Psycholinguistic Approaches to Language Processing in Heritage Speakers

Patrick A. Bolger and Gabriela C. Zapata
University of Alberta

Abstract
This paper focuses on the dearth of language-processing research addressing Spanish heritage speakers in assimilationist communities. First, we review key offline work on this population, and we then summarize the few psycholinguistic (online) studies that have already been carried out. In an attempt to encourage more such research, in the next section, we review the various techniques that psycholinguists commonly use in studies of language processing. We restrict our review to approaches that should be within the reach of most researchers. In our final section, we provide points of departure for further psycholinguistic investigations into this population. We do this in the hope that, one day, we may be able to determine with some degree of certainty the challenges that this population of Spanish speakers faces when processing their first language online, in real time.

Introduction

Since the end of the 19th century, psycholinguistics has contributed to our understanding of first language (L1) storage and processing, and, in the last thirty years, it has illuminated the study of bilingualism (see Garnham, Garrod, & Sanford, 2006 for a review of the history of psycholinguistics). The psycholinguistic literature on bilingualism is also quite comprehensive at this point, and it has greatly contributed to our knowledge on the matter (see Field, 2008; Kroll & Sunderman, 2003; and Schwartz & Kroll, 2006 for reviews of the existing literature).

However, there is one population of bilinguals that very few researchers in psycholinguistics have addressed, namely, Spanish heritage speakers in assimilationist English communities.[1] This population essentially encompasses most Spanish-heritage speakers in the United States.[2] Unlike other early bilingual populations such as Catalan-Spanish speakers, for instance, these heritage speakers often do not reach the same level of linguistic ability in their two languages.[3] As a group they may also exhibit varying levels of proficiency in all or several skill areas (i.e., speaking, reading, writing, and listening). Usually, such heritage speakers are second- or third-generation immigrants who grow up in monolingual Spanish or bilingual Spanish-English households (in which one or both parents speak Spanish). In these situations, they are exposed to Spanish or both English and Spanish since birth or before the age of 5, but receive most of or all their schooling in English. That is, the experiences of these speakers are different from those of their monolingual, late-, and balanced-bilingual peers on various levels, including sociolinguistic variables such as L1/minority status, the use of Spanish in the family, the age onset of L2 acquisition (or onset of bilingualism), L1 input availability, and L1 literacy. This unique situation often impedes them from attaining native-level abilities in their L1. As stated by Montrul in much of her work on heritage speakers (e.g., Montrul, 2002, 2004b, 2006a, 2006b, 2008; Montrul & Bowles, 2009; Montrul, Foote, & Perpiñán, 2008), they share linguistic

characteristics with not only their L1 monolingual counterparts because of their early onset of L1, but also with L2 learners because of their varying degrees of proficiency in Spanish acquisition.

Even though in the last decade a relatively significant number of studies have investigated heritage speakers' grammars, most of them have relied on offline techniques such as grammaticality judgments (GJ). The advantages of this technique have been clear for some time. One advantage is the precision with which one can target not only the specific constructions one has hypothesized, but also specific violations of those constructions. Another advantage is that they probe the speaker's knowledge (competence) in the language, relatively unfettered by limitations in the speaker's productive or receptive abilities. However, the traditional offline GJ task does not offer much information about language processing in real time (online processing), which is how language functions must normally be carried out.[4] Juffs (2001) has also noted that this task may not tell how a participant arrived at a particular judgment, and may allow the participant to incorporate metalinguistic knowledge into their judgments (p. 208). Naturally, one way to address these deficiencies is to use more online tasks, such as those used in psycholinguistic research, to test grammatical knowledge. Indeed, we see the two methods as complementary.

Accordingly, the purpose of this article is to explore how the field of psycholinguistics has contributed, and what it might further offer, to our understanding of heritage speakers' language abilities. The first section of the paper is devoted to the existing literature on this group's grammatical representations, an area that shows great promise for future psycholinguistic studies, but has so far been analyzed almost entirely with offline studies. Next, we focus on the few psycholinguistic studies that have been carried out with Spanish heritage speakers in assimilationist communities. The third section provides information on the online methods used by psycholinguists for data collection, focusing on methods that are within the potential reach of most academic researchers. In the last section of the paper, we provide suggestions for further psycholinguistic investigations on this topic.

Selected Studies on Heritage Speakers' Grammars

Silva-Corvalán's (1991, 1994) and Zentella's (1997) seminal work with heritage speakers in Los Angeles and New York in the 1990s drew attention to Spanish-English bilingual populations that had previously been mostly unexplored, and they inspired a large number of studies with heritage speakers in the U.S. These first studies were very important because they helped us define who heritage speakers were, and they pointed to apparent differences between the Spanish grammatical knowledge of these speakers and that of their monolingual counterparts. In addition, the data from these studies seemed to suggest that the grammars of heritage speakers were characterized on the one hand by convergence, defined by Bullock and Toribio (2004) as "the enhancement of inherent structural similarities found between two linguistic systems [in this case Spanish and English]" (p. 91). On the other hand, they were also characterized by divergence from monolingual norms. It was evident at the time that heritage speakers were different, but not *how* they were different. Studies in subsequent years have sought to answer this question by investigating a variety of structures.

Perhaps the most prolific research program with heritage speakers of Spanish in the U.S. is that of Silvina Montrul and her colleagues (e.g., 2002; 2004b; Montrul & Bowles, 2009; Montrul, Foote, & Perpiñán, 2008). Part of this body of research has pioneered the examination of early grammars with online techniques (Montrul, 2006b, reviewed in the next section). The rest of the studies are based on a variety of offline methods. They have focused on linguistic structures that, due to their nature, are governed by properties of the core syntax and those at the interface between syntax and discourse/pragmatics. This focus is important because Silva-Corvalán's (1991, 1994) work and the work of researchers with heritage speakers in other languages (e.g., Sorace's work) have suggested that although the core syntax of heritage speakers may exhibit the same properties as that of monolinguals, the semantic and discourse-pragmatic knowledge that govern particular grammatical structures may differ for heritage speakers. As we will argue in the final part of this paper, it is the processing of these grammatical structures that seems particularly worthwhile to investigate in this population of bilinguals, since such structures seem to distinguish them from monolingual and other bilingual populations (e.g., late bilinguals, bilinguals of other languages).

In one of her first studies on heritage speakers, Montrul (2002) focused on the distinctions in aspect between Spanish Preterite and Imperfect tenses. There were four groups of participants: adult native monolinguals; early bilinguals who had acquired both English and Spanish before the age of 3 (simultaneous bilinguals); early bilinguals who had learned English between ages 4 and 7 (early child L2 learners); and 8 late child L2 learners. The second and third groups are classic heritage speakers. The participants completed four offline production and comprehension tasks: a written morphology recognition task, an oral production task, a sentence conjunction task, and a truth value judgment task.

Montrul (2002) found that the later one's exposure to English, the less likely one is to have difficulty with the Preterite/Imperfect distinction. Specifically, the simultaneous bilinguals were most affected, followed by early child L2 learners, and then late child L2 learners. In other words, heritage speakers were heavily influenced. This confirms results from previous studies, suggesting that even if early onset of acquisition is important, it is not enough for full attainment, especially if the L1 is in a minority situation, and L1 input is reduced.

More specifically, the second through fourth tasks suggested overall that the areas of most difficulty were stative verbs in the Preterite, followed by achievement verbs in the Imperfect, both of which are marked forms realized under only certain conditions. Montrul (2002) claims that these aspects of grammatical knowledge are more prone to erosion and incomplete acquisition among heritage speakers, since acquisition of them is more dependent on frequency of input.

The issues raised in this work were further investigated by Montrul (2004b) in her study on heritage speakers' production of Spanish subjects and objects, which are both pragmatically and semantically determined. Specifically, in Spanish, null subjects are the norm, and while the overt expression of subjects is allowed, the appropriateness of their presence depends on discourse context and pragmatics factors (e.g., topic, change of referent, etc.). Similarly, direct and indirect objects (and their associated clitics) are also governed by syntax and semantics (e.g., the

distribution of the preposition *a* with animate versus inanimate objects). The results from transcribed picture narratives showed that both monolinguals and heritage speakers used these structures in their retellings, but the heritage speakers, especially those in the intermediate group, were less accurate in general than the monolinguals, despite being aware of the grammatical structures, suggesting that they had not completely mastered their pragmatic determinants (Montrul, 2004b). The results of this study mirrored those reported by Silva-Corvalán (1994) and Zapata, Sánchez, and Toribio (2005) in their studies on overt/null subjects and object expressions.

Montrul further examined heritage speakers' knowledge of object expression in Spanish in a more recent study (Montrul & Bowles, 2009). The focus of this study was Differential Object Marking (DOM), which refers to the obligatory use in Spanish of the preposition *a* before certain direct objects and in related structures. The results of two acceptability judgment tasks and an oral narrative showed significant differences between the heritage-speaker and monolingual participants in their knowledge and production of DOM, with heritage speakers ranking grammaticality significantly differently from monolinguals, and omitting obligatory DOM structures more often. However, it is important to mention that, in spite of these problems, heritage speakers did seem to have robust knowledge of indirect objects.

Overall, the results seem to confirm those from previous studies. First, grammatical structures that involve semantics are unstable in heritage speakers.[5] Montrul and Bowles (2009) believe that these results might be evidence of the important role of L1 input in early bilingual acquisition. A dearth of such input, though perhaps sufficient for early L1 learners to acquire core properties of syntax, is insufficient for them to acquire the aspects of syntax that lie at the nexus of syntax and other domains (e.g., semantics, pragmatics). In such cases, L1 learners of Spanish might incorporate more L2-like (in this case, English-like) constructions into Spanish. This is convergence. Indeed, for Silva-Corvalán (1991, 1994), convergence seemed to also be a key feature of these heritage speakers' grammars.

A final, related study by Montrul, Foote, & Perpiñán (2008) compared Spanish monolinguals, heritage speakers, and late L2 learners on gender agreement. The participants all completed two written tasks and one oral task. The first task consisted of sentences that were presented visually (on a computer screen) in a situational context, accompanied by three different pictures. The subjects had to read the sentence, then choose the picture to which the implicit noun (represented by a blank space) referred, taking into account not only the given context, but also the gender and number of the referent in the sentence. The second written task required participants to work with a passage with 40 blanks, each with two options, and to choose the correct option according to context. The participants' oral production was tested with an oral picture-description task in which they were asked to describe images in photographs.

The results suggested that L2 learners were more accurate than heritage speakers in the written tasks, but not in the production task, where the heritage speakers were more accurate than the L2 learners. But the three tasks together showed that the two groups made similar errors, not only with feminine forms (which were often interpreted as masculine), but also with nouns with exceptional or less canonical endings. This suggests that the two groups did not exhibit major

differences in their knowledge, a result that was also found by Bruhn de Garavito (2002) in her study of Canadian early and late bilinguals' knowledge of Spanish word order.

Montrul, Foote, and Perpiñán (2008) hypothesized that the tasks may have highlighted the differing cognitive strengths of the two groups with respect to language mode. Specifically, the written tasks may have tapped into the L2 participants' explicit (metalinguistic) knowledge, which is typical of instructed, L2 acquisition. In contrast, the oral task involved more automatic processing and the kind of implicit knowledge that is the result of naturalistic language acquisition, which would be the norm for heritage speakers. Based on their results, Montrul, Foote, and Perpiñán concluded that although both groups have knowledge of grammatical gender in Spanish, the two might differ in terms of its storage and processing.

The studies presented in this section suggest that, although heritage speakers of Spanish seem to be aware that the usage of certain constructions in Spanish follows particular rules, they are often uncertain with respect to exactly *when* those rules should apply. This uncertainty translates not only into performing less accurately on those structures compared to native speakers, but also sometimes compared to late learners of Spanish, depending on language mode. The studies also constitute excellent examples of the most important work that has been done with Spanish heritage speakers' grammars. Not only do these studies exemplify good research practices (e.g., they have all taken into account individual differences such as age of onset of bilingual acquisition and proficiency levels), but they have also effectively linked experimental work with linguistic theory. In addition, they have provided us with much information about how the grammar of heritage speakers differs or resembles that of monolinguals on the one hand, and that of late bilinguals on the other. However, there is still much to be done. In particular, the offline techniques on which these studies are based limit what we can infer. In order to widen the scope of interpretation, it is important to test some of this work with online methods.

In the next section we discuss how psycholinguistics has contributed to our understanding of language processing among heritage speakers. To keep the scope of this article reasonable, we will focus on studies that are most relevant to Spanish-English heritage speakers in the U.S., an assimilationist culture.

Selected Psycholinguistic Studies on Spanish-English Bilinguals

Before embarking on a review of psycholinguistic studies done on Spanish-English bilinguals, we think it will be useful to frame the main characteristics of this population within an appropriate theoretical model. Ullman's (2001) Declarative/Procedural Model has implications for L1 attrition, and indirectly, incomplete acquisition, something that we suspect uniquely characterizes the heritage speakers we are highlighting. Generally speaking, this model allocates lexical knowledge to declarative memory (located in the temporal lobe of the brain) and grammatical knowledge to procedural memory (located more frontally). Since it resides in declarative memory, lexical knowledge is more malleable in adulthood and, therefore, more susceptibility to the relative frequencies of its overall membership.

Importantly for bilingualism, this susceptibility might include L1 susceptibility to the relative frequencies of L2 lexical items. In other words, the L1 lexicon may be fairly susceptible to being

suppressed or "eroded" through relatively intensive use of L2 lexical items, something very likely to occur when the L2 is the dominant language in the community. In contrast, the L1 grammar may be spared for the most part because it resides in procedural memory, which is automatized early on, and therefore less plastic and less susceptible to changes in the external linguistic environment of the bilingual, like emigration to a more L2-dominant culture, or immersion in an L2-dominant schools system (see also Hernandez & Li (2007) for a review of the declarative/procedural split and how it accounts for lexical- versus syntactic-processing differences in bilingualism, depending on age of acquisition). This theory of how memory storage across linguistic domains differs could easily be extended to account for the finding that among L1 attriters with a strong L1 background, the L1 lexicon decays faster than the L1 grammar (Montrul, 2008; Köpke, 2002).

Ullman's (2001) model could also be extended to predict that the L1 attrition in young bilinguals that might accompany such changes in the external linguistic environment would be more severe and more comprehensive across various linguistic domains. Indeed, this seems to be the case (Köpke, 2004; Montrul, 2008; Pallier et al., 2003). In other words, attrition is more severe among the younger learners because not only lexical (declarative) but also grammatical (procedural) learning are still taking place when L1 attrition starts. And in extreme cases (e.g., very young children who lose all contact with their L1, say, through adoption), the L2 grammar might take over in procedural memory altogether, essentially reversing (psychologically) the L1 and L2. In less extreme cases, the model would not be inconsistent with the possibility that core aspects of L1 grammatical processing would remain intact while more peripheral aspects (ones learned later) would be taken over by the L2. Montrul (2008) estimates that the age at which the potential for near-native second language acquisition maximizes and language loss minimizes is around age 9, which she also believes (pp. 267-268) is consistent with both Ullman's model and her own claims. Indeed, Montrul (2008) makes the case that bilingual children are most vulnerable to L1 attrition before the age of 9, whereas Ullman's (2001) Declarative/Procedural model provides a viable mechanism through which this attrition could occur.

Since most Spanish-English heritage speakers in the U.S. will eventually integrate into a highly English-dominant society, Ullman's model could be extended to predict that those with an early L2-English onset will exhibit peripheral grammatical-processing difficulties in Spanish as adults. This prediction seems to align with the facts (Montrul, 2008). For now, suffice it to say that what distinguishes the heritage speakers of interest in this paper from those with stronger L1 backgrounds is not so much lexical processing as it is grammatical processing.

As a result, we will refer to lexical processing in this group in just two cases. They are, it turns out, two examples of only a handful of studies that have used psycholinguistic techniques, specifically focusing on adult, Spanish-English heritage speakers in the U.S. (other examples include Kohnert, Hernandez, & Bates, 1998; Nicol, Teller, & Greth, 2001; and Silverberg & Samuel, 2004). After discussing the two selected studies, we will discuss two further psycholinguistic studies that have looked at grammatical processing in this population.[6] Indeed, grammatical processing seems to make this group's challenges unique and worthy enough for specialized discussion. We discuss these studies in the final section of this paper with suggestions for further research.

The first of the four studies looking at Spanish heritage speakers in the U.S. is Schwieter (2008), who investigated how bilinguals control which language is being produced at any one time. This doctoral dissertation involved the participation of heritage speakers and English-dominant late bilinguals (divided into low- and high-proficiency L2 learners based on classroom and university designations). The participants completed two online behavioral tasks: 1) word translation with distracters that were either semantically related or unrelated words or pictures, and 2) picture naming with semantically related versus unrelated distracters in the same modality (words versus pictures).

The translation results showed that the three participant groups performed in similar ways, with related pictures facilitating response latencies in both tasks, and related words inhibiting them in both tasks. These results were evidence in favor of La Heij's (2005) Concept Selection Hypothesis. However, when all groups were collapsed into high- and low-proficiency groups according to a verbal fluency measure (not classroom/university status), Schwieter (2008) found that, in contrast to the high-proficiency group, the low-proficiency group did not get facilitation from related pictures. Similarly, he also found that switching costs in the picture-naming task were greater back into L1 for the lower-proficiency group than for the higher proficiency group. The overall results suggested that proficiency, not bilingual type, was the key factor in determining how bilinguals access concepts versus words during processing, with higher proficiency speakers mediating language selection directly at the conceptual level (as in La Heij's (2005) Concept Selection Hypothesis), and lower proficiency speakers mediating language selection at the lexical level, as would be predicted in Green's (1998) Inhibitory-Control model.

The other psycholinguistic study that touches upon lexical representations in Spanish-English heritage speakers in the U.S. is a study by Moreno and Kutas (2005) that used event-related potentials (ERPs), or "brainwaves"[7] that are time-locked to stimuli and are averaged over many trials and participants (see Online Methods below for a more complete description of the methodology). Based on the finding by Weber-Fox and Neville (1996) that L2 learners who began to learn their L2 relatively late in life (i.e., after the age of 16) showed a delayed semantics-sensitive "brainwave" (an N400) in that language, Moreno and Kutas looked at semantic processing in Spanish and English among Spanish-dominant heritage speakers (i.e., those who started school without knowing any English), and among English-dominant heritage speakers (who had been exposed to some English before school). The researchers presented both groups of bilinguals with English and Spanish sentences containing semantic anomalies. The sentences were presented one word at a time on a screen (*rapid serial visual presentation*, RSVP). After reading the entire sentence, the participants had to indicate whether it was a proper sentence in that language. They found that the onset of the semantics-sensitive N400 "brainwave" was delayed for each group in their non-dominant language relative to their dominant language. This delay could be explained by the relatively late exposure to English among the Spanish-dominant participants, but not among the English-dominant participants, who were also characterized by early exposure to Spanish. The only difference was that this latter group had more English influence early on, ultimately leading to English dominance because of its dominance in the culture as well. Thus, with respect to semantics, it appears that Spanish

heritage speakers in the U.S. can end up processing their L1 in a way similar to how late bilinguals process their L2 if their exposure to English comes relatively early.

In terms of our knowledge of heritage speakers, the studies by Schwieter (2008) and Moreno and Kutas (2005) are important because they are first steps in a promising line of research, suggesting that age of onset of bilingualism or type of acquisition does not necessarily provide early bilinguals with a persistent advantage in lexical processing and production in their L1. They may, in fact, have certain characteristics in common with those of late L2 learners. This line of research is consistent with more traditional linguistic research suggesting that this population has undergone incomplete acquisition. It is also consistent with the predictions of Ullman's (2001) model of bilingual development and its implications for early L1 attrition. Namely, the lexicon will surely suffer in cases of incomplete acquisition.

But naturally, psycholinguistics also addresses areas other than the lexicon, grammatical processing in particular (for reviews, see Juffs, 2001; Kormos, 2006; & Schwartz & Kroll, 2006). Grammatical processing among Spanish heritage speakers in the U.S. may be particularly vulnerable to age-of-acquisition effects. Studies looking at age of acquisition have mostly focused on such effects among L2 learners and other late bilinguals, but to our knowledge, only a very small number have looked at the heritage speakers we are concerned with in this paper.

Foote (2007; 2010) used an online sentence-fragment completion task to investigate possible differences in the production of Spanish and English subject-verb agreement by not only Spanish-English bilinguals who learned Spanish either early (heritage speakers) or late (late learners), but also Spanish and English monolinguals. Foote also examined the effects of age of English acquisition and proficiency on both bilingual groups' agreement production. The material in Foote's study included sentence fragments in English and Spanish that consisted of noun phrases with a singular head noun followed by a prepositional phrase. In half of the fragments the head nouns consisted of conceptually plural, but grammatically singular referents, such as *the label on the bottles* (where multiple labels are entailed from the multiple bottles they are on), whereas in the other half, the head nouns had referents that were both conceptually and grammatically singular (e.g., *the road to the mountains*). All the fragments were accompanied by pictures to trigger readings that were either conceptually plural or singular. The fragments were presented aurally, accompanied by the corresponding picture on a computer screen. The participants were asked to repeat the fragments and to transform them into a sentence as quickly as they could. The recorded responses were later analyzed for accuracy, but no reaction times were taken into account.[8] The overall results of the three experiments in the study showed no differences in the way either group of bilinguals versus monolinguals produce subject-verb agreement, nor in the way in which bilinguals produced agreement in either of their two languages. However, when Foote considered the relationship between age of acquisition and proficiency, she discovered that these two factors did affect production, particularly in the bilinguals' degree of sensitivity to conceptual effects. Specifically, heritage speakers and intermediate learners seemed to be more affected by the task manipulations than late learners and advanced learners, respectively. Also of interest is that heritage speakers seemed to make more errors than late learners with the same proficiency level, which Foote speculates was probably a result of the way these two groups had learned Spanish (naturalistic versus instructed). This

finding fits well with the views of Montrul (2008), who claims that heritages speakers of Spanish who learn English very in early childhood are more vulnerable to incomplete acquisition of grammatical rules.

Finally, Montrul (2006b) investigated how heritage speakers process English and Spanish unaccusative and unergative verbs, comparing their performance to that of monolingual native speakers of both languages. She focused on verbs that reflected Sorace's (2000) semantic unaccusative hierarchy. According to Sorace's hierarchy, core unaccusative verbs such as *caer* [to fall] are easier to process than more peripheral verbs, such as *faltar* [to lack] because they are less susceptible to grammatical variability. Such differences in ease of processing, Montrul hypothesized, should be reflected in the participants' response latencies in an online task.

The materials in Montrul's study were in both English and Spanish, with the English as translations of the Spanish. The materials included sentences with verbs in each category of Sorace's (2000) hierarchy (e.g., unaccusative core, less core, periphery, etc.) in four different conditions based on the language-specific properties of these verbs in Spanish (e.g., preverbal or postverbal subject). Participants engaged in a self-paced reading task, where they read a sentence presented on a computer screen in four unpunctuated fragments, one fragment at a time, and they indicated with a button press when they had finished reading the fragment (see a more detailed description of this task in the next section). Since Spanish was the weaker language for the heritage speakers, Montrul hypothesized that they would take relatively more time processing Spanish than monolingual Spanish speakers, and that they would be faster in English than in Spanish. Naturally, she also wanted to investigate whether the heritage speakers' processing of English sentences would be different from that of monolingual speakers.

There were similarities and differences in the results between the heritage speakers on the one hand, and the two monolingual groups on the other. For example, the heritage speakers took more time to respond to the stimuli in both languages, so they were slower than the monolinguals overall. However, the speed with which heritage speakers processed Spanish unaccusatives and unergatives was similar to that of monolingual speakers. In contrast, some differences were evident in the English task, as the heritage speakers displayed faster latencies to "less core" verbs (in Sorace's (2000) hierarchy) than English monolinguals, who performed in the same way for all of Sorace's semantic subclasses. The pattern of performance by the heritage speakers in English was similar to the pattern in Spanish, and, hence, that of monolingual Spanish speakers.

The results of this study (including the offline tasks not reported here) showed that heritage speakers' global knowledge of unaccusativity in Spanish and English was comparable to that of monolingual native speakers in both languages. These results are evidence that heritage speakers had acquired the "core" syntax, shared across languages. However, when considering the data from both the on- and offline tasks, the heritage speakers seem to exhibit differences in those aspects of grammar that are more language specific, and, therefore, less core and more peripheral. In this particular case, the heritage speakers may differ with respect to the lexico-semantic representations of some verbs in the unaccusativity hierarchy (Montrul, 2006b).

This study and the one reviewed before it (Foote, 2007), though unrelated with respect to the grammatical feature at hand, are unique in their attempt to improve our understanding of the grammatical processing of Spanish-English heritage speakers in the U.S. The two reviewed before that (Schwieter, 2008; Moreno & Kutas, 2005) are also unique in their attempt to improve our knowledge of other aspects of language processing in this population. Interestingly, the findings from each of the studies are consistent with both Montrul (2008) directly, and Ullman (2001) indirectly (i.e., as extended to the heritage speaker). Namely, both perspectives predict that being exposed significantly to the L2 in an L2-dominant culture before the acquisition of the L1 has stabilized somewhat can have particularly negative consequences for ultimate acquisition of the L1.

Again, however, we argue that it is important to carry out more research using online techniques in order to complement offline techniques. A total of four studies of online processing is woefully insufficient. With this in mind, we begin the second half of this paper by covering some of the online techniques that can be feasibly used by many researchers. In turn, this review provides points of departure for future online research into the grammatical processing of heritage speakers, a topic reserved for the end of this paper.

Online Methods

The most important contribution of the field of psycholinguistics is its ability and potential to tap into the cognitive mechanisms that are involved in real-time language processing (Carreiras & Clifton, 2004). In order to investigate processing, psycholinguists tend to rely much less on offline (relatively time-insensitive) techniques such as grammaticality judgments, and instead focus on online (relatively time-sensitive) experimental methods and tools that can account for the temporal and, increasingly, spatial characteristics of the mind/brain as it comprehends or produces language. The following sections will focus on online methods. We will discuss approaches that are not only currently productive in the literature, but also appropriate currently for heritage speakers and within the reach, potentially, of most researchers interested in the topic. Namely, these are behavioral methods, eyetracking, and event-related potentials.[9] As we describe these approaches, we will also highlight their relative advantages and disadvantages with respect to what they measure.

Behavioral Methods. Over the years, researchers in the field of psycholinguistics have created a great variety of behavioral techniques to measure language processing. In fact, it is impossible to cover them all here, nor is it possible to predict new variations in the future. Instead, we will attempt to break them down into their component parts, the idea being that the reader will gain insight into what these techniques measure and how they might contribute to the study of heritage speakers. To conserve space, the behavioral methods discussed below include only those that are used primarily with adults or older children.[10]

Next to naturalistic data collection, (linguistic) behavioral methods are the oldest (dating back to the late 19th century), least expensive, and most well documented of all psycholinguistic techniques. It is not easy to formulate a definition that covers all such methods. However, a reasonable one would be tasks that elicit conscious, overt, physical responses that are not only external and easily observed, but also constrained by time and prone to response error.

Somewhat more concretely, the participant is typically asked to respond quickly in some physical way (e.g., a button press, or by uttering something) to a visual or auditory stimulus, at which point a computer (or other specialized recording device) measures how long she took to respond (response latency, usually measured in milliseconds) and whether she made an error in that response.

Additionally, behavioral tasks vary roughly along three dimensions: (1) the nature of the required response; (2) how the stimuli are presented; and (3) the conditions under which the task is executed. The nature of the response required varies in the following ways (among others that are not relevant for most language studies): producing (e.g., picture naming, color naming, reading words aloud, writing), confirmation (e.g., self-paced reading), and classification (e.g., lexical decision, semantic categorization, sentence matching). Producing encompasses a broad class of response types including picture naming, reading aloud, and translation. With the possible exception of writing, production tasks come fairly naturally to people. Accordingly, the data have a higher degree of ecological validity. Picture naming is often used to tap into semantic processes, whereas reading aloud is usually used to tap into orthographic and phonological processes. Writing is more problematic on a number of levels, but we are beginning to see some strides in this area (e.g., Wengelin et al., 2009).

Confirmation is a relatively simple type of response. All it requires is that the participant confirm that some mental process has been completed. In psycholinguistics, the most commonly used confirmation technique is self-paced reading. Although it is often referred to as "a poor man's eyetracker," it has been shown to be a very successful technique in language processing (Mitchell, 2004). Perhaps a better term would be "a clever researcher's inexpensive alternative to eyetracking." The relative strengths of this technique, therefore, are its low cost paired with often relatively good correlation with more ecologically valid, natural-reading tasks (e.g., Traxler, Pickering, & McElree, 2002). Interestingly, due to the increase in the amount of texting and reading from small devices (which now includes a great deal of swipe technology), Mitchell (2004) speculates that in some ways, self-paced reading may be even more ecologically valid in the modern world than conventional sentence reading in an eyetracker.

But probably the most frequently exploited dimension of behavioral experiments are classification techniques (including categorization, judgment, etc.). Representative tasks include lexical decision ("Is it a word?", tapping into lexical access), semantic categorization ("Is it bigger than a bread box?", tapping into semantic processing), and sentence matching ("Are the two sentences the same?", tapping into grammatical knowledge). Some researchers (e.g., McElree & Griffith, 1995) have even attempted online (speeded) grammaticality-judgment tasks, the backbone of theorizing in the generative tradition.

It is difficult to describe the exact advantages of categorization; there is too much variation here. But perhaps the unifying characteristic is that it can distill complex linguistic/cognitive processes into simple responses that are not only amenable to analysis, but also still represent the underlying linguistic/cognitive complexity. The other response manipulations described above also do this, but in less desirable ways. For instance, a weakness with production data is that the data are difficult and problematic to analyze, whereas a weakness with confirmation is the trust

that one must confer upon the participant that she carried out the task as requested. When designed correctly, categorization tasks can ameliorate many shortcomings.

The second dimension, stimulus-presentation manipulations, involves how the participants perceive the experimental stimuli. They encompass a great variety of techniques. The most relevant manipulations are as follows: language mode, priming, degradation, timing, and distraction. Language mode determines whether the stimulus is in the spoken or visual (written or signed) mode. Depending on the experiment, stimulus presentation can be restricted to only one mode (e.g., self-paced reading), presented in either (visual or aural auditory lexical decision), or presented in both (cross-modal, e.g., participants listening to sentences while making a lexical decision on a word presented visually). With cross-modal experiments, researchers can often make the experimental design more complex without adversely affecting participant response, since working memory seems to have independent components for language mode (Baddeley & Hitch, 1974).

Priming is a technique in which two stimuli are presented sequentially, the goal being to ascertain the effect of the first stimulus (the prime) on the participant's response to the second stimulus (the target). This technique has been one of the most widely used of stimulus manipulations. Since primes can be related to targets in many ways (orthographically, phonologically, semantically, etc.), the paradigm has given us tremendous insight into the organization of the lexicon, as well as in other domains, like syntax. One of the earliest finds was semantic priming (Meyer & Schvaneveldt, 1971), where it was found that primes that are semantically related to targets facilitate reaction times to targets. Other types of priming include identity priming, form priming, and phonological priming (see Forster, 1999, for an in-depth discussion of the technique.).

Degradation is any manipulation that makes a stimulus more difficult to perceive. The most commonly used of these is masking, where, instead of allowing a stimulus to remain indefinitely on the screen (or to be replaced by a blank space), it is instead replaced quickly by another stimulus. Usually, the point is to control how long a participant can see a stimulus which, in turn, can tap into subconscious, bottom-up language processing. Another more recent variant of this is progressive demasking (Dufau, Stevens, & Grainger, 2008), which gives the participant more control over the duration of the mask, helping to control for individual variation in awareness thresholds. It is also possible to manipulate the luminance of a screen, or its foreground and background colors, to make a stimulus more difficult to discern (e.g., Tzur & Frost, 2007). Finally, outright distortion of the stimulus itself is possible, as in the case of letter-transposition studies (e.g., Guerrera & Forster, 2008) and alternating-case presentations (i.e., AlTeRnAtInG cAsE) (Coltheart & Freeman, 1974). The main advantage of the degradation manipulation is that a researcher can highlight the bottom-up or top-down processes that participants use to complete the task, since they are doing it under conditions where full, or complete, processing is restricted.

Timing can also be an important part of experimental design. The most prominent manipulation of time is during masked priming experiments were stimulus-onset asynchrony (SOA) is varied. SOA is the duration of time that a stimulus is shown before another stimulus appears, like the time between the presentation of two letter strings in a priming experiment (e.g., doctor >

NURSE). In a typical masked-priming paradigm with written words, the prime may be presented for as short as a single retrace on a monitor (e.g., 16.67 ms on a monitor with a 60 Hz refresh rate). However, prime durations of 40-60 ms are much more common in the literature. There is no upward limit on how long the prime can be presented, but in studies with native readers, any prime shown for longer than 70 ms is bound to be perceived in some way by most participants (Forster & Veres, 1998). One of the more commonly cited findings from the masked-priming literature is a series of studies by Ludovic Ferrand and Jonathan Grainger (e.g., Ferrand & Grainger, 1993; 1994), who found that priming at short prime durations seems to elicit orthographic effects, whereas priming at longer priming durations seems to elicit phonological effects (but see Holyk & Pexman, 2004; Tzur & Frost, 2007).

Distraction might also be used in an experiment. The advantage of this is that it can test whether two processes are independent, or whether there is an effect of one process on another. An example of this manipulation is the picture-word interference paradigm, where participants need either to name a picture while a letter string (usually a word) is presented on or near the picture itself. The participant would be asked to name the picture and ignore the word, or vice-versa. This particular manipulation has been very fruitful in looking at the effect of concepts (operationalized as pictures) on words in research on production for both monolinguals (e.g., Schriefers, Meyer, & Levelt, 1990; Starreveld & La Heij, 1996) and bilinguals (e.g., Costa, Miozzo, & Caramazza, 1999).

The last dimension is the set of conditions under which the task is executed. These conditions include, among others, response speed (e.g., Rapid Automatized Naming, or RAN; Denckla & Rudel, 1974), response inhibition (e.g., go/no-go; Siakaluk, Buchanan, & Westbury, 2003), and task complication (e.g., the reading-span task (Daneman & Carpenter, 1980), and response-contingency paradigms ["do x if a is true; do y if b is true") (Van Turennout, Hagoort, & Brown, 1997). These tend to be specialized manipulations that either address specific research questions, or simplify the task for special populations, or create trials where processing occurs in the absence of overt behavior (important sometimes for ERP research; see below).

At first blush, it is easy to criticize behavioral tasks in terms of ecological validity; no one normally carries out such tasks in their daily life. However, such criticism is counterbalanced by an important gain in another type of ecological validity. The element of time pressure (usually conveyed in the instructions and/or in the experiment itself as a timeout mechanism) gives researchers a window into how people naturally solve simple "problems" (in our case, "language problems") without recourse to reflection. In this sense, then, behavioral tasks can tap into language processing in "real time," which is really the main strength of psycholinguistics. Again, the startup costs for doing behavioral research are minimal. In fact, there are two widely used, free software programs to run most such experiments: *DMDX* (Forster & Forster, 2003) on the MS Windows operating system, and *PsyScope* (Cohen, MacWhinney, Flatt, & Provost, 1993) on the operating system developed by Apple Computer.

Although the startup and maintenance costs are higher, the two experimental paradigms covered next, particularly eyetracking, have the potential to make even further gains in ecological

validity. However, as the reader will see, that potential is clearly not their sole value in the study of language processing.

Eyetracking. Eye movements (and our ability to infer cognitive operations from them) occur specifically because of the limited visual acuity of the fovea, which has the proper density of cone-shaped photoreceptors necessary for detailed information processing, like letter or color identification. The fovea spans only about 2 degrees around the center of fixation (360 degrees representing the circular span of the eyeball). During normal reading and the processing of information in the static visual sense, therefore, the eyes must move in order to place information from different regions into foveal view. The eye movements we use to do this are known as saccades. Fixations occur when the eyeballs come to a relative halt (they are never completely still). Very little or no visual information is processed during a saccade, leaving most of the perceptual processing to fixations.

Eyetracking technology can monitor these eye movements with very good temporal and spatial precision. In being so precise, the technology gives us unprecedented opportunities to infer mental processes with a wide variety of new variables such as how long an item (e.g., a word or a picture) is fixated the first time it is gazed at, how long it is gazed at overall, how many times an item is fixated, how long the saccade is from one item to another, etc.

The two dominant experimental paradigms in eyetracking research are normal reading and the visual world paradigm. Normal reading is just that, and requires no further explanation. However, the visual world paradigm does. In such an experiment, the participant is viewing some sort of scene or a screen with objects or characters in it. The scene can range from fairly realistic (e.g., photos) to fairly abstract (e.g., line drawings on different parts of a screen). During a trial, the participant usually does one of two things while viewing a scene displayed on a computer screen: performing some task on the objects on the screen (e.g., *click on the picture of the dog*); or making some judgment about the congruity of the sentence and the scene. A typical manipulation involves the placement of some relevant distracter in the scene to see whether the participant's eyes are drawn to it more than a control. This paradigm has been productive in studies of single-language processing, particularly in spoken language, but so far has not extended very far into the study of bilingual language studies (but see the research collaboration of Viorica Marian and Michael Spivey (e.g., Marian & Spivey, 2003; Spivey & Marian, 1999). Just as eyetracking with normal reading has a less expensive, behavioral alternative in the self-paced reading paradigm, so the visual world paradigm has a less expensive, behavioral alternative in the tracking of arm movements during the manipulation of a computer mouse (see Spivey, Grosjean, & Knoblich, 2005). In fact, this paradigm can pick up on certain features of continual processing that are missed during saccadic behavior in equivalent eye-tracking experiments (Spivey, Grosjean, & Knoblich, 2005, p. 10,393).

Other than the increased precision and variety of measurements, eyetracking technology has contributed at least one major experimental manipulation that is unique: namely, the gaze-contingent display change, introduced first by Reder (1973) and applied to reading by McConkie and Rayner (1975, 1976). By taking advantage of either very fast display presentations or saccadic suppression (our inability to perceive visual information during a saccade), it is possible

to change the display in front of a participant without the participant noticing very easily, or even at all. In doing so, one can investigate the effect of partial parafoveal or peripheral processing of stimuli. This manipulation has been highly productive in the investigation of reading behavior. But to date there have been few studies of bilinguals that have used these paradigm (but see Pollatsek, Bolozky, Well, & Rayner, 1981, for an early study with Hebrew-English bilinguals).

In sum, most psycholinguists would agree that eyetracking provides extremely useful insights into language processing. As the technology has become more available by becoming both less expensive and more user-friendly over the years, eyetracking research into bilingualism has become more common, and will continue to do so.

Event-Related Potentials. Event-related potentials (ERPs) are electroencephalograms (EEGs, or in layperson's terms, "brainwaves") that have event codes (usually condition identifiers) inserted into the signal at time points associated with particular stimulus events in an experiment. When one averages the signals around the event codes, the common, underlying signals emerge, which subsequently are compared for significant differences. If particular signal patterns emerge again and again under similar conditions across experiments, they might become known as *components* (Otten & Rugg, 2005).

Components can be analyzed across conditions for amplitude, duration, and/or latency. A component's amplitude or latency can vary, and might represent either processing difficulty (or at least an unexpected event), or processing complexity (Kutas, Van Petten, & Kluender, 2006). However, amplitude, duration, and latency often interact in certain conditions (e.g., something that is very unusual takes longer to resolve, etc.). ERPs offer a window into mental processes with a temporal precision unmatched by any other experimental paradigm.[11]

ERP experiments also give language researchers the ability to obtain cognitive effects in the absence of almost any overt behavior on the part of the participant. This ability is unique to ERPs. In cases where responses are required, the response is often delayed until after a trial or block of trials. This technology has made it possible to measure online processing directly, without the measurement error that accompanies overt, non-linguistic behaviors.

In the study of language, the most prolific component has been known as the N400. It was discovered by Kutas and Hillyard (1980a, 1980b). This is a negative peak in the signal that begins to emerge at about 200 ms after stimulus onset, and peaks at about 400 ms (hence, the name). It originally appeared in semantically anomalous sentences at the point of the anomaly (i.e., upon the presentation of the word *sky* in the sentence *The little boy found a little sky in his pocket*.). As a result, the initial hypothesis was that the N400 represented difficulty with semantic integration. It is now believed to represent processing of any potentially meaningful event (Kutas & Federmeier, 2000; Kutas, Federmeier, Coulson, King, & Münte, 2000).

Two other important components for language processing are the Left Anterior Negativity (LAN) and the P600, both of which seem to represent syntactic processing in some way (Friederici, 2002; Hagoort, Brown, & Groothusen, 1993; Kutas, Van Petten, & Kluender, 2006; Osterhout, 1990; Osterhout & Holcomb, 1992). The LAN is an anterior (frontal lobe) waveform

with a negative trend between 300 and 400 ms after stimulus onset, whereas the P600 is a posterior (parietal lobe) waveform with a positive trend, peaking on average between 500 and 800 ms after stimulus onset. They both have been found to appear after grammatical violations such as after the word *understands* in **He is really understands it*. The exact functions that the two components represent are still under debate (Kutas, Van Petten, & Kluender, 2006).

There are more components useful to the study of language processing, but the scope of this paper prohibits such a review. Overall, ERPs offer a view of the timing of the mind and brain that cannot currently be seriously challenged by any other experimental paradigm in psycholinguistics. A review of recent ERP literature on bilinguals and bilingualism is provided by Moreno, Rodríguez-Fornells, and Laine (2008).

Naturally, each experimental paradigm comes with a history of research that determines which variables must be accounted or controlled for, either at the design or analysis level. These vary a great deal from paradigm to paradigm, and as such, are beyond the scope of this paper. However, a growing trend in psycholinguistics is to minimize the dichotomization of variables at the design phase (especially variables that are naturally continuous, like frequency), and instead account for the effects of such variables later during analysis, often through mixed effects modeling. See Baayen, Davidson, and Bates (2008) for a justification of this approach.

In this section, we discussed the wide variety of behavioral tasks that are available to researchers in the field of bilingualism and heritage speakers, although they have not been used much in this latter domain for various reasons. This wide variety of tasks is complemented by eyetracking and ERPs, where new insights have been and will be provided by new ways to peer into the mind and brain. In contrast to behavioral research, the main consideration for researchers interested in eyetracking and ERP research is startup and maintenance costs, which are fairly high. Nonetheless, they do and will offer many insights into language processing among heritage speakers.

The next section of the paper will provide suggestions for future research, as well as highlighting other important issues in how psycholinguistic research might be carried out with this particular population.

Suggestions for Further Research

So far, we have discussed several issues pertaining to the language processing of Spanish-English heritage speakers in assimilationist, English-dominant communities. After briefly covering offline studies that have pinpointed the aspects of language that may distinguish these bilinguals from bilinguals of other types (e.g., late bilinguals, bilinguals of other languages), we reviewed some of the few psycholinguistic studies that have been carried out with this group. Then we reviewed the psycholinguistic techniques that might be appropriate and/or feasible for many or most researchers interested in this population of bilinguals. Two conclusions can be drawn. First, generally speaking, what seems to distinguish this group may not be L1 attrition at all, but rather incomplete L1 acquisition, at least for many of them (Montrul, 2008). Second, and more specifically, the particular level of language that seems most affected is the semantic/pragmatic determinants of peripheral syntactic structures. This makes sense in light of

Ullman's (2001) model discussed above. If L1 acquisition is inhibited early in development, there is a good chance that not all the necessary syntactic structures would have been in place before procedural memory started to mature (e.g., through synaptic pruning) for more efficient processing, and consequently, declining plasticity. This account would predict that, through convergence, adult Spanish-English heritage speakers in the U.S. would exhibit persistent English-like characteristics in their grammatical processing of Spanish.

However, only a handful of studies have attempted to verify this view. There is a great deal of research that needs to be carried out still. Therefore, we will use the current section to explore techniques that we can use to fill this gap. Our discussion will be restricted to approaches that tap morpho-syntactic processing, since that is what seems to distinguish this group. Also, it will be restricted to techniques that show the most general promise; in other words, it will exclude techniques that cannot test a variety of structures. Finally, it will be restricted, moreover, to relatively inexpensive methods, which excludes functional neuroimaging and magnetoencephalography (MEG).[12] The reader should bear in mind that what follows is a list of potential points of departure for exploration. It would be presumptuous of us to set out a defined research agenda at this point since it is not yet totally clear what the relevant questions are. But we hope to provide a starting point nonetheless.

Some of the behavioral techniques reviewed above have been applied to syntactic processing. In terms of syntactic production data, the dominant paradigm is syntactic priming (Hartsuiker & Pickering, 2008). In this paradigm, a participant must listen to a sentence (or sentences) in one language, and then produce a sentence in another language. For instance, Hartsuiker, Pickering, and Veltkamp (2004) had bilingual participants engage in a pseudo-communication task with the confederate partner describing "pictures" (scripted sentences, in reality) in Spanish according to different structural conditions (e.g., active, passive, intransitive), and the participant describing her own picture in English freely with whatever structure she naturally generated. The goal was to look for evidence of cross-linguistic syntactic influence, which they found. Meijer and Fox Tree (2003) also found cross-linguistic syntactic priming using a similar methodology, not with the confederate participant, but rather with intervening sentences in the other language to which the participants had to respond.

One weakness of this sort of task is that it is somewhat offline in the sense that it is difficult to pressure participants to utter relatively novel sentences and measure the onset of those sentences without considerable measurement error. Instead, utterances are coded for their grammatical structure, and then compared by proportion of utterance to other structures. Participants usually are not under time pressure, so the task may conceal other cognitive operations that influenced the outcome.

This paradigm would be useful and easy to employ in addressing the grammatical processing of Spanish-English heritage speakers. Importantly, one can use a variety of grammatical structures in either English or Spanish to prime production in the other. It would be very illuminating to use this paradigm to determine which cross-linguistic constructions heritage speakers would be most sensitive to. But to our knowledge, this has not been done.

There are several behavioral techniques that address grammatical processing yet do not use utterances as dependent measures. The most widely used of these is probably the self-paced reading task (Just, Carpenter, & Woolley, 1982; Mitchell, 2004). In this task, a written segment (a word or string of words) of a sentence is presented on a screen one segment at a time. The participant is asked to read the segment (naturally incorporating the context of any preceding segment[s], if any), and to press a key to indicate that he has read and understood it, at which point the next segment appears. Researchers often use a moving window technique, where symbols representing letters (e.g., -------) in the sentence are progressively replaced with words in step with the participant's key presses; meanwhile, the preceding (already read) text turns back into abstract symbols. This approach, which mimics normal reading, has been highly prolific, not only in studies of monolingual processing, but also bilingual processing (Dussias & Cramer Scalz, 2008; Hoover, 1992; Hoover & Dwivedi, 1998; Igoa, Carreiras, & Meseguer, 1998). Probably the most intense scrutiny has been focused on syntactic ambiguity resolution, where it seems that speakers of some languages like English prefer to assume that relative clauses always modify the most recent noun in a given noun phrase, a "low-attachment" preference, whereas speakers of other languages, like Spanish, are relatively more comfortable assuming that relative clauses can modify more distant nouns in the phrase, a "high-attachment" preference (Cuetos & Mitchell, 1988; Dussias & Sagarra, 2007). In fact, this issue has been studied extensively using both self-paced reading and eyetracking.

Self-paced reading is a very important alternative to testing normal reading with eyetracking since the former can be used to pilot almost anything that can be considered in the latter (with the notable exception of parafoveal preview paradigms using gaze-contingent display changes). The main difference between the two approaches seems to be in the granularity of analysis, not in the types of structures one can test (Frenck-Mestre, 2005). For this reason, it is always worth considering the self-paced reading technique for normal reading since it is much less costly than eyetracking, and therefore, wiser in many cases (Mitchell, 2004).

Unfortunately, except for Montrul (2006b), little has been done using self-paced reading with Spanish-English heritage speakers in the U.S. A simple experiment using this paradigm would look for response differences between monolingual and heritage speakers on Spanish sentences that were either ungrammatical (e.g., incorrect DOM, or gender agreement) or grammatically awkward (e.g., excessive subject insertion). In fact, far more grammatical structures could be tested using this paradigm.

Of course, the converse is also true: anything that can be tested using self-paced reading can be tested using an eyetracker. In addition to the simple response-time measures that self-paced reading offers, eyetracking offers dozens of new variables that tell us about how long the eyes were looking at a particular word, how many letter spaces the eyes jumped on a particular saccade during reading, etc. In short, eyetracking offers a lot more, if such technology is available to the researcher. In fact, one might very well find effects that would have been missed in a self-paced reading task (Frenck-Mestre, 2005).

With respect to the testing of grammatical structure using eyetracking, there is a great deal of eyetracking research on syntactic ambiguity resolution during normal reading for monolinguals

(e.g., Carreiras & Clifton, 1999; see Frenck-Mestre & Pynte (2000) for a review) as well as bilingual processing (e.g., Dussias & Sagarra, 2007; Frenck-Mestre & Pynte, 1997). This parallel with the self-paced reading literature emphasizes the point made earlier with respect to normal reading: self-paced reading and eyetracking complement each other. High- versus low-attachment of relative clauses may or may not be relevant to Spanish-English heritage speakers in the U.S. We look forward to research in this area. But however interesting, this phenomenon is essentially a specific cross-linguistic peculiarity, filling only a small corner of the syntactic phenomena potentially relevant to heritage speakers.

It is only very recently that researchers have used eyetracking technology in normal reading to investigate alternative areas of grammatical processing in Spanish (e.g., Keating, 2009, for gender agreement). We expect more in the future, but among researchers investigating grammatical processing, one woefully underutilized technique within eyetracking during normal reading is the gaze-contingent display change. We feel that this technique offers a great deal of promise. For instance, it would be quite easy to present incorrect Preterite versus Imperfect verb forms in the parafovea of Spanish-speaking participants, only to change it back to the preferred verb form some short time after the verb is fixated (i.e., the fast priming technique used in Sereno & Rayner, 1992). The same could be done, possibly, with gender agreement. Trueswell and Kim (1998) used this technique to address the early processing of argument structure among monolingual English speakers, but it seldom has been used to look into other syntactic phenomena. We believe that this paradigm is an untapped resource for further research into grammatical processing among heritage speakers. Indeed, it is ability in eyetracking to use gaze-contingent display changes as well as increased granularity of analysis that help outweigh its costs relative to self-paced reading, at least with respect to normal reading.

A more recent alternative to both self-paced reading and eyetracking that addresses the relative weaknesses of both approaches is the maze task (Forster, Guerrera, & Elliot, 2009). In this task, participants see a series of screens. On each screen they need to indicate which of two alternative words (presented on the right and left) is the more grammatical continuation of the sentence that has been unfolding up to that point due to the very responses that the participant has made accurately (incorrect responses cancel the trial). This task encourages deep processing of sentences, which cannot be guaranteed easily with self-paced reading or eyetracking during normal reading, and it also allows a researcher to manipulate the distracters such that the poor alternative may address some specific syntactic issue. In Spanish, for instance, one might present a verb opposite the screen from a subject, where the subject was preferred in Spanish due to pragmatics. This would mean that the verb in this case would be dispreferred. But, as noted above in the review of Montrul (2004b) a Spanish-English heritage speaker might have difficulty with this sort of pragmatic/discourse-determined use of subject dropping (or insertion in this case).

The maze task also might be appropriate to address object expression in Spanish. In our lab in Canada, we are currently studying the processing of Spanish double-object constructions by monolingual speakers, early bilinguals with different levels of proficiency, and intermediate and advanced L2 learners. In fact, our study, based on Cuervo's (2007) work with L2 Spanish learners, incorporates the use of both the maze task and eye-tracking techniques. It will be

interesting to see how maze task will develop. We suspect that it holds a good deal of promise. This paradigm is more complex to program than a self-paced reading task, but the financial commitment is the same (low).

But eyetracking offers more than just the recording of eye movements during normal reading. The visual world paradigm (Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy , 1995), summarized above, offers a way to investigate a variety of grammatical structures during spoken-language comprehension. In contrast to normal reading, experiments in this paradigm are less constrained by the rigid structure of print in normal reading, and more open to the creativity of researchers. That is, it is easier to generate novel ways to test a variety of structures. This paradigm is easily applied to both children (Trueswell, Sekerina, Hill, & Logrip, 1999) and adults (Spivey, Tanenhaus, Eberhard, & Sedivy, 2002). Most conveniently, it is an ideal paradigm to investigate the interface between syntax and semantics or pragmatics (the areas ostensibly most vulnerable to incomplete acquisition among heritage speakers) since these latter variables can often be depicted visually. For instance, it would be fairly easy to follow up on the study by Montrul, Foote, and Perpiñán (2008) with this paradigm. One could create static visual scenes with objects that varied in grammatical gender to see where participants' eyes looked during the presentation of aural sentences in Spanish with full objects. One would expect that the groups with more automatic gender processing would initiate earlier gazes to the correct referent based on the offset of the definite article (marked for gender). One could also test for knowledge of Preterite versus Imperfect with visual scenes in which objects consistent with one or the other form of the verb were both present on the screen (see Altmann & Kamide, 2007, for an example a similar manipulation in English). One would expect that the participants with more stable representations of those verb forms would gaze earlier at the congruous item. Clearly, we suspect that the visual-world paradigm will be a very productive one with respect to how Spanish-English heritage speakers in the US process Spanish grammar.

The last experimental paradigm we will discuss, albeit briefly, is event-related potentials (ERPs). As noted above, Moreno and Kutas (2005) found a delayed N400 effect in the non-dominant language of Spanish-English heritage speakers, suggesting that lexical processing in Spanish had suffered particularly among the English-dominant speakers. However, despite the fact that they were using a semantic-anomaly paradigm, they still found a P600 in the dominant language (though not in the non-dominant language). Recall that this effect is usually associated with syntactic anomalies. The next step is then obvious: use a similar sample of participants in a syntactic-anomaly paradigm. Some recent steps have been taken in this direction with late L2 learners (e.g., Bowden, 2007; Isel, 2007; Mueller, Hirotani, & Friederici, 2007; Tokowicz & MacWhinney, 2005; Weber & Lavric, 2008), but not among Spanish-English heritage speakers. It might be possible to vary the subtle syntactic features related to semantics and pragmatics that seem to affect this group in a syntactic-violation paradigm. In fact, since there seems to be a semantic component to what the heritage speakers seem to be missing in grammatical processing, one might even find N400 effects in such a syntactic-anomaly paradigm.

The most promising characteristic of ERPs is temporal precision; an ERP study may detect subtleties that would otherwise be missed in behavioral or even eyetracking studies. But it may also be the case that the ERP components that traditionally emerge from such paradigms require

fairly non-subtle violations to do so. In any case, this line of research is worth pursuing. Our feeling is that one would have to start with a fairly high number of items in a restricted set of grammatical conditions in order to ensure sufficient power for detection.

In this paper we have covered four areas: (1) summaries of key linguistic studies carried out with Spanish heritage speakers in English assimilationist communities; (2) summaries of the few psycholinguistic studies that have carried out on the same; (3) a review of the various techniques that psycholinguists commonly use; and (4) suggestions for further psycholinguistic investigations into this population. Clearly, there is a dearth of psycholinguistic research into this community of speakers, who seem to be particularly characterized by difficulties in the processing of non-core, semantically driven syntax, a characterization that fits in with modern memory-based theories of monolingual and bilingual processing. By coupling the strengths of formal, linguistic description and psycholinguistic methods, we may one day be able to determine with some degree of certainty the challenges that this population of Spanish speakers face when processing their first language online, in real time.

References

Abutalebi, J., Cappa, S. F., & Perani, D. (2001). The bilingual brain as revealed by functional neuroimaging. *Bilingualism: Language and Cognition, 4*, 179-190.

Altmann, G. T. M., & Kamide, Y. (2007). The real-time mediation of visual attention by language and world knowledge: Linking anticipatory (and other) eye movements to linguistic processing. *Journal of Memory and Language, 57*, 502-518.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*, 390-412.

Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 8, pp. 47-89). New York: Academic Press.

Bowden, H. W. (2007). *Proficiency and second-language neurocognition: A study of Spanish as a first and second language*. (Unpublished doctoral dissertation). Georgetown University, Washington, DC.

Bruhn de Garavito, J. (2002). Verb raising in Spanish: A comparison of early and late bilinguals. In B. Skarabela, S. Fish, & A. Do (Eds.), *Proceedings of the 26th annual Boston University conference on language development* (pp. 84-94). Somerville, MA: Cascadilla.

Bullock, B. E., & Toribio, A. J. (2004) Introduction: Convergence as an emergent property in bilingual speech. *Bilingualism: Language and Cognition*, 91-93.

Carreiras, M., & Clifton, C., Jr. (1999). Another word on parsing relative clauses: Eyetracking evidence from Spanish and English. *Memory & Cognition, 27*, 826-833.

Carreiras, M., & Clifton, C., Jr. (2004). On the online study of language comprehension. In M. Carreiras & C. Clifton, Jr. (Eds.), *The online study of sentence comprehension: Eyetracking, ERPs and beyond* (pp. 1-14). New York: Psychology Press.

Cohen J. D., MacWhinney B., Flatt M., & Provost, J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments, and Computers, 25*, 257-271.

Coltheart, M., & Freeman, R. (1974). Case alternation impairs word identification. *Bulletin of the Psychonomic Society, 3*, 102-104.

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.

Costa, A., & Santesteban, M. (2004). Lexical access in bilingual speech production: Evidence from language switching in highly proficient bilinguals and L2 learners. *Journal of Memory and Language, 50,* 491-511.

Cuervo, M. C. (2007). Double objects in Spanish as a second language: Acquisition of morphosyntax and semantics. *Studies in Second Language Acquisition, 29,* 583-615.

Cuetos, F., & Mitchell, D. C. (1988). Cross-linguistic differences in parsing: Restrictions on the use of the late closure strategy in Spanish. *Cognition, 30,* 73–105.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19*(4), 450-466.

Denckla, M. B., & Rudel, R. (1974). Rapid "automatized" naming of pictured objects, colors, letters and numbers by normal children. *Cortex, 10*(2), 186-202.

Dufau, S., Stevens, M., & Grainger, J. (2008). Windows executable software for the progressive demasking task. *Behavior Research Methods, 40,* 33-37.

Dussias, P. E., & Cramer Scaltz, T. R. (2008). Spanish–English L2 speakers' use of subcategorization bias information in the resolution of temporary ambiguity during second language reading. *Acta Psycholgica, 128,* 501-513.

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

Ferrand, L., & Grainger, J. (1993). The time course orthographic and phonological code activation in the early phases of visual word recognition. *Bulletin of the Psychonomic Society, 31,* 119-122.

Ferrand, L., & Grainger, J. (1994). Effects of orthography are independent of phonology in masked form priming. *The Quarterly Journal of Experimental Psychology, 47A*, 365-382.

Field, J. (2008). Face to face with the ghost in the machine: Psycholinguistics and TESOL. *TESOL Quarterly, 42*, 361-374.

Foote, R. (2007). A psycholinguistic investigation of agreement in Spanish and English monolinguals and bilinguals (Doctoral dissertation), Retrieved from Dissertation Abstracts International. (68, 0545).

Foote, R. (2010). Age of acquisition and proficiency as factors in language production: Agreement in bilinguals. *Bilingualism: Language and Cognition, 13*(2), 99-118.

Forster, K. I. (1999). The microgenesis of priming effects in lexical access. *Brain and Language, 68,* 5-15.

Forster, K. I., & Forster, J. C. (2003). DMDX: A windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers, 35*, 116-124.

Forster, K. I., Guerrera, C., & Elliot, L. (2009). The maze task: Measuring forced incremental sentence processing time. *Behavior Research Methods, 41,* 163-171.

Forster, K. I., & Veres, C. (1998). The prime lexicality effect: Form priming as a function of prime awareness, lexical status, and discrimination difficulty. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*, 498-514.

Friederici, A. D. (2002) Toward a neural basis of auditory sentence processing. *Trends in Cognitive Science, 6*(2), 78-84.

Frenck-Mestre, C. (2005). Eye-movement recording as a tool for studying syntactic processing in a second language: A review of methodologies and experimental findings. *Second Language Research, 21,* 175-198.

Frenck-Mestre, C., & Pynte, J. (1997). Syntactic ambiguity resolution while reading in second and native languages. *The Quarterly Journal of Experimental Psychology, 50A,* 119–48.

Frenck-Mestre, C., & Pynte, J. (2000) Resolving syntactic ambiguities: cross-linguistic differences? In M. De Vincenzi & V. Lombardo (Eds.), *Cross-linguistic perspectives on language processing* (pp. 119-148). Dordrecht: Kluwer.

Garnham, A., Garrod, S., & Sanford, A. (2006). Observations on the past and future of psycholinguistics. In M. J. Traxler & M. A. Gernsbacher (Eds.), *Handbook of psycholinguistics* (2nd ed.) (pp. 1-18). Amsterdam; Boston: Elsevier/Academic Press.

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

Guerrera, C., & Forster, K. I. (2008). Masked form priming with extreme transposition. *Language and Cognitive Processes, 23*, 117-142.

Hagoort, P., Brown, C. M., & Groothusen, J. (1993). The syntactic positive shift (SPS) as an ERP measure of syntactic processing. *Language and Cognitive Processes, 8*, 439–483.

Hartsuiker, R. J., & Pickering, M. J. (2008). Language integration in bilingual sentence production. *Acta Psychologica, 128,* 479–489.

Hartsuiker, R. J., Pickering, M. J., & Veltkamp, E. (2004). Is syntax separate or shared between languages? Cross-linguistic syntactic priming in Spanish-English bilinguals. *Psychological Science, 15,* 409-414.

Hernandez, A. E., & Li, P. (2007). Age of acquisition: Its neural and computational mechanisms. *Psychological Bulletin, 133,* 638-650.

Holyk, G. G., & Pexman, P. M. (2004). The elusive nature of early phonological priming effects: Are there individual differences? *Brain and Language, 90,* 353-367.

Hoover, M. L. (1992). Sentence processing strategies in Spanish and English. *Journal of Psycholinguistic Research, 21,* 275-299.

Hoover, M. L., & Dwivedi, V. D. (1998). Syntactic processing by skilled bilinguals. *Language Learning, 48,* 1-29.

Igoa, J. M., Carreiras, M., & Meseguer, E. (1998). A study on late closure in Spanish: Principle-grounded vs. frequency-based accounts of attachment preferences. *Quarterly Journal of Experimental Psychology, 51A,* 561-592.

Isel, F. (2007). Syntactic and referential processes in second-language learners: Event-related brain potential evidence. *Neuroreport, 18,* 1885-1889.

Juffs, A. (2001). Psycholinguistically oriented second language research. *Annual Review of Applied Linguistics, 21,* 207-220.

Just, M. A., Carpenter, P. A., & Woolley, J. D. (1982). Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General, 111*, 228-238.

Keating, G. D. (2009). Sensitivity to violations of gender agreement in native and nonnative Spanish: An eye-movement investigation. *Language Learning, 59,* 503-535.

Kohnert, K. J., Hernandez, A. E., and Bates, E. (1998). Bilingual performance on the Boston Naming Test: Preliminary norms in Spanish and English. *Brain and Language, 65*, 422-440.

Köpke, B. (2002). Activation thresholds and non-pathological first language attrition. In F. Fabbro (Ed.), *Advances in the neurolinguistics of bilingualism*: *Essays in honor of Michel Paradis* (pp. 119-142). Undine: Forum.

Köpke, B. (2004). Neurolinguistic aspects of attrition. *Journal of Neurolinguistics, 17,* 3-30.

Kormos, J. (2006). *Speech production and second language acquisition*. Mahwah, NJ: Lawrence Erlbaum Associates.

Kroll, J. F., & Sunderman, G. (2003). Cognitive processes in second language learners and bilinguals: The development of lexical and conceptual representations. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 104-129). Malden, MA: Blackwell.

Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Science, 4*, 463-470.

Kutas, M., Federmeier, K. D., Coulson, S., King, J. W., & Münte, T. F. (2000). Language. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of psychophysiology* (2nd ed., pp. 576–601). Cambridge: Cambridge University Press.

Kutas, M., & Hillyard, S. A. (1980a). Event-related brain potentials to semantically inappropriate and surprisingly large words. *Biological Psychology, 11*, 99–116.

Kutas, M., & Hillyard, S. A. (1980b). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science, 207*(4427), 203-205.

Kutas, M., Van Petten, C. K., & Kluender, R. (2006). Psycholinguistics electrified II (1994-2005). In M. J. Traxler & M. A. Gernsbacher (Eds.), *Handbook of psycholinguistics, 2nd ed*. (pp. 659-724). New York: Elsevier.

La Heij, W. (2005). Selection processes in monolingual and bilingual lexical access. In J.

F. Kroll & A. M. B. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 289-307). New York: Oxford.

Larsson, J. P., Constan, F. V., Sebastian-Galles, N., & Deco, G. (2008). Lexical plasticity in early bilinguals does not alter phoneme categories: I. neurodynamical modeling. *Journal of Cognitive Neuroscience, 20,* 76-94.

Marian, V., & Spivey, M. (2003). Competing activation in bilingual language processing: Within- and between-language competition. *Bilingualism: Language and Cognition, 6,* 97-115.

McConkie, G. W., & Rayner, K. (1975). The span of the effective stimulus during a fixation in reading. *Perception & Psychophysics, 17*, 578–86

McConkie, G. W., & Rayner, K. (1976). Asymmetry of the perceptual span in reading. *Bulletin of the Psychonomic Society, 8*, 365–68

McElree, B., & Griffith, T. (1995). Syntactic and thematic processing in sentence comprehension: Evidence for a temporal dissociation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 134-157.

McKee, C. (1996). Online methods. In D. McDaniel, C. McKee, & H. Smith Cairns (Eds.), *Methods of assessing children's syntax* (pp. 189-208). Cambridge, MA: The MIT Press.

Meijer, P. J. A., & Fox Tree, J. E. (2003). Building syntactic structures in speaking: A bilingual exploration. *Experimental Psychology, 50*, 184-195.

Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology, 90,* 227-234.

Mitchell, D. C. (2004). Online methods in language processing: Introduction and a historical review. In M. Carreiras & C. Clifton, Jr. (Eds.), *The online study of sentence comprehension: Eyetracking, ERPs and beyond* (pp. 15-32). New York: Psychology Press.

Montrul, S. (2002). Incomplete acquisition and attrition of Spanish tense/aspect distinctions in adult bilinguals. *Bilingualism: Language and Cognition,* 39-68.

Montrul, S. (2004a). Psycholinguistic evidence for split intransitivity in Spanish second language acquisition. *Applied Psycholinguistics, 25,* 239-267.

Montrul, S. (2004b). Subject and object expression in Spanish heritage speakers: A case of morphosyntactic convergence. *Bilingualism: Language and Cognition, 7,* 125-142.

Montrul, S. (2006a). Incomplete acquisition as a feature of bilingual and L2 grammars. In R. Slabakova, S. A. Montrul, & P. Prévost (Eds.), *Inquiries in linguistic development: In honor of Lydia White* (pp. 335-359). Philadelphia: John Benjamins.

Montrul, S. (2006b). On the bilingual competence of Spanish heritage speakers: Syntax, lexical-semantics and processing. *International Journal of Bilingualism, 10,* 37-69.

Montrul, S. (2008). *Incomplete acquisition in bilingualism: Re-examining the age factor*. Amsterdam: John Benjamins.

Montrul, S., Foote, R., & Perpiñán, S. (2008). Gender agreement in adult second language learners and Spanish heritage speakers: The effects of age and context of acquisition. *Language Learning, 58,* 503-553.

Montrul, S., & Bowles, M. (2009). Back to basics: Incomplete knowledge of Differential Object Marking in Spanish heritage speakers. *Bilingualism: Language and Cognition, 12,* 363-383.

Moreno, E. M., & Kutas, M. (2005). Processing semantic anomalies in two languages: An electrophysiological exploration in both languages of Spanish-English bilinguals. *Cognitive Brain Research, 22,* 205-220.

Moreno, E. M., Rodríguez-Fornells, A., & Laine, M. (2008). Event-related potentials (ERPs) in the study of bilingual language processing. *Journal of Neurolinguistics, 21,* 477-508.

Mueller, J. L., Hirotani, M., & Friederici, A. D. (2007). ERP evidence for different strategies in the processing of case markers in native speakers and non-native learners. *BMC Neuroscience, 8*(18). Retrieved from http://www.biomedcentral.com/content/pdf/1471-2202-8-18.pdf

Nicol, J. L., Teller, M.., & Greth, D. (2001). Production of verb agreement in monolingual, bilingual, and second-language speakers. In J. L. Nicol (Ed.), *One mind, two languages: Bilingual language processing* (pp. 117-133). Malden, MA: Blackwell Publishers.

Osterhout, L. (1990). Event-related brain potentials elicited during sentence comprehension. (Unpublished doctoral dissertation). Tufts University, Medford, MA.

Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language, 31,* 785–806.

Otten, L. J., & Rugg, M. D. (2005). Interpreting event-related brain potentials. In T. C. Handy (Ed.), *Event-related potentials: A methods handbook* (pp. 3-16). Cambridge, MA: The MIT Press.

Pallier, C., Dehaene, S., Poline, J. -B., LeBihan, D., Argenti, A. -M., Dupoux, E., & Mehler, J. (2003). Brain imaging of language plasticity in adopted adults: Can a second language replace the first? *Cerebral Cortex, 13*(2), 155-161.

Pollatsek, A., Bolozky, S., Well, A. D., and Rayner, K. (1981). Asymmetries in the perceptual span for Israeli readers. *Brain and Language, 14,* 174-180.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124,* 372-422.

Reder, S. M. (1973) On-line monitoring of eye position signals in contingent and noncontingent paradigms. *Behaviour Research Methods & Instrumentation, 5,* 218–228.

Schriefers, H., Meyer, A. S., & Levelt, W. J. M. (1990). Exploring the time course of lexical access in language production: Picture–word interference studies. *Journal of Memory & Language, 29,* 86-102.

Sebastian-Gallés, N., Rodriguez-Fornells, A., de Diego-Balaguer, R., & Diaz, B. (2006). First- and second-language phonological representations in the mental lexicon. *Journal of Cognitive Neuroscience, 18,* 1277-1291.

Sereno, S. C., & Rayner, K. (1992). Fast priming during eye fixations in reading. *Journal of Experimental Psychology: Human Perception and Performance, 18,* 173-184.

Schwartz, A. I., & Areas Da Luz Fontes, A. B. (2008). Cross-language mediated priming: Effects of context and lexical relationship. *Bilingualism: Language and Cognition, 11*, 95-110.

Schwartz, A. I., & Kroll, J. F. (2006). Language processing in bilingual speakers. In M. J. Traxler & M. A. Gernsbacher (Eds.), *Handbook of psycholinguistics* (2nd ed.) (pp. 967-999). Amsterdam; Boston: Elsevier/Academic Press.

Schwieter, J. W. (2008). A psycholinguistic investigation of language selectivity in bilingual speech production. (Doctoral dissertation). Retrieved from Dissertation Abstracts International, (68, 3828).

Sereno, S. C., & Rayner, K. (1992). Fast priming during eye fixations in reading. *Journal of Experimental Psychology: Human Perception and Performance, 18*, 173-184.

Siakaluk, P. D., Buchanan, L., & Westbury, C. (2003). The effect of semantic distance in yes/no and go/no-go semantic categorization tasks. *Memory & Cognition, 31*(1), 100-113.

Silva-Corvalán, C. (1991). Spanish language attrition in a contact situation with English. In H. Seliger & R. Vago (Eds.), *First language attrition* (pp. 151-171). Cambridge: Cambridge University Press.

Silva-Corvalán, C. (1994). *Language contact and change.* Oxford: Oxford University Press.

Silverberg, S., & Samuel, A. G. (2004). The effect of age of second language acquisition on the representation and processing of second language words. *Journal of Memory and Language, 51,* 381-398.

Slotnick, S. D. (2005). Source localization of ERP generators. In T. C. Handy (Ed.), *Event-related potentials: A methods handbook* (pp. 149-166). Cambridge, MA: The MIT Press.

Sorace, A. (2000). Gradients in auxiliary selection with intransitive verbs. *Language, 76,* 859-890.

Spivey, M. J., Grosjean, M., & Knoblich, G. (2005). Continuous attraction toward phonological competitors. *Proceedings of the National Academy of Sciences, 102*(29), 10393-10398.

Spivey, M., & Marian, V. (1999). Crosstalk between native and second languages: Partial activation of an irrelevant lexicon. *Psychological Science, 10,* 281-284.

Spivey, M. J., Tanenhaus, M. K., Eberhard, K. M., & Sedivy, J. C. (2002). Eye movements and spoken language comprehension: Effects of visual context on syntactic ambiguity resolution. *Cognitive Psychology, 45,* 447-481.

Starreveld, P. A., & La Heij, W. (1996). Time-course analysis of semantic and orthographic context effects in picture naming. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 22*, 869-918.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science, 268*(5217), 1632–1634.

Tokowicz, N., & MacWhinney, B. (2005). Implicit and explicit measures of sensitivity to violations in second language grammar: An event-related potential investigation. *Studies in Second Language Acquisition, 27*, 173-204.

Toribio, A. J., & Nye, C. (2006). Restructuring of reverse psychological predicate in bilingual Spanish. In J. Montreuil & C. Nishida (Eds.), *New perspectives in Romance linguistics, Volume I: Morphology, syntax, semantics, and pragmatics. Selected papers from the 35th Linguistic Symposium on Romance Languages (LSRL)* (pp. 263-277). Amsterdam: John Benjamins.

Traxler, M. J., Pickering, M. J., & McElree, B. (2002). Coercion in sentence processing: evidence from eye-movements and self-paced reading. *Journal of Memory and Language, 47*, 530-547.

Trueswell, J. C., & Kim, A. E. (1998). How to prune a garden path by nipping it in the bud: Fast priming of verb argument structure. *Journal of Memory and Language, 39,* 102-123.

Trueswell, J. C., Sekerina, I., Hill, N., & Logrip, M. (1999). The kindergarten-path effect: Studying on-line sentence processing in young children. *Cognition, 73,* 89–134.

Tzur, B., & Frost, R. (2007). SOA does not reveal the absolute time course of cognitive processing in fast priming experiments. *Journal of Memory and Language, 56,* 321-335.

Ullman, M. T. (2001). The neural basis of lexicon and grammar in first and second language: the declarative/procedural model. *Bilingualism: Language and Cognition, 4*(1), 105–122.

Van Turennout, M., Hagoort, P., & Brown, C. M. (1997). Electrophysiological evidence on the time course of semantic and phonological processes in speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*(4), 787-806.

Weber, K., & Lavric, A. (2008). Syntactic anomaly elicits a lexico-semantic (N400) ERP effect in the second language but not the first. *Psychophysiology, 45*, 920-925.

Weber-Fox, C. M., & Neville, H. J. (1996). Maturational constraints on functional specializations for language processing: ERP and behavioral evidence in bilingual speakers. *Journal of Cognitive Neuroscience, 8*(3), *231-256.*

Wengelin, Å., Torrance, M., Holmqvist, K., Simpson, S., Galbraith, D., Johansson, V., & Johansson, R. (2009). Combined eyetracking and keystroke-logging methods for studying cognitive processes in text production. *Behavior Research Methods, 41,* 337-351.

Zapata, G. C., Sánchez, L., & Toribio, A. J. (2005). Contact and contracting Spanish. *International Journal of Bilingualism, 9*, 377-395.

Zentella, A. C. (1997). *Growing up bilingual: Puerto Rican children in New York*. Malden, MA: Blackwell.

Notes

1. The term assimilationist refers to a society where conformity to the dominant culture is expected. This expectation applies to language as well as customs.

2. Note that this group is important not because it is unique, but rather because it is an accessible population that represents some of the less accessible populations.

3. Note that a variety of psycholinguistic studies have examined Spanish early bilinguals (Catalan-Spanish) in Spain (e.g., Costa, Miozzo, & Caramazza, 1999; Costa & Santesteban, 2004; Hernández, 2009; Larsson, Constan, Sebastian-Galles, & Deco, 2008; Sebastian-Galles, Rodriguez-Fornells, de Diego-Balaguer, & Diaz, 2006), but this population's bilingualism exhibits different social and linguistic characteristics than those of heritage speakers in the U.S.

4. See *Online Methods* section below, where the notion of *online processing* is more carefully described.

5. The results of their study were similar to those reported by Toribio and Nye's (2005) in their investigation of heritage speakers' interpretation and production of psychological predicates. The participants in Toribio and Nye's study also exhibited robust knowledge of Agreement and Case licensing, properties of the core syntax, but "vulnerability in the lexico-semantic module…and in the domain of discourse/pragmatics" (p. 273), which seems to mirror the findings in all the studies described in this section.

6. Some recent psycholinguistic studies have investigated Spanish-English bilinguals (e.g., Schwartz & Areas Da Luz Fontes, 2008; Silverberg & Samuel, 2004), but they have not focused on early bilinguals or heritage speakers per se.

7. It is not considered appropriate to call brain potentials "brainwaves" anymore. Hence, the quotation marks. However, this layperson's term is still the easiest way to describe the phenomena to a novice. For this reason, we use it here. For a more accurate description of these potentials, see the *Online Methods* section below.

8. This is one case where response times would be ridden with measurement error associated with the time it takes participants to create a sentence. Error is the appropriate measure in this case.

9. fMRI and MEG are two further brain-imaging technologies that have become more common in studies of language processing. However, these techniques are quite expensive and thus probably not well suited, at this time, to deal with heterogeneous groups such as heritage speakers where large sample sizes are most appropriate. So for that reason, and in order to keep the scope of this article reasonable, we will not cover them here.

10. This is especially true if one includes researchers in first-language and literacy development under the domain of psycholinguistics. In order to keep the scope of this article reasonable, however, we will refer the reader to other papers that summarize the techniques available to younger age groups (e.g., McKee, 1996).

11. For reasons that are beyond the scope of this paper, the spatial resolution of ERPs is problematic: it is not easy to tell where the source of any given component is in the brain. That said, there is a healthy literature on how to improve this weakness in the paradigm (see Slotnick, 2005, for a review).

12. For a review of the functional research showing how age of acquisition affects bilingualism, see Hernandez and Li (2007). And see Abutalebi, Cappa, and Perani (2001) for a different view.