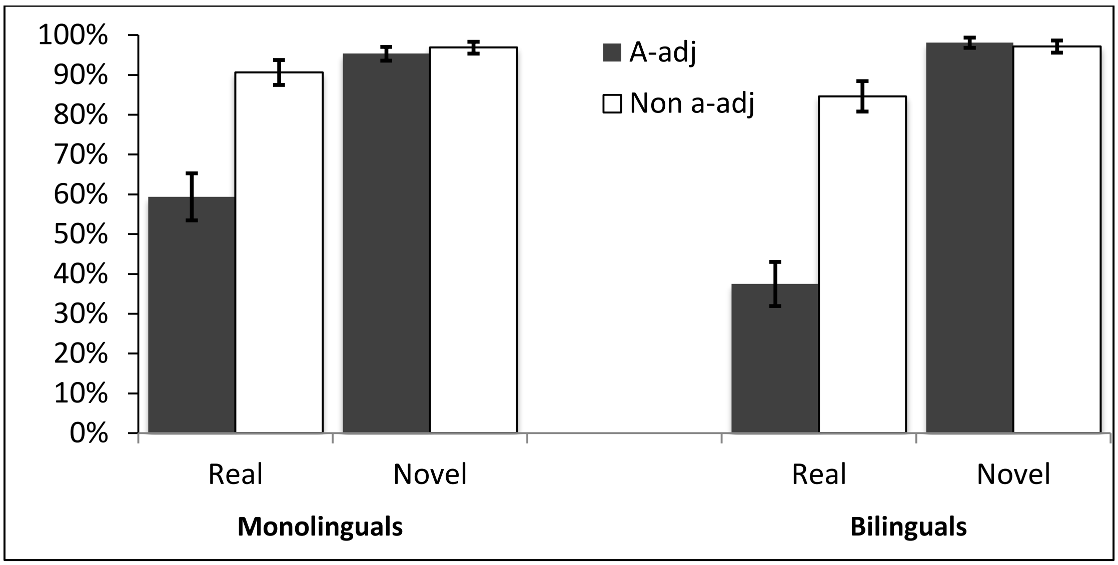

Figure 2. Percentage of prenominal use by condition and by group. Spanish-English bilinguals and German-English bilinguals are collapsed as one group. Error bars indicate standard error.

[1] An initial model comparing the three groups separately revealed a Group by Novelty interaction. A follow-up model only comparing Spanish and German bilinguals revealed no group interactions with any other factors.

Table 4. Parameters for mixed logistic regression model. Parameter estimates indicate the likelihood of a response using a prenominal construction.

# Observations: 908	Groups: 2	# Participants: 58			
Fixed Effects					
		Estimate	SE	*Z*	*p*
Intercept		2.67	1.51	1.77	0.076
Adjective Novelty = Real		-0.94	1.48	-0.64	0.525
Adjective Class = Non A		2.73	0.45	6.0	< 0.001
Group		1.28	1.17	1.09	0.273
Adjective Novelty*Group		-2.42	1.16	-2.09	0.037
Random Effects					
Groups		Variance	Std. Dev.		
Participant	Intercept	1.29	1.134		
	Adjective Class	1.17	1.084		

The final model is summarized in Table 4. Parameters estimates for the likelihood of a prenominal response are given in the form of log-odds space, but to make the results more accessible, they are interpreted in the context of the percentages of prenominal use shown in Figure 2. There was no main effect of adjective novelty, but a main effect of adjective class suggests that the likelihood of a prenominal response is higher with sentences containing non a-adjectives compared to sentence containing a-adjectives. No main effect of group was observed, but there was a significant interaction between group and adjective novelty. This suggests that, with respect to real words, the likelihood of a prenominal response was in general lower for bilinguals relative to monolinguals. To follow up on this interaction, we performed a series of simple effects analyses to test whether group differences emerged with real a-adjectives and/or real non-adjectives. These models had Group as the only fixed effect, while allowing for random intercepts to vary by participant and trial. With real a-adjectives, there was a main effect of

Group, β = –1.20, *SE* = .45, *z* = –2.656, *p* = 0.008, but no main effect was observed with real non a-adjectives, β = –1.05, *SE* = .63, *z* = –1.68, *p* = 0.09. This is reflected by the fact that the percentage of prenominal use with real a-adjectives is 59% for monolinguals and 38% for bilinguals, relative to real non a-adjectives which yielded 91% (monolinguals) and 85% (bilinguals).

To further investigate this difference between the two groups, filler trials were also coded for whether participants used prenominal or relative clause constructions. This would be particularly revealing given that filler trials containing 3rd-person-singular verbs were expected to be used with relative clauses at all times. Otherwise, participants would have had to produce an anomalous sentence (e.g., *the sings* fox moved to the apple*). But as it turns out, there were several participants who generated these kinds of errors (see Figure 3). We performed a subsequent simple effects analysis to address this issue. In this model, the dependent variable was again binomial (prenominal vs. relative clause placement), but here we only looked at fillers containing verb targets. This yielded a total of 364 items to analyze, 34 of which were excluded from the analysis for containing incorrect responses. The model also included Group (monolinguals vs. bilinguals) as a fixed effect, while allowing for random intercepts by participant and trial. The results showed a main effect of Group, β = –4.061, *SE* = 1.248, *z* = –3.252, *p* = 0.001, suggesting that the likelihood of a prenominal response with verbs is lower in bilinguals than in monolinguals. This is also reflected in Figure 3, showing an almost categorical distinction between the two groups with respect to the percentage of verb errors. Visual inspection revealed that the five monolinguals who generated these errors 100% of the time also prenominalized real a-adjectives 100% of the time. For monolinguals (N = 13) who generated these errors between 10% and 65% of the time, the percentage of prenominal use for real a-adjectives was 72%. For the rest of the monolingual participants (N = 12), the percentages go down to 36%.

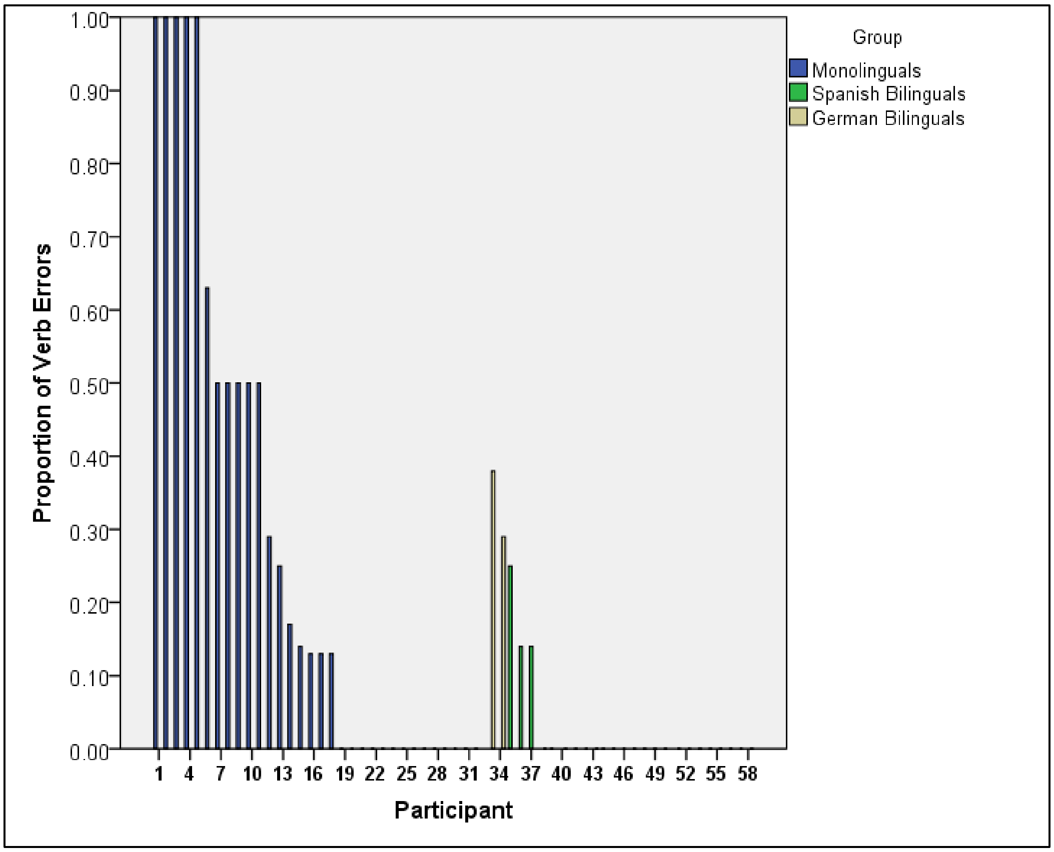

Figure 3. Proportion of verb errors (used prenominally) across all participants. Participants 31 through 58 correspond to the bilingual groups.

3.2. Analysis of latency data

Reaction time (RT) data were also analyzed to compare processing speed across conditions and groups. In addition to response errors, trials were excluded if they began with a hesitation, if they contained self-corrections, or if no response was detected by the microphone. For this analysis, we excluded participants that generated errors with filler (verb) trials more than 30% of the time (monolinguals: N = 11; German bilinguals: N = 1). We also excluded five monolinguals and one Spanish-English bilingual due to equipment malfunction. The remaining

RT dataset was cleaned using a procedure to remove absolute and relative outliers. First, correct trials that had RTs above 3000 ms or below 300 ms were excluded. Any remaining trials were excluded if they fell outside the 2.5 standard deviation range around the mean RT for each participant. Thus, 10% of the reaction time data across all conditions consisted of outlier responses. For this analysis, a repeated measures ANOVA was performed, with Adjective Novelty (real vs. novel) and Adjective Class (a-adjectives vs. non a-adjectives) as within-subject factors, and Group (monolingual vs. Spanish bilingual vs. German bilingual) as a between-subject factor.

Mean latency responses are shown in Figure 4. There was a main effect of Adjective Novelty, F (1, 37) = 5.935, p = .020, η_p^2 = .138, a main effect of Adjective Class, F (1, 37) = 11.739, p = .002, η_p^2 = .241, and an interaction between Adjective Novelty and Class, F (1, 37) = 4.325, p = .045, η_p^2 = .105. No group interactions were found, but there was an overall main effect of Group, F (2, 37) = 5.915, p = .006, η_p^2 = .242. Post hoc tests with a Bonferroni correction revealed that Spanish-English bilinguals were overall slower in producing correct sentences relative to monolinguals, p = .015, and German-English bilinguals, p = .011. No differences were found between monolinguals and German-English bilinguals. To break down the interaction, follow-up repeated measures ANOVAs were performed for each Adjective Novelty condition, with Adjective Class as the within-subjects factor, and Group as the between-subjects factor. In the real-word condition, there was a main effect of Adjective Class, F (1, 37) = 12.946, p = .001, η_p^2 = .259, suggesting that real a-adjectives (monolinguals: M = 1176.85; Spanish bilinguals: M = 1571.42; German bilinguals: M = 1139.38) were overall slowest to produce than real non a-adjectives (monolinguals: M = 1012.95; Spanish bilinguals: M =

1335.03; German bilinguals: *M* = 1038.94). No main effects were found in the novel-word condition.

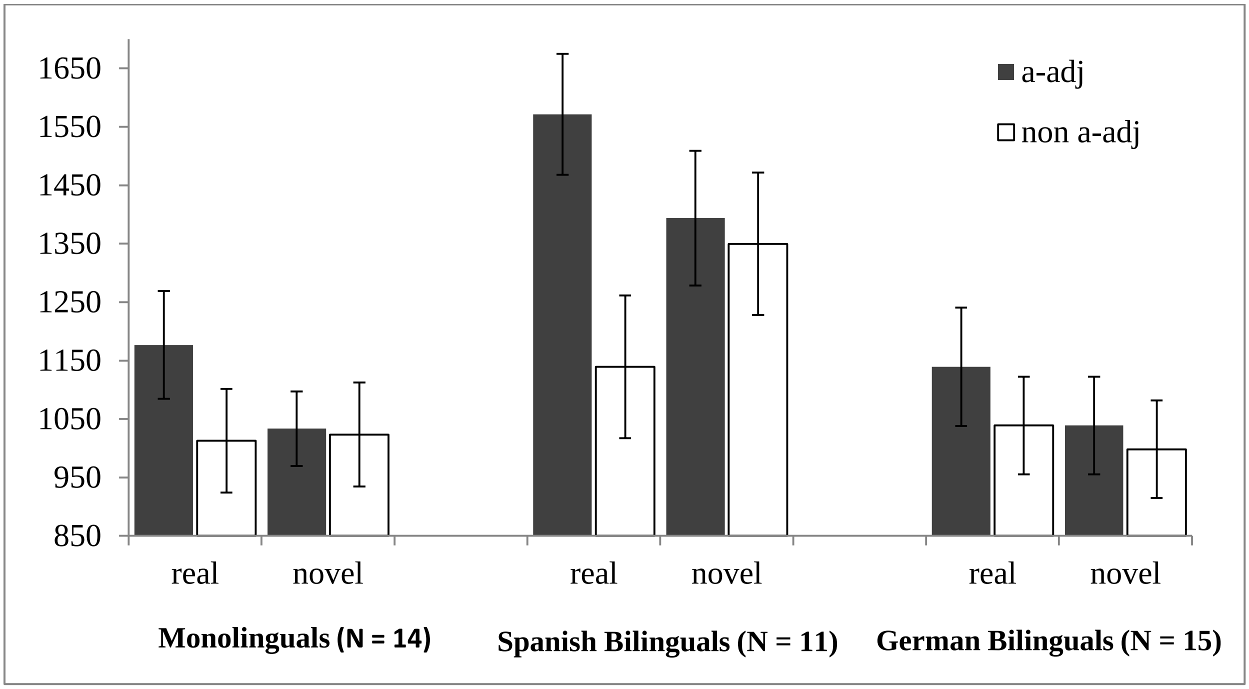

Figure 4. Mean latency (ms) at the onset of sentence production by condition and by group. Error bars represent standard error.

To investigate whether lexical access also influences processing speed in sentence planning, we performed a subsequent repeated measures ANOVA including Word Class (real a-adjective vs. verbs) as a within-subjects factor, and Group as a between-subjects factor. For sentences with real a-adjectives, we included responses containing both prenominal and relative clause constructions. For sentences with verbs, we only included correct responses (i.e., with a relative clause construction). If anything, this would bias the results against the idea that lexical information interacts with syntactic information, given that prenominal responses are presumably planned faster. Mean latency responses for these results are shown in Figure 5. There was a main effect of Word Class, F (1, 37) = 7.014, p = .012, η_p^2 = .159, suggesting that sentences containing real a-adjectives were overall slower to produce than sentences containing verbs. There was no Group by Word Class interaction, but again, there was a significant main effect of Group, F (2, 37) = 5.827, p = .006, η_p^2 = .240. Post hoc analyses with a Bonferroni correction revealed that Spanish-English bilinguals were the slowest to produce sentences relative to monolinguals, p = .011, and relative to German-English bilinguals, p = .002.

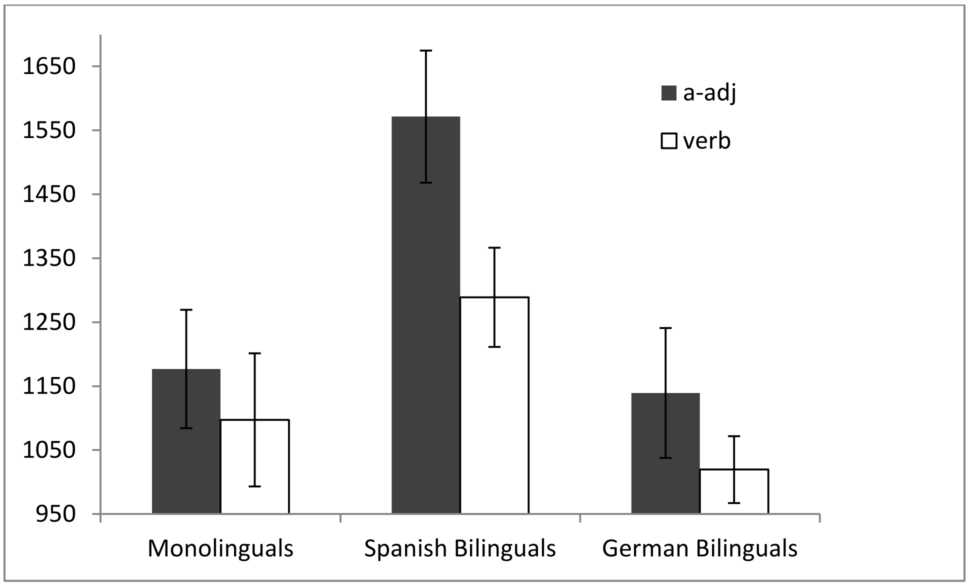

Figure 5. Mean latency (ms) at the onset of sentence production for real a-adjectives and verbs by group. Error bars represent standard error.

4. General Discussion

4.1. Does bilingualism modulate the use of a-adjectives?

In the present study, we investigated the extent to which speaking two languages affects how bilinguals plan and use subtle grammatical constraints in the L2. The results showed that both monolinguals and highly proficient bilinguals are sensitive to the a-adjective constraint, replicated the main pattern of results reported by Boyd and Goldberg (2011) with respect to the real-word condition. However, the results also showed that the magnitude of this pattern was modulated by group. Specifically, bilinguals as a whole were more stable in their preference for these patterns, to the point where the percentages with real a-adjectives were virtually identical to those found in Boyd and Goldberg (2011). In contrast, monolinguals showed remarkable variability in the extent to which they avoided prenominal use with real a-adjectives. A closer inspection in the dataset revealed that about half of the monolinguals systematically generated errors with filler trials that contained 3rd person singular verbs as target labels (e.g., *the sings* fox moved to the apple*), and that an increasing use of these errors may have led to more prenominalization with respect to real a-adjectives. A similar pattern was observed with a few bilinguals (N = 5), although none of them generated these errors 100% of the time. The fact that both monolinguals and bilinguals exhibited these unexpected patterns suggests that the degree of task difficulty was to some extent responsible for interfering with the level of access to the actual language representations. Recall that, unlike Boyd & Goldberg (2011), the exposure block in our study only contained examples of prenominal constructions. Although these examples were only for novel words, they may have introduced a strong bias towards prenominalization in general. Other manipulations that we introduced in this study was the time allowed to generate sentences, and the inability to predict picture/target actions across trials. All of these components presumably introduced additional demands that interfered with the access and planning of correct sentences. This yields another important point: the fact that both bilingual groups behaved similarly in this respect suggests that they were able to overcome these demands more readily via the use of distinct cognitive resources.

The question then is: which factors may have helped bilinguals overcome these difficulties? One possibility is that bilinguals simply had enhanced metalinguistic knowledge

(Bialystok, 2001). Under this view, bilinguals showed a more stable preference for using a-adjectives because they have stronger language representations in both languages. While this is a plausible explanation, and it was one that we had originally adopted to generate predictions, the presence of the error-data suggests that the differences between the two groups are primarily driven by task-specific demands, and not necessarily because of prior language knowledge. A second possibility is that bilinguals performed the task differentially using a more proactive strategy, whereas monolinguals (at least half of them) may have used a more reactive strategy that led them to prenominalize more often. Proactive control has been hypothesized to reflect the ability to engage predictive functions in order to anticipate upcoming events, while reactive control may instead engage processes that enable selection of alternatives that have been already activated (Morales, Gómez-Ariza, & Bajo, 2013). One prediction stemming from this account is that, if it is the case that cognitive control is actively being used, then speakers who show more proactive abilities in a non-linguistic task (e.g., AX-CPT; Rosvold, et al., 1956) would be predicted to generate the least amount of errors (and more stable patterns for real a-adjectives). Overall, the cognitive control account would seem more directly relevant to the current results, although neither of these explanations should be taken as mutually exclusive accounts.

Another interesting, but unexpected, outcome of this study was that neither bilinguals nor monolinguals generalized the a-adjective constraint to novel a-adjectives, with both groups using them prenominally across the board (e.g., *the ablim fox*). In contrast, the results of Experiment 1 in Boyd and Goldberg (2011) showed that, without a preemptive context, speakers *partially* generalized the a-adjective constraint into novel a-adjectives because of Categorization. That is, once speakers have access to the mental representation of the a-adjective category, some of its lexical-semantic properties will spill over, or extend, to new members that share similar lexical features. However, it is important to note that in that study, speakers were exposed to prenominal and relative clause examples in the exposure block, and the experimenter guided the production block by providing phonological information about the labels (i.e., saying them out loud in front of the participant). For the purposes of our study, we did not facilitate syntactic cues in the exposure block, and participants had to generate their own phonological cues during the production block (i.e., they were asked to say the labels out loud upon seeing them on the screen). With respect to these self-generated phonological cues, we did have several participants who generated 'incorrect' phonological forms for novel a-adjectives (e.g., for the target *ablim*, some speakers produced it in the form of [əblɪm'], while others did so in the form of [eɪ'blɪm]), which may have weakened the 'attraction' effect of real a-adjectives. This suggests that Categorization may be dependent on the presence of specific linguistic (syntactic and/or phonological) cues[2]. It is possible that by reintroducing the syntactic cue alone we would observe Categorization effects, but that this could also be contingent on whether the person generates the correct phonological form. The idea of generalization abilities in the context of this paradigm has been put into question recently (Yang, in press), suggesting that speakers do not generalize on the basis of Categorization, and instead, simply generalize on the basis of prior morphosyntactic knowledge. However, on the basis of this account, we would predict that our participants should have shown some indication of generalization irrespective of the manipulations we used in the exposure block. Instead, the present findings align best with the idea that speakers generalize on the basis of exposure to input that is linguistically relevant to a given context.

[2] It is also reasonable to assume that, in addition to these linguistic cues, the demands imposed by the other manipulations may have contributed to the lack of generalization. However, we conducted a second experiment with the German-English group that was identical (in the exposure block) to Experiment 2 in Boyd and Goldberg (2011) to see whether we could observe the learning effects of statistical preemption. Our findings mirrored the results found in the original study, suggesting that the generalization results in the present study are driven by the linguistic cues. This second experiment is not discussed in detail here, because we only have data for one group of speakers.

4.2. Are demands imposed by language-specific features when speaking in the L2?

The latency measures collected in the present study also revealed a set of complex and interesting results. On the one hand, Spanish-English bilinguals were found to be maximally slowest in generating sentences across all conditions, relative to monolinguals and German-English bilinguals. On the other hand, German-English bilinguals were overall just as fast as monolinguals in producing these sentences. In principle, the fact that Spanish-English bilinguals experienced a general slowdown supports the notion of a frequency-lag in the L2 because of less experience in each language (Gollan et al., 2008), and that this frequency-lag extends to the realm of grammatical information (Runnqvist et al., 2013). However, because German-English bilinguals and monolinguals were so similar in their latencies, it seems less likely that the underlying mechanism generating these results was a frequency-lag. If anything, the German-English group, being immersed in the native and more dominant L1, should have shown stronger slowdown effects. Because of this, the pattern of results do not align with the notion of a global-inhibitory component either, since it would also predict German-English bilinguals to be maximally slowest as a function of speaking the L2 while being immersed in the L1. It is possible that these effects of global inhibition do not affect highly proficient bilinguals in the same way as L2 learners (e.g., Linck et al., 2009), because they are simply more experienced at negotiating the use of two languages. However, the fact that Spanish-English bilinguals scored higher when performing the Verbal Fluency task in the L2 relative to the L1, and the fact that the German-English bilinguals scored higher in the L1 relative to the L2, suggest that even highly proficient bilinguals are susceptible to global inhibitory effects of immersion.

One alternate explanation for these latency patterns is that both groups of bilinguals simply experienced cross-language activation of syntactic information at the onset of planning, and that the selection process yielded either interference or facilitation. For Spanish-English bilinguals, this meant that they had to constantly suppress a highly salient construction in Spanish (i.e., postnominal) in order to produce correct sentences using prenominal and relative clause constructions. German-English bilinguals, on the other hand, may have experienced less interference because prenominal and relative clause constructions are also prototypical alternatives in German. Although German relative clauses do not share linear word order with English (e.g., *Der Junge, der schlafend ist…* literally could be translated as *The boy, who asleep is…*), what may be more important here is the extent to which different syntactic choices relate to one another because of their function, and not just because of their form. The Construction Grammar framework fits well with this idea because it assumes that constructions (e.g., morphemes, words, phrases, etc.) are created on the basis of form-meaning/function pairings (Goldberg, 2013). This would also help explain, at least partly, why previous studies have found priming of syntactic structures across languages when there is no linear word order overlap (e.g, Fleischer et al., 2012; Shin & Christianson, 2009; Weber & Indefrey, 2009). As plausible as this interpretation may be, future research will need to investigate this issue using online measures that are more temporally sensitive (e.g., eye-tracking and/or event related potentials) in order to identify the different processing components that are involved when planning a sentence.

A second possible explanation for the latency patterns is that the demands imposed by these constructions are different for each bilingual group, because the way in which they use their two languages on a daily basis is different. The Adaptive Control Hypothesis (Green & Abutalebi) proposes that the domain-general control processes regulating the use of two (or more) languages change, or adapt, to particular kinds of bilingual experiences. To illustrate, some bilinguals are able to use one language in certain contexts or with some people, but have to use the other language in other contexts/with other people. In order to navigate through these scenarios effectively, these bilinguals would have to become skilled at suppressing the unintended language quickly. Conversely, some bilinguals are able to switch (i.e., codeswitch) from one language to another with many interlocutors in any given conversation. For these speakers to be able to engage in this kind of behavior fluidly, they would need to become skilled at maintaining the two languages active. According to Green and Abutalebi (2013), each of these situations would yield different cognitive profiles. To investigate this issue preliminarily, we included a series of basic items in the Language History Questionnaire asking whether they code-switch or not. All twelve Spanish-English participants reported codeswitching on a regular basis, but none of the German-English participants reported engaging in such behavior. To the extent that these basic demographics are genuine, we could conclude that Spanish-English bilinguals were slower given that they have more experience keeping the two languages maximally active, which would go against the task demands in our experiment. German-English bilinguals, on the other hand, were faster because they were already more experienced at suppressing the unintended language, which would be an advantageous strategy to possess in this particular case. Similar to our previous discussion, however, we would like to emphasize that the last two accounts presented here are not mutually exclusive by any means.

4.3. Does lexical information influence access to syntactic information in production?

Because the present study used 3rd person singular verbs as part of the filler stimuli, we were able to ask whether the latencies for real a-adjectives were simply driven by the fact that they required a more demanding structure (i.e., relative clause), or whether there was something unique about the processing of a-adjectives, in addition to the demands imposed by relative clauses. The results indeed showed that sentences containing real a-adjectives were slower to generate than sentences containing verbs, even when including both prenominal and relative clause responses for real a-adjectives. This finding is consistent with the bilingual production model proposed by Hartsuiker et al. (2004) in which lexical items (i.e., verbs) activate syntactic information, and is also consistent with recent studies showing that lexical access affects the processing of syntactic information in L2 sentence comprehension (Hopp, 2015). Our findings add several pieces of information that build from these previous studies. First, we find that the lexical-syntactic interaction within and across languages is not restricted to special cases like verbs, which can carry more 'grammatical' information than other lexical items. Instead, this interaction seems to be present across different kinds of lexical items. Second, because all three groups in our study showed this pattern, our findings suggest that this phenomenon may be a consequence of a general principle that operates in the system. In addition to the contributions this study makes with respect to the bilingual literature, the present results also support the statistical preemption hypothesis in that speakers activate competing alternatives. That is, at the onset of planning, speakers would partially activate the 'a-adjective + prenominal' construction, but this would be suppressed by the co-activation of the stronger 'a-adjective + relative clause' construction. Conversely, the verbs used in our study would not engage competing alternatives in the same way with respect to syntactic placement. If there were competing alternatives, these would be maximally suppressed by the task demands (i.e., by not being able to change the form of the target label, speakers would be constrained to answer using relative clauses with verbs).

5. Conclusion

To conclude, the present study yielded three main findings. First, we found that bilingualism modulates the use of subtle grammatical constraints in the L2. In the presence of a sentence production task that imposed cognitive demands, bilinguals showed more steady performance patterns than monolinguals. This was evidenced by the fact that bilinguals as a whole committed less errors, and were more consistent in generating the a-adjective constraint. This set of findings is consistent with the emerging literature showing that bilinguals exploit cognitive control differentially than monolinguals, and that this has consequences for language use. Second, we also found differences between the two bilingual groups when examining their latency abilities (i.e., slower latencies for Spanish-English bilinguals, and faster latencies for German-English bilingual comparable to monolinguals). Spanish-English bilinguals may have experienced more cross-language interference because of the need to suppress irrelevant syntactic information and/or because of their expertise in keeping the two languages available. Conversely, German-English bilinguals may have experienced less cross-language interference because of the languages having more 'grammatical' alignment and/or because of the bilinguals' expertise at suppressing the unintended language. This is a fascinating line of research that needs more attention from the bilingual research community. Future studies should investigate the extent to which we can dissociate the consequences of co-activation vs. the consequences of the interactional context in which a bilingual uses his/her languages. Finally, we found that lexical information can influence the activation of syntactic information in sentence planning. We showed that usage-based linguistic theories can be used as a tool to examine different aspects of language processing in bilinguals. Two general limitations of the present study were the small sample size for each bilingual group, and the small number of trials used across conditions. Future research investigating the questions outlined in this study should also use different paradigms that have more robust designs, and should also look at different bilingual populations.

References

Abutalebi, J., Annoni. J. M., Seghier, M., Zimine, I., Lee-Jahnke, H., Lazeyras, F., Cappa, S. F., & Khateb, A. (2008). Language control and lexical competition in bilinguals: an event-related fMRI study. *Cerebral Cortex*, 18, 1496-505

Abutalebi, J., & Green, D. (2008). Control mechanisms in bilingual language production: Neural evidence from language switching studies. *Language and Cognitive Processes* 23(4), 557 - 582.

Abutalebi, J., & Green D. (2007). Bilingual language production: The neurocognition of language representation and control. *Journal of Neurolinguistics*, 20, 242-275.

Alderson, J. C., Clapham, C., & Steel, D. (1997). Metalinguistic knowledge, language aptitude and language proficiency. *Language Teaching Research*, 1(2), 93-121.

Bates, E., & MacWhinney, B. (1987). Competition, variation, and language learning. In *Mechanisms of language acquisition*, edited by B. MacWhinney. Hillsdale, NJ: Lawrence Erlbaum Associates.

Bialystok, E. (2001). Metalinguistic aspects of bilingual processing. *Annual Review of Applied Linguistics*, 21(1), 169.

Bialystok, E. (1999). Cognitive complexity and attentional control in the bilingual mind. *Child Development*, 70, 636–644.

Bialystok, E. (1990) Communication Strategies: A Psychological Analysis of Second Language Use, Oxford: Basil Blackwell.

Bialystok, E., Craik, F. I. M., Green, D. W., & Gollan, T. H. (2009). Bilingual minds. *Psychological Science in the Public Interest*, 10, 89–129.

Boyd, J. K., & Goldberg, A. E. (2011). Learning what not to say: The role of statistical preemption and categorization in a-adjective production. *Language*, vol. 87, pp. 55-83, 2011.

Brenders, P., Van Hell, J. G., & Dijkstra, T. (2011). Word recognition in child second language learners: Evidence from cognates and false friends. *Journal of Experimental Child Psychology*, *109*, 383-396.

Bybee, J. L. (2013). Usage-based theory and exemplar representation. In Thomas Hoffman and Graeme Trousdale (eds.) *The Oxford Handbook of Construction Grammar*, pp. 49-69. Oxford University Press.

Bybee, J. L. (2010). Language, usage and cognition. Cambridge: Cambridge University Press.

Bybee, J. L. (2008). Usage-based grammar and second language acquisition. In *Handbook of cognitive linguistics and second language acquisition*, edited by P. Robinson and N. C.Ellis. London: Routledge.

Bruck M., & Genesse, F. (1995). Phonological awareness in young second language learners. *Journal of Child Language, 22,* 307-324.

Carlson, S. M, & Meltzoff, A. N. (2008). Bilingual experience and executive functioning in young children. *Developmental Science* 11 (2): 282–298.

Chambers, C. G., & Cooke, H. (2009). Lexical competition during second-language listening: sentence context, but not proficiency, constrains interference from the native lexicon. *Journal of Experimental Psychology, Learning, Memory, and Cognition*, 35, 1029–1040.

Christoffels, I. K., & De Groot, A. M. B. (2005). Simultaneous interpreting: A cognitive perspective. In J. F. Kroll, and A. M. B. de Groot, (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 454–479). New York: Oxford University Press.

Costa, A. (2005). Lexical access in bilingual production. In J. F. Kroll, and A. M. B. De Groot, (Eds.), *Handbook of bilingualism: Psycholinguistic approaches*, pp. 308–325. New York: Oxford University Press.

Costa, A., Caramazza, A., & Sebastián-Gallés, N. (2000). The cognate facilitation effect: Implications for models of lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26, 1283–1296.

Costa, A., Hernández, M., & Sebastián-Gallés, N. (2008). Bilingualism aids conflict resolution: Evidence from the ANT task. *Cognition*, 106, 59–86.

Costa, A., Miozzo, M., & Caramazza, A. (1999). Lexical selection in bilinguals: Do words in the bilingual's two lexicons compete for selection? *Journal of Memory and Language*, 41, 365–397.

De Groot, A. M. B., Delmaar, P., & Lupker, S. J. (2000). The processing of interlexical homographs in translation recognition and lexical decision: Support for non-selective access to bilingual memory. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 53A*, 397–428.

Dijkstra, T. (2005). Bilingual visual word recognition and lexical access. In J.F. Kroll & A.M.B. De Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 179–201). New York: Oxford University Press.

Dijkstra, A., Van Jaarsveld, H., & Ten Brinke, S. (1998). Interlingual homograph recognition: Effects of task demands and language intermixing. *Bilingualism: Language and Cognition*, 1, 51–66.

Dussias, P.E., & Sagarra, N. (2007). The effect of exposure on syntactic parsing in Spanish-English bilinguals. *Bilingualism: Language and Cognition*, 10, 101–116.

Duyck, W., Assche, E., Drieghe,D., & Hartsuiker, R. J. (2007). Recognition by bilinguals in a sentence context: Evidence for nonselective lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33, 663–679.

Elston-Güttler, K. E., Gunter, T. C., & Kotz, S. A. (2005). Zooming into L2: Global language context and adjustment affect processing of interlingual homographs in sentences. *Cognitive Brain Research*, *25*(1), 57-70.

Emmorey, K., Borinstein, H. B., Thompson, R. L., & Gollan, T. H. (2008a). Bimodal bilingualism. *Bilingualism: Language and Cognition*, 11, 43–61.

Galambos, S.J., & Goldin-Meadow, S. (1990). The effects of learning two languages on levels of metalinguistic awareness. *Cognition*, 34: 1-56.

Gennari, S. P., & MacDonald, M. C. (2009). Linking production and comprehension processes: The case of relative clauses. *Cognition*, 111, 1–23.

Goldberg, A. E. (2013). Constructionist Approaches to Language. In Thomas Hoffmann and Graeme Trousdale (eds.) *Handbook of Construction Grammar*. Oxford University Press.

Goldberg, A.E. (2011). Corpus evidence of the viability of statistical preemption. Cognitive Linguistics 22 1: 131-154.

Goldberg, A.E. (2011). Are a-adjectives like afraid prepositional phrases underlying and does it matter from a learnability perspective? Princeton University.

Goldberg, A. E. (2006). Constructions at Work: the nature of generalization in Language. Oxford: Oxford University Press.

Goldberg, A. E. (1995). Constructions: A construction grammar approach to argument structure. Chicago: The University of Chicago Press.

Gollan, T. H., Montoya, R. I., Cera, C., & Sandoval, T.C. (2008). More use almost always means a smaller frequency effect: Aging, bilingualism, and the weaker links hypothesis. *Journal of Memory and Language*, 58, 787–814.

Gollan T.H., & Silverberg N.B. (2001). Tip-of-the-tongue states in Hebrew-English bilinguals. *Bilingualism: Language and Cognition*. 2001; 4:63–84.

Green, D. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition*, 1, 67–81.

Grosjean, F. (1989). Neurolinguists, beware! The bilingual is not two monolinguals in one person. *Brain and Language*, 36, 3–15.

Gullifer, J. W. (2015). *Using syntactic priming to identify cross-language constraints in bilingual language processing*. Unpublished doctoral dissertation, Pennsylvania State University, University Park, PA.

Hartsuiker, R. J., Pickering, M. J., & Veltkamp, E. (2004). Is Syntax Separate or Shared Between Languages? *Psychological Science (Wiley-Blackwell)*, *15*(6), 409-414.

Hermans, D., Bongaerts, T., De Bot, K., & Schreuder, R. (1998). Producing words in a foreign language: Can speakers prevent interference from their first language*? Bilingualism: Language and Cognition*, 1, 213–229.

Hopp, H. (2015). The timing of lexical and syntactic processes in second language sentence comprehension. *Applied Psycholinguistics*, 1-28.

Hoshino, N., & Kroll, J. F. (2008). Cognate effects in picture naming: Does cross-language activation survive a change of script? *Cognition*, 106, 501–511.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language, 59*(4), 434-446.

Kroll, J.F., Bobb, S.C., Misra, M., & Guo, T. (2008). Language selection in bilingual speech: Evidence for inhibitory processes. *Acta Psychologica*, 128, 416–430.

Kroll, J.F., Bobb, S.C., &Wodniecka, Z. (2006). Language selectivity is the exception, not the rule: Arguments against a fixed locus of language selection in bilingual speech. Bilingualism: *Language and Cognition*, 9, 119–135.

Kroll, J. F., Dussias, P. E., Bice, K., & Perrotti, L. (2015). Bilingualism, mind, and brain. In M. Liberman & B. H. Partee (Eds.), *Annual Review of Linguistics*.

Kroll, J. F., & Gollan T. H. (2014). Speech planning in two languages: what bilinguals tell us about language production. In *The Oxford Handbook of Language Production*, ed. V Ferreira, M Goldrick, M Miozzo, pp. 165–81. Oxford, UK: Oxford Univ. Press

La Heij, W., Bruyn, E., De Elens, E., Hartsuiker, R., & Helaha, D. (1990). Orthographic facilitation and categorical interference in a word-translation variant of the stroop task. *Canadian Journal of Psychology*, 44, 76–83.

Lakoff, G. (1987). Women, fire, and dangerous things: What categories reveal about the mind. Chicago: University of Chicago Press.

Langacker, Ronald W. (1987). Foundations of cognitive grammar: Vol. 1. Theoretical prerequisites. Stanford, CA: Stanford University Press.

Libben, M., & Titone, D. (2009). Bilingual lexical access in context: Evidence from eye movements during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(2), 381-390.

Linck, J. A., Kroll, J. F., & Sunderman, G. (2009). Losing access to the native language while immersed in a second language: Evidence for the role of inhibition in second language learning. *Psychological Science, 20,* 1507-1515.

MacDonald, M.C., & Seidenberg, M. S. (2006). Constraint satisfaction accounts of lexical and sentence comprehension. In M. J. Traxler, and M. A. Gernsbacher, (Eds.), *Handbook of psycholinguistics, 2nd ed.* (pp.581-611). Amsterdam: Elsevier.

Midgley, K. J., Holcomb, P. J., & Grainger, J. (2009). Masked repetition and translation priming in second language learners: A window on the time-course of form and meaning activation using ERPs. *Psychophysiology, 46*, 551–565.

Miller, N. A., & Kroll, J. F. (2002). Stroop effects in bilingual translation. *Memory and Cognition*, 30, 614–628.

Morales, J., Gómez-Ariza, C.J., & Bajo, M.T. (2013). Dual mechanisms of cognitive control in bilinguals and monolinguals. *Journal of Cognitive Psychology, 25*, 531-546.

Morford JP, Wilkinson E, Villwock A, Piñar P, Kroll JF. 2011. When deaf signers read English: Do written words activate their sign translations? *Cognition* 118:286–92

Pickering, M. J., & Ferreira, V. S. (2008). Structural Priming: A Critical Review. *Psychological Bulletin*, *134* (3), 427–459. http://doi.org/10.1037/0033-2909.134.3.427

R Development Core Team (2010). R: A Language and Environment for Statistical Computing Vienna: R Foundation for Statistical Computing.

Rosvold, H. E., Mirsky, A. F., Sarason, I., Bransome, J., Edwin, D., & Beck, L. H. (1956). A continuous performance test of brain damage. *Joumal of Consulting Psychology*, 20, 343-350.

Runnqvist, E., Gollan, T. H., Costa, A., & Ferreira, V. S. (2013). A disadvantage in bilingual sentence production modulated by syntactic frequency and similarity across languages. *Cognition*, *129* (2), 256-263.

Schoonbaert, S., Hartsuiker, R., & Pickering, M. (2007). The representation of lexical and syntactic information in bilinguals: Evidence from syntactic priming. *Journal of Memory and Language*, 56(2), 153-171.

Schwartz, A., & Kroll, J. (2006). Bilingual lexical activation in sentence context. *Journal of Memory and Language*, *55*(2), 197-212.

Schwartz, A. I. & Van Hell, J. G. (2012). Bilingual visual word recognition in sentence context. In J. Adelman (Ed.), *Visual Word Recognition* (pp. 131-150). Psychology Press.

Thierry, G., & Wu, Y. J. (2007). Brain potentials reveal unconscious translation during foreign language comprehension. *Proceeding of National Academy of Sciences*, 104, 12530–12535.

Van Assche, E., Duyck, W., Hartsuiker, R. J., & Diependaele, K. (2009). Does bilingualism change native language reading? Cognate effects in a sentence context. *Psychological Science, 20*, 923-927.

Van Hell, J. G., & Dijkstra, T. (2002). Foreign language knowledge can influence native language performance in exclusively native contexts. *Psychonomic Bulletin & Review*, 9, 780–789.

Von Studnitz, R., & Green, D. W. (2002). Interlingual homograph interference in German–English bilinguals: Its modulation and locus of control. Bilingualism: *Language and Cognition*, 5, 1–23.