

Article

Comparing Single-Word Insertions and Multi-Word Alternations in Bilingual Speech: Insights from Pupillometry

Michael A. Johns ^{1,*}  and Paola E. Dussias ²¹ Institute for Systems Research, University of Maryland, College Park, MD 20910, USA² Department of Spanish, Italian and Portuguese, Pennsylvania State University, State College, PA 16802, USA

* Correspondence: maj@umd.edu

Abstract: Prominent sociolinguistic theories of language mixing have posited that single-word insertions of one language into the other are the result of a distinct process than multi-word alternations between two languages given that the former overwhelmingly surface morphosyntactically integrated into the surrounding language. To date, this distinction has not been tested in comprehension. The present study makes use of pupillometry to examine the online processing of single-word insertions and multi-word alternations by highly proficient Spanish-English bilinguals in Puerto Rico. Participants heard sentences containing target noun/adjective pairs (1) in unilingual Spanish, (2) where the Spanish noun was replaced with its English translation equivalent, followed by a Spanish post-nominal adjective, and (3) where both the noun and adjective appeared in English with the adjective occurring in the English pre-nominal position. Both types of language mixing elicit larger pupillary responses when compared to unilingual Spanish speech, though the magnitude of this difference depends on the grammatical gender of the target noun. Importantly, single-word insertions and multi-word alternations did not differ from one another. Taken together, these findings suggest that morphosyntactic integration is not the defining feature of single-word insertions, at least in comprehension, and that the comprehension system is tuned to the distributional properties of bilingual speech.

Keywords: codeswitching; language mixing; lone items; nonce borrowings; pupillometry



Citation: Johns, Michael A., and Paola E. Dussias. 2022. Comparing Single-Word Insertions and Multi-Word Alternations in Bilingual Speech: Insights from Pupillometry. *Languages* 7: 267. <https://doi.org/10.3390/languages7040267>

Academic Editors: Julia Hofweber, Jan Patrick Zeller and Jeanine Treffers-Daller

Received: 4 August 2022

Accepted: 13 October 2022

Published: 20 October 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

One goal of sociolinguistic approaches to the study of bilingualism has been to distinguish instances of codeswitching—broadly construed as the fluid use of two or more languages within a single utterance or linguistic unit—from other contact-induced phenomena. Numerous models (Myers-Scotton 1993), constraints (Bhatt and Bolonyai 2011), hypotheses (Poplack et al. 1988; Backus 2014), and typologies (Muysken 2000, 2013) have been proposed that attempt to define and delimit codeswitching. Despite a large body of literature and voluminous scholarly activity on the subject, however, researchers have yet to agree upon a single definition of codeswitching and on whether it is qualitatively different from other language contact phenomena (Poplack 2018, p. 1). This is perhaps most evident with respect to the status of single-word insertions (i.e., lone other-language items, henceforth, LOLIs) of one language into another. A decades-long debate has been waged over whether such insertions constitute bona fide codeswitches, or whether they are the result of a distinct process (Poplack and Sankoff 1984; Poplack et al. 1988; Sankoff et al. 1990; Poplack 2012).

First proposed by Poplack et al. (1988), the nonce borrowing hypothesis posits a categorical distinction between LOLIs and multi-word alternations where the two languages are morphosyntactically distinct. In this approach, the latter constitute codeswitches, while the former are termed nonce borrowings or “spontaneous borrowings [that] assume the morphological and syntactic identity of the recipient language prior to achieving the social

characteristics of established loanwords [emphasis original]" (Poplack 2012, p. 645). Thus, the defining characteristic of a nonce borrowing is its morphosyntactic integration into the surrounding other-language context, a feature that distinguishes it from codeswitching—defined by Poplack as “the juxtaposition” of elements from two languages that retain the morphosyntactic characteristics of their lexifier language (Poplack 1993, p. 355). This distinction between codeswitches and nonce borrowings proposed by Poplack and colleagues is not without criticism: some researchers posit that LOLIs and multi-word alternations are the result of the same fundamental process (Myers-Scotton 1993, pp. 205–7) while others argue against the spontaneity of nonce borrowings and in favor of a continuous approach driven by factors such as frequency (Stammers and Deuchar 2012) and saliency (Backus 2014). To avoid confusion, we will adopt the term *language mixing* as an umbrella term referring to any instances in which a bilingual uses their two languages within the same utterance or conversation (both LOLIs and multi-word alternations thus fall into this category).

To date, most studies of LOLIs have been conducted solely on corpus data, but psycholinguistic studies examining LOLIs remain scant (though see Johns et al. 2019 for one example, as well as Tsoukala et al. 2021 for a computational approach.). This is due, in part, to the differing questions that drive each field of research: while sociolinguistic approaches seek to understand the social and structural factors that constrain language mixing, psycholinguistic approaches have focused on the cognitive mechanisms underlying bilingual language control. This division is not absolute, however, nor are the questions themselves unrelated to one another. For example, recent psycholinguistic models such as the Production-Distribution-Comprehension (MacDonald 2013) and the P-Chain (Dell and Chang 2014) models posit an intimate link between the distributional properties of the input and the processing strategies individuals adopt in order to comprehend it. As such, there has been a recent push to bring psycholinguistic data to bear on sociolinguistic models of language mixing and vice versa (Han et al. 2022; Tomić and Kaan 2022; Gosselin and Sabourin 2021; Johns and Steuck 2021; Kheder and Kaan 2021; Kroll et al. 2021; Beatty-Martínez et al. 2020; Beatty-Martínez and Dussias 2017). With respect to LOLIs, this is of particular importance given that sociolinguistic proposals such as the nonce borrowing hypothesis make explicit claims about the cognitive mechanisms underlying LOLIs vis à vis multi-word alternations (Poplack and Dion 2012, p. 309).

The present study builds upon the growing body of literature linking the social and cognitive aspects of language mixing (e.g., Johns et al. 2019; Beatty-Martínez and Dussias 2019; Valdés Kroff et al. 2017; Guzzardo Tamargo et al. 2016; Hofweber et al. 2016) and takes a first step at systematically comparing LOLIs to multi-word alternations during online comprehension in highly proficient early Spanish-English bilinguals. The study presented here also makes use of a methodology relatively new to the study of sentence processing—pupillometry—to test predictions about the cognitive architecture of bilingualism derived from the sociolinguistic literature on codeswitching. Before describing the present study, we discuss in more depth the nonce borrowing hypothesis, and provide justification for why studying this linguistic behavior through the lens of language comprehension—despite it being a bilingual speech production phenomenon—is a suitable and desirable step.

Before discussing the literature surrounding LOLIs, it is first important to describe the terminology used in this paper. First, we use the term ‘lone other-language item’, or LOLI, to describe single-word insertions of one language into another. This term is purely descriptive and is not intended to convey any theory-specific information. This is in contrast to the term ‘nonce borrowing’, which refers specifically to the theoretical concept put forward by Poplack and colleagues (Poplack et al. 1988; Sankoff et al. 1990; Poplack 2012; Poplack and Dion 2012). We also use multi-word alternations (henceforth, MWA) to describe one of our experimental conditions (see Section 3.2) where the sentence starts in Spanish but finishes in a multi-word stretch of English. Lastly, while the terms ‘single-word insertion’ and ‘multi-word alternation’ are based loosely on Muysken’s (2013) typology of

language contact phenomenon, we again opt for the terms LOLI and MWA as *descriptive*, rather than theoretical, terms.

2. Background

2.1. Lone Other-Language Items

The first large-scale comprehensive analysis of LOLIs was conducted by [Poplack et al. \(1988\)](#) on the spontaneous speech of 120 French-English bilinguals in the Ottawa-Hull region of Canada. The resulting two-million-word corpus yielded nearly 20,000 instances of single English-origin insertions occurring in otherwise French speech. While these insertions constituted less than one percent of all words in the corpus, the authors were surprised to find that their behavior was largely homogenous. For example, English-origin nouns uttered at least twice in the corpus were consistently assigned French grammatical gender and surfaced with near ubiquitous use of the French null plural affix. Likewise, all but ten of the 20,000 English-origin insertions appeared in syntactic positions congruent with French and incongruent with English. Thus, despite their English origin, these LOLIs—be they hapax legomena or highly frequent—overwhelmingly surfaced with French morphosyntactic features “starting quite early, that is, as soon as minimal frequency of use can be detected” (*ibid.* p. 67). The authors termed this process nonce borrowing, referring to the “abrupt and categorical” morphosyntactic integration observed for such items that liken them to established loanwords ([Poplack 2012](#), p. 645; [Sankoff et al. 1990](#), p. 64). Similar patterns of nonce borrowing have been argued for lone English-origin insertions in Tamil ([Sankoff et al. 1990](#)), Turkish ([Adalar and Tagliamonte 1998](#)), Ukrainian ([Budzhak-Jones 1998](#)), Igbo ([Eze 1998](#)), Persian ([Samar and Meechan 1998](#)), Acadian French ([Turpin 1998](#)), Korean ([Shin 2002](#)), and Spanish ([Torres Cacoulios and Aaron 2003](#)), demonstrating the pervasive nature of this phenomenon across bilingual speech. Indeed, LOLIs have been shown to be “by far the predominant manifestation of bilingual mixing in every language pair empirically studied” ([Poplack and Dion 2012](#), pp. 280–81)—be they morphosyntactically integrated or instances of a community-specific bilingual strategy (e.g., the use of kinship terms from English in otherwise Spanish utterances; see [Aaron 2015](#)). These LOLIs of one language into an utterance of the other language have been given the umbrella term of lone other-language items (LOLIs).

The nonce borrowing hypothesis not only makes predictions about the morphosyntactic features of LOLIs; it also provides a cognitive rationale for why speakers opt for such a strategy in the first place:

“We can only speculate that speakers eschew code-switching single words, in the sense of switching both lexicons and grammars (as they do when engaging in multiword code-switching), because the cognitive and processing costs of doing so for a lone other-language item are appreciably greater than those incurred by simply allowing the already activated grammar to continue operating, handling native and etymologically foreign forms in the same way.” ([Poplack and Dion 2012](#), p. 309)

In other words, [Poplack and Dion \(2012\)](#) argue that codeswitching requires that both the lexicons and grammars are switched, while for LOLIs, it is only the lexicon that is switched (to insert the lexical item from the other language) while the grammar of the recipient language remains active. Because only the lexicon is switched, the LOLI adopts the morphosyntactic properties of the recipient language, which serves as a facilitative mechanism to avoid the purportedly cognitively demanding process of switching both lexicon and grammar for only one word, and the concomitant costs.

While evidence exists for these so-called switch costs (e.g., [Meuter and Allport 1999](#); [Costa and Santesteban 2004](#); [Hernandez et al. 2001](#); [Bobb and Wodniecka 2013](#); [Bultena et al. 2015a, 2015b](#), among many others), recent studies have cast doubt on the idea that switching languages, particularly voluntarily or spontaneous language-switching, necessarily incurs processing costs. For example, [Blanco-Elorrieta and Pykkänen \(2017, p. 17\)](#) found that both behavioral costs and neural effects in the executive control network were eliminated

when participants voluntarily switched languages (see also [Gollan and Ferreira 2009](#)). Similar findings have been found during reading, both at the sentence level ([Dussias 1997](#); [Gullifer et al. 2013](#)) and with respect to LOLIs ([Johns et al. 2019](#)) as well as in spontaneous production ([Johns and Steuck 2021](#)).

In light of these recent findings, the present study seeks to determine if LOLIs incur similar processing demands when compared to multi-word codeswitches, or if the morphosyntactic integration of LOLIs distinguishes them from multi-word codeswitches. For reasons that will be discussed below, we opt to examine comprehension as opposed to production to better measure the cognitive difficulty of these respective categories during online processing. To accomplish this, 58 highly proficient early Spanish-English bilinguals listened to unilingual Spanish sentences and sentences containing LOLIs and multi-word alternations while pupil size—an indicator of cognitive load—was recorded.

2.2. How Production Shapes Comprehension

The nonce borrowing hypothesis makes predictions about the cognitive difficulty associated with switching languages based on data from spontaneous bilingual production; however, it is currently unclear how to accurately measure cognitive load online during spontaneous production. The difficulty in assessing cognitive load during bilingual production is further compounded by potential confounds related to ecological validity, naturalness, and volition. Measuring cognitive load during language comprehension, however, has proven to be not only fruitful in understanding the relationship between a bilingual's two languages and the cognitive mechanisms that regulate them, but also in examining the relationship between production and comprehension.

Mounting evidence has shown an intimate connection between production patterns and comprehension strategies, particularly with respect to codeswitched speech. One influential model that has sought to capture this relationship is [MacDonald's \(2013\)](#) Production-Distribution-Comprehension (PDC) model. The PDC argues that cognitive limits on language planning and production, such as those related to memory and retrieval, shape the distributional properties of language. Comprehenders adapt to these distributional properties, with comprehension facilitated when the input matches what is expected based on prior experience, and impaired when it does not. As [MacDonald \(2013, p. 1\)](#) states:

“In language . . . the input to the perceiver is itself the consequence of language behavior—it is the utterances produced by other language users, who have their own cognitive systems presumably shaped by their own experiences.”

For example, object relative clauses are less frequent in speech than subject relative clauses, and generally more difficult to process ([Traxler et al. 2002](#)). However, when given greater exposure to object relative clauses, comprehenders process them similarly to subject relative clauses ([Wells et al. 2009](#)). This example demonstrates how comprehension is guided in large part by an individual's experience with language production.

The connection between distributional patterns in production and comprehension strategies has also been attested when examining codeswitched speech. [Guzzardo Tamargo et al. \(2016\)](#) analyzed a corpus of spontaneous Spanish-English bilingual speech and found that switches between the Spanish progressive auxiliary *estar* ‘to be’ and an English gerund (e.g., *están cleaning*, ‘they are cleaning’) are much more frequent than structurally similar switches between the Spanish perfect auxiliary *haber* ‘to have’ and an English participle (e.g., *han cleaned*, ‘they have cleaned’). The authors then asked whether this production asymmetry was reflected in individuals' online processing of these two different types of switches. Using eye-tracking, the authors found significantly longer reading times in the codeswitched perfect construction than the codeswitched progressive construction, confirming that the production asymmetry found in spontaneous speech was reflected in individuals' online processing.

Similar patterns have been attested when Spanish-English bilingual code-switchers process a mixed noun phrase composed of a Spanish determiner followed by an English noun. Corpus studies have reported a so-called “masculine default strategy,” which is

characterized by an overwhelming tendency for speakers to use the masculine Spanish determiner *el* in conjunction with an English noun, regardless of the gender of its Spanish translation equivalent (Herring et al. 2010; Valdés Kroff 2016; Cruz 2021). Take, for example, this utterance from the Bangor Miami corpus of Spanish-English bilingual speech (Deuchar et al. 2014):

pero no tenían el *flag out there*?

‘but they didn’t have the flag out there?’ (sastre9.fem2)

In this example, the English word *flag* translates into Spanish as the feminine noun *la bandera*, but nonetheless is uttered with the masculine determiner *el*. Following up on these findings in spontaneous production, Valdés Kroff et al. (2017) used a visual world paradigm to examine monolingual and bilingual speakers’ ability to use grammatical gender encoded in Spanish determiners to predict an upcoming noun. In this task, participants saw two-picture displays where both pictures could represent objects with the same grammatical gender in Spanish (e.g., both masculine or both feminine; the “same-gender condition”) or with different genders (e.g., one masculine and one feminine; the “different-gender condition”). In the different-gender condition, the authors predicted that participants would look to the target item sooner upon hearing the determiner (*el* or *la*). This is so because it is in this condition that the gender information in the determiner is informative (i.e., participants do not need to wait until they hear the target noun to direct their gaze to it). While this was the case for the monolingual Spanish speakers, who showed facilitation for both the masculine and feminine determiners in the different-gender condition, the bilingual speakers with experience with codeswitching showed facilitation with feminine determiners but not with the masculine ones. The authors attributed this finding to the bilinguals’ experience with codeswitching, and in particular with their experience using the masculine determiner as a default form.

Lastly, Johns et al. (2019) sought to determine how presenting codeswitched stimuli separately from Spanish and English unilingual stimuli vis à vis presenting them interleaved with unilingual stimuli affected online processing. The authors found that presenting both codeswitched and unilingual stimuli together in the same experimental session facilitated the processing of LOLIs compared to presenting them independently from unilingual stimuli. They argued that this is due to the overall distribution of codeswitching in spontaneous speech: it is unlikely for a bilingual to encounter long stretches of speech where every utterance contains a codeswitch, as evidenced by findings from corpus data showing that in spontaneous speech, codeswitching is relatively infrequent (Torres Cacoullos and Travis 2018; Guzzardo Tamargo et al. 2016, p. 142; Poplack et al. 1988, p. 57; Aaron 2015, p. 461). Thus, given bilinguals’ experience with codeswitching as a relatively bursty phenomenon (see Guzmán et al. 2017), presenting codeswitch stimuli in a manner congruent with the bilingual speakers’ prior experience facilitated processing.

This same reasoning—that distributional properties in the input in turn affect individuals’ online processing strategies—can be applied to the study of LOLIs, allowing us to test predictions based on production by examining comprehension. With respect to LOLIs, spontaneous production data illustrate their predominance vis à vis other forms of language mixing (Poplack and Dion 2012, pp. 280–81) and their overwhelming tendency to appear morphosyntactically integrated. If morphosyntactic integration serves a facilitative role in production, then the sheer frequency of its usage should lead to similar facilitative effects in comprehension. Likewise, the nonce borrowing hypothesis makes specific predictions about the difficulty associated with LOLIs and multi-word alternations: while the former “involves recourse to one grammar only, that of the recipient language” (Torres Cacoullos and Aaron 2003, p. 289), the latter requires “switching both grammars and lexicons” (Poplack and Dion 2012, p. 309), a process which is assumed to be more cognitively demanding. Because it is currently unclear how to accurately measure cognitive load online during spontaneous production, we turn to comprehension as a means to understand the cognitive difficulty associated with LOLIs and multi-word alternations.

2.3. Using Pupillometry to Assess Online Processing

To measure cognitive load during online language comprehension, the present study makes use of a technique relatively new to the language sciences: pupillometry. Psychological and neurological work over the past several decades has shown that the pupillary response is linked not only to changes in ambient luminance, but also to aspects of the sympathetic nervous system (Goldwater 1972) as well as the locus coeruleus and the norepinephric system (LC-NE; Samuels and Szabadi 2008; Aston-Jones and Cohen 2005). The LC-NE has been associated with memory retrieval (Beatty and Kahneman 1966; Attar et al. 2013), selective attention (Foot and Morrison 1987), and arousal (Bradshaw 1967). For example, when selective attention is engaged either due to increased cognitive demands or attentional requirements, the pupil involuntarily dilates due to its connection to the LC-NE. Recently, pupillometry has been applied to a variety of language-related processes, such as effortful speech processing (Kuchinsky et al. 2013), lexical retrieval (Schmidtke 2014), bilingual cognate facilitation (Guasch et al. 2017), and the processing of language mixing (Byers-Heinlein et al. 2017), highlighting its sensitivity to a variety of language processing phenomena (see Schmidtke 2018). Under this approach, an increase in pupil size with respect to a particular linguistic stimulus is assumed to be indicative of greater cognitive load resulting from an increase in the allocation of attentional resources (Gabay et al. 2011; Alnæs et al. 2014).

While sparse, there exist some prior studies that use pupillometry to examine bilingual language processing. For example, Byers-Heinlein and colleagues (2017) used pupillometry to assess how bilingual infants process nouns of one language in a carrier phrase of the other language (e.g., “Find the *chien!*”, ‘find the dog’). They found increased pupil sizes compared to a unilingual condition when the carrier phrase was in the dominant language and the target noun was in the non-dominant language, but no changes in pupil size when the carrier phrase was in the non-dominant language and the target in the dominant language. More recently, Beatty-Martínez and colleagues (2021) used pupillometry to investigate how different types of bilingual process language-mixed speech. In their study, the authors found that bilinguals who tend to keep their languages to separate communicative contexts showed larger pupil responses to language-mixed stimuli, while bilinguals who tend to mix languages more frequently showed no difference.

3. Materials and Methods

3.1. Participants

Sixty-eight early Spanish-English bilinguals from the University of Puerto Rico, Río Piedras campus were recruited for this study, with eight being excluded as they were unable to complete both sessions of the experiment and two being excluded due to missing and/or corrupted data. Fifty-eight participants were thus included in the analyses. All participants completed a detailed language history questionnaire that asked about all languages the participants know or were learning, how often each was used and in which contexts (e.g., with friends, at home, at school), self-rated proficiency in each language, and self-reported codeswitching tendencies. Participants were highly proficient in both languages (see Table 1). All participants reported regularly engaging in codeswitching, self-reporting their highest use of codeswitching in their free time (7.27/9, where 9 indicates ‘Always’), followed by at home (6.89/9), school (6.87/9), and work (4.66/9). Participants also completed a lexical decision task as an online measure of lexical access in both English and Spanish. Participants saw 100 letter strings in English and Spanish, each language presented in its own block, consisting of 50 real words and 50 pseudowords that were phonotactically legal in their respective language (e.g., English: *veem*, Spanish: *panselo*). Pseudowords were created in Wuggy (Keuleers and Brysbaert 2010). Participants were instructed to indicate, as quickly and accurately as possible, if the letter string was a real word of English or Spanish (depending on the block). The dprime function in the psych package (v. 0.5.0; Makowski 2018) was used to calculate d-prime scores for each language and are presented in Table 1. Of the self-rated proficiency measures and the d-prime scores,

two-sample paired *t*-tests revealed participants rated themselves more proficient in Spanish for self-rated speaking, reading, and understanding abilities. There were no differences for self-rated writing abilities nor for d-prime scores.

Table 1. Self-rated proficiency and lexical decisions scores (standard deviations in parentheses).

	Speaking	Reading	Writing	Understanding	D-Prime
English	8.93/10 (0.88)	9.24/10 (0.84)	8.81/10 (1.22)	9.43/10 (0.82)	3.27 (0.64)
Spanish	9.41/10 (0.79)	9.52/10 (0.73)	9.00/10 (1.15)	9.71/10 (0.56)	3.16 (0.58)
<i>p</i> -value	<0.01 *	0.01 *	0.18	<0.01 *	0.28

Note: “*” refers to $p < 0.05$

3.2. Stimuli

Experimental stimuli consisted of a noun/adjective pair that always appeared at the end of the sentence. This site was chosen given the differing word orders between English and Spanish: while English adjectives are prenominal, Spanish adjectives are largely post-nominal (e.g., “el gato perezoso”, literally ‘the cat lazy’). As a result, morphosyntactically integrated lone English nouns would surface with post-nominal Spanish adjectives (“el cat perezoso”) while multi-word alternations would preserve the English pre-nominal word order (“el lazy cat”). A total of 90 Spanish nouns, each matched with a unique Spanish adjective, were selected. Half of the nouns were masculine, and the other half were feminine. The 90 noun/adjective pairs were divided into three lists of 30 items, with nouns matched by frequency between the three lists (A vs. B: $t = 0.28$, $p = 0.78$; A vs. C: $t = 0.57$, $p = 0.57$; B vs. C: $t = 0.31$, $p = 0.76$). Frequencies were taken from the Puerto Rico sub-section of the *Corpus del español* (Davies 2016). Lastly, the noun/adjective pair appeared half the time as the direct object of the main verb (e.g., *El novio le regaló unas pantallas plateadas*, ‘the boyfriend bought him/her some silver earrings’) and half the time in a prepositional phrase that was an adjunct of the main verb (e.g., *Los prisioneros querían escaparse de la cárcel segura*, ‘the prisoners wanted to escape from the safe prison’)¹.

Experimental stimuli appeared in three conditions: the Unilingual Spanish condition, where the entire sentence was presented in Spanish (“el ladrón robó el anillo brillante”, ‘the robber stole the shiny ring’); the lone other-language item (LOLI) condition, where the target noun was replaced with its English translation equivalent (“el ladrón robó el ring brillante”); and the multi-word alternation (MWA) condition, where the entire noun/adjective pair was replaced with its English equivalent (“el ladrón robó el shiny ring”). Within each condition, half of the items had feminine target nouns and half had masculine target nouns; in addition, half had the target noun/adjective pairs as a direct object of the main verb and half in a prepositional phrase that was an adjunct to the main verb.

For both the LOLI and MWA conditions, the gender of the preceding determiner was always congruent with the Spanish translation of the target noun. In other words, when the Spanish translation of the English target noun was masculine, it was preceded by a Spanish masculine determiner; if the Spanish translation was feminine, it was preceded by a Spanish feminine determiner. Note that, while infrequent, the Spanish feminine determiner is used with English nouns whose Spanish translation equivalent is feminine (e.g., Valdés Kroff 2016; Cruz 2021). In addition, Beatty-Martínez and Dussias (2017, p. 179) found no differences in the processing of English nouns whose Spanish translation equivalent was feminine (e.g., “spoon”, ‘la cuchara’) when preceded by a feminine versus a masculine determiner (e.g., “la spoon” versus “el spoon”). Given this, we chose to keep the determiner congruent with the gender of the Spanish translation equivalent in order to compare the more frequent use of the masculine determiner with a following English noun to the less frequent feminine determiner. This was also done to determine if the gender of the determiner interacted with morphosyntactic integration across the LOLI and MWA

conditions. Example stimuli are given in Table 2, and a full list of stimuli is available as Supplementary Materials.

Table 2. Example stimuli.

Gender	Condition	Pre-Target	Target
Feminine	Unilingual LOLI MWA	El novio le regaló unas	pantallas plateadas. ¹ earrings plateadas silver earrings.
Masculine	Unilingual LOLI MWA	El ladrón robó el	anillo brillante. ² ring brillante. shiny ring.

¹ ‘the boyfriend bought him/her some silver earrings’; ² ‘the thief stole the shiny ring’.

Stimuli were recorded by a native Puerto Rican Spanish-English bilingual codeswitcher using a Shure SM35 head-worn condenser cardioid microphone and were normalized for intensity using Praat (Boersma and Weenink 2022). Language-mixed stimuli were not spliced, instead produced as-is by the speaker in order to avoid any unnatural-sounding artefacts that may influence comprehension, given that speakers have been shown to produce cues as to an upcoming switch in languages (e.g., Fricke et al. 2016). The speaker was instructed to produce the stimuli as she normally would were she speaking with another Spanish-English bilingual codeswitcher, with the caveat that the English elements should be produced with English-like qualities. In other words, we asked that the speaker pronounce, e.g., ‘silver’ as close to [silvə] as possible, as opposed to [silver] with Spanish-like qualities.

Participants were presented with 30 experimental items from each of the three conditions for a total of 90 experimental items. Each participant thus heard all 90 adjective/noun pairs only once. Items were counterbalanced across participants such that, while all 90 noun/adjective pairs occurred in all three conditions, a given participant heard each noun/adjective pair in only one of these conditions. An additional 120 filler items were created and recorded by the same speaker and were randomly interspersed with experimental items. Half of the filler items were in unilingual Spanish and half contained multi-word alternations that did not occur at the same site as those in the experimental items. Filler items were designed to be structurally similar to the experimental items, e.g., consisting of a subject, noun, and object (or optional adjunct). Example filler items are given in Table 3.

Table 3. Example Filler Items.

Unilingual Fillers	Codeswitched Fillers
El pintor abrió el cuaderno para dibujar. ‘The painter opened the notebook to draw.’	El entrenador hizo el ejercicio with the whole class. ‘The trainer did the exercise with the whole class.’
La investigadora trabajadora encontró la evidencia. ‘The hardworking investigator found the evidence.’	La voluntaria compró los juguetes for the child. ‘The volunteer bought the toys for the child.’

Stimuli were pseudorandomized such that a maximum of two experimental stimuli could occur in succession and were presented in two sessions that occurred on two separate days: the first session consisted of 30 unilingual stimuli along with 60 unilingual filler items, while the second session consisted of 30 LOLI stimuli, 30 MWA stimuli, and 60 codeswitched fillers. The first session also included an additional 30 items that were analyzed for a different experiment but are not included in the present study (see Johns and Dussias 2021). The time between sessions could vary but maximally took place within 3 days of each other. Separating the unilingual stimuli and the language-mixed stimuli into two sessions was done to isolate any effects of the language-mixed stimuli from the unilingual stimuli, providing a more stable baseline condition representative of a unilingual

‘mode’ of discourse (Green and Abutalebi 2013; Green and Wei 2014). Nonetheless, separating stimuli in this way is not entirely reflective of spontaneous production (e.g., Johns et al. 2019); we return to this issue and its ramifications in the discussion.

3.3. Design

Each trial consisted of the following (see Sirois and Brisson 2014; Schmidtke 2018 for details regarding constraints in designing pupillometric studies):

1. A 1000 ms neutral period without audio displaying a drawing of an ear in a central 300-by-300-pixel interest area.
2. The stimulus, which began playing after the 1000 ms baseline period; the drawing of the ear remained on-screen such that the luminance did not change.
3. The target period, beginning at the onset of the target noun/adjective pair and extending 3000 ms; the drawing of the ear remained on-screen.
4. The cue to repeat the sentence aloud, beginning after the 3000 ms offset period and indicated by the drawing of the ear changing to the drawing of a mouth.

Participants were asked to repeat the sentence allowed when cued for two reasons: first, it ensured that the participants were attending to the stimuli; second, the pupillary response is strongest when the participant must make an overt response, be it oral or behavioral (Sirois and Brisson 2014). Figure 1 illustrates the trial design. For filler items, which did not contain a target, the image of the ear remained on screen for 2000 ms after the sentence finished playing.

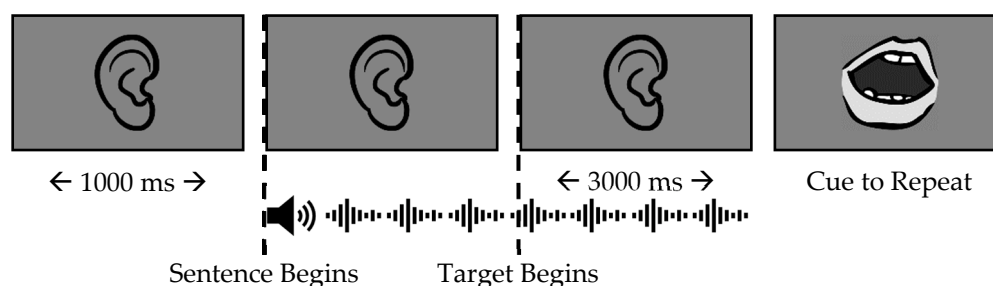


Figure 1. Trial Design.

3.4. Procedure

Participants were seated in a quiet room with a steady light source. Each session began with informed consent, followed by the eye-tracking task. Data were collected on an EyeLink Portable Duo eye-tracker (SR Research) recording at 1000 Hz in head-stabilized mode using the right eye pupil and corneal reflection. Participants first completed a brief practice followed by two experimental blocks of 30 experimental items and 30 filler items each. Participants were encouraged to take a break between each block. Calibration was performed before the practice and before the start of the second block (maximum error < 1.0 degrees; average error < 0.5 degrees). After the eye-tracking task, participants completed the language history questionnaire and the lexical decision task. For the lexical decision task, the order of the blocks (English, Spanish) was counterbalanced by participant. At the end of each session, participants were paid at a rate of 10 USD/hour. A short debriefing was also held at the end of the second session.

3.5. Data Pre-Processing and Cleaning

Data were extracted using SR Research DataViewer (v. 4.2.1) using a ‘Time Course (Binning) Analysis’ with the interest period beginning 300 ms pre-stimulus onset and ending 3000 ms post-stimulus onset. Data were automatically binned into 20 ms time bins, and the following variables were extracted for each bin:

1. The average right-eye pupil size across all non-blink samples.
2. The average right-eye x- and y-gaze position across all non-blink samples.

3. The proportion of samples that were in a blink event.
4. The proportion of samples that were in a saccade event.
5. The proportion of samples that fell in the central interest area.
6. The proportion of samples that fell outside of the central interest area.
7. The proportion of samples that fell off-screen.

The average pupil size in the 300 ms pre-stimulus onset period was used to baseline correct the pupil size during the 3000 ms post-stimulus onset period. Baseline correction was performed by subtracting each trial's average pupil size in the pre-stimulus onset period from the average pupil size in each bin of the post-stimulus onset period (see [van Rij et al. 2019](#), pp. 3–6). Trials where the total proportion of samples that occurred in a blink or saccade event or outside of the central interest area exceeded 25% in either the pre-stimulus onset or post-stimulus onset period were excluded from further analysis, resulting in 10.33%² of trials being excluded. The final dataset consisted of 5611 trials of the original 6959.

3.6. Analysis

Pupil size was modeled as a time-dependent variable using generalized additive mixed-effects models (GAMMs) with the `bam` function in the `mgcv` package (v 1.8-40; [Wood 2011](#); see also [Wood 2017](#) for a comprehensive overview) in R (v 4.1.3; [R Core Team 2022](#)). Conditions were modeled using ordered factor difference smooths ([Wieling 2018](#)), which allow distinguishing between differences in both the intercept (i.e., height of the curve) and trajectories (i.e., the shape of the curve) across conditions. The model also contained a smooth term capturing the interaction between the x- and y-gaze positions, used to account for the effects of gaze position on pupil size ([Gagl et al. 2011](#)), as well as random reference/difference smooths by participant and by item ([Sóskuthy 2021](#), p. 14). The ordered factor difference smooths were constructed in order to compare the effects of gender (masculine vs. feminine) across the three main conditions (unilingual Spanish, LOLI, and MWA). In order to obtain all of the necessary comparisons across these conditions, the model was subsequently relevelled.

Model criticism was performed using the `check` function in the `mgcViz` package (v 0.1.9, [Fasiolo et al. 2020](#)), and visualizations were performed using the `itsadug` package (v 2.4, [van Rij et al. 2020](#)). All smooths were specified to use thin-plate regression splines, and the model was specified with a scaled-*t* distribution. Lastly, an autoregressive-1 (AR-1) model was included to account for autocorrelation in the residuals which is characteristic of time-series data like the pupillary response ([Baayen et al. 2022](#)). A rho value of 0.95 was found to sufficiently reduce autocorrelation (autocorrelation at lag 1 = 0.19). Full R code can be found at <https://osf.io/2cbmd/>; accessed on 31 July 2022.

4. Results

Table 4 below provides a summary of the significant effects, while Tables S2–S5 in the supplemental materials provide detailed model summaries for the relevelled model. Fitted smooths are visualized in Figure 2, while estimated parametric (i.e., overall height) values are presented in Figure 3. There were significant overall height differences for masculine nouns when comparing the MWA and Unilingual conditions, and for feminine nouns when comparing the LOLI and Unilingual conditions. In both cases, the MWA and LOLI conditions elicited larger overall pupillary responses when compared to their respective Unilingual conditions (see Figure 3). There was also a significant interactive parametric effect of gender between the MWA and LOLI conditions. For masculine nouns, there are no discernable differences in the overall height of the pupillary response between the MWA and LOLI conditions. For feminine nouns, however, LOLIs elicit larger pupillary responses than the MWA condition (see Figure 3). There were also significant non-linear differences between the LOLI and Unilingual conditions for both masculine and feminine nouns. In both cases, the LOLI condition elicited a more sustained pupillary response compared to

the Unilingual condition, though this was more pronounced for feminine nouns due to the additional difference in overall height (see Figure 2).

Table 4. Summary of significant effects.

Parametric Coefficients	Estimate	SE	t-Value	p-Value	
MWA, M-Unilingual, M	13.399	6.145	2.181	0.029	*
LOLI, F-Unilingual, F	14.933	6.314	2.365	0.018	*
(MWA, F-M)–(LOLI, F-M)	−14.619	7.407	−1.974	0.048	*
Smooth Terms	EDF	Ref.df	F-Value	p-Value	
LOLI, M-Uni., M	8.310	10.746	2.136	0.015	*
LOLI, F-Uni., F	7.772	10.105	2.262	0.012	*

Note: “*” refers to $p < 0.05$.

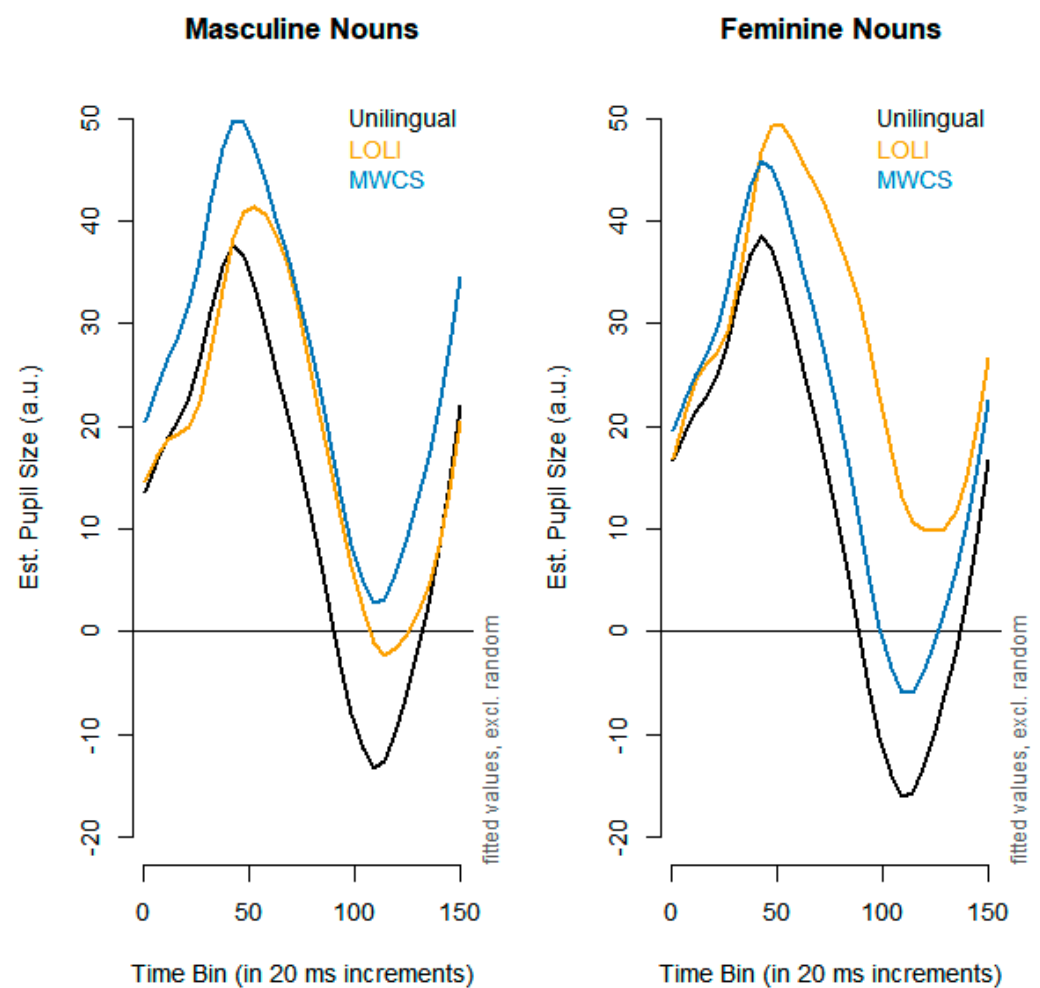


Figure 2. Fitted smooths.

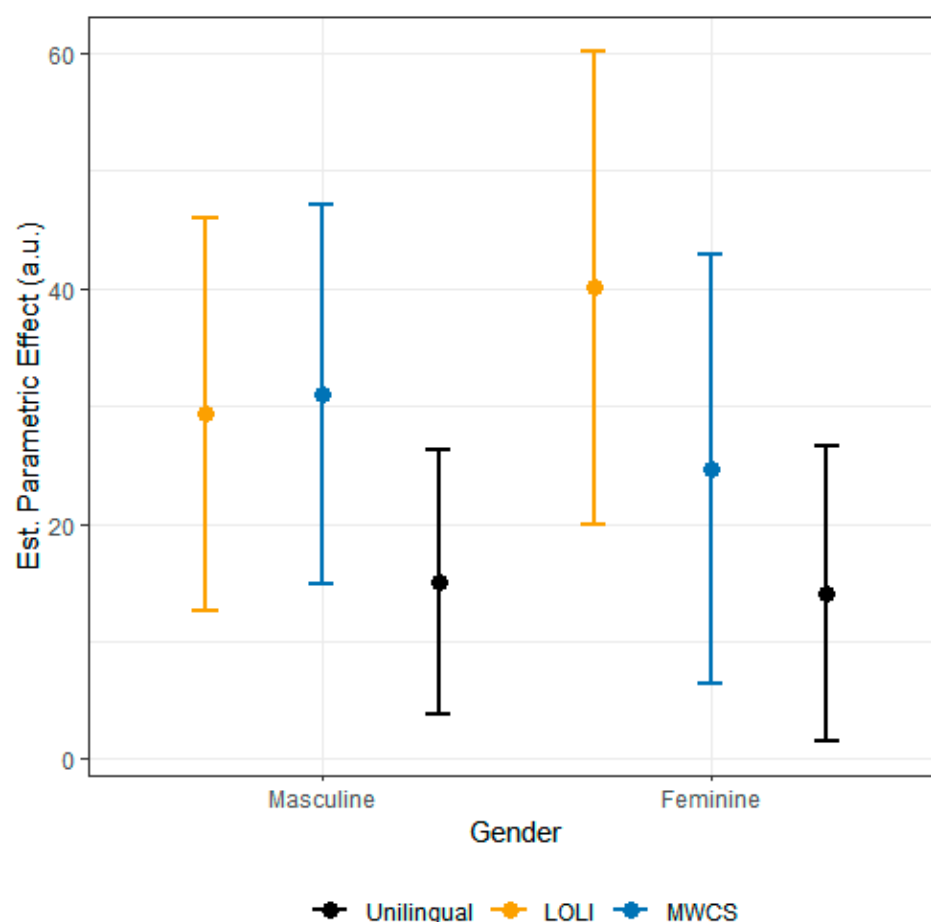


Figure 3. Fitted parametric (intercept) values.

5. Discussion

Recall that prior research on LOLIs in spontaneous bilingual speech has posited that LOLIs are less cognitively demanding to produce than multi-word alternations, particularly when they are morphosyntactically integrated into the surrounding language (e.g., [Poplack and Dion 2012](#), p. 309). Considering that LOLIs are both the most frequent form of language mixing and are overwhelmingly morphosyntactically integrated in production, it follows that they should also exhibit some sort of facilitative benefits in processing when compared to multi-word alternations, if at least because they are more frequent. The results of the present study paint a more nuanced picture, however, where the processing of LOLIs and multi-word alternations vis à vis unilingual speech is, in part, modulated by the grammatical gender of the target noun (or its English translation equivalent).

To summarize the findings discussed above, masculine nouns in the MWA condition elicited overall larger pupillary responses compared to the Unilingual condition, while masculine nouns in the LOLI condition differed only in their trajectory when compared to the Unilingual condition, eliciting larger pupillary responses towards the middle of the epoch only (see Figure 2). For feminine nouns, only those in the LOLI condition differed from the Unilingual condition, eliciting both a larger overall pupillary response as well as a more sustained pupillary response towards the end of the epoch (see Figure 2). While feminine nouns in the MWA condition did not differ from those in the Unilingual condition, neither did they differ from feminine nouns in the LOLI condition.

With respect to the LOLIs, it is not surprising that feminine target nouns, preceded by a feminine Spanish determiner, elicited overall larger pupillary responses compared to masculine nouns, preceded by a masculine Spanish determiner. Recall that previous research has established that switches between a Spanish determiner and an English noun overwhelmingly surface with the masculine Spanish determiner—regardless of the gram-

matical gender of the Spanish translation of the English noun—and that switches between a feminine Spanish determiner and an English noun are rare (e.g., [Valdés Kroff 2016](#); though see [Cruz 2021](#), who found higher rates in a Spanish-English bilingual community in southern Arizona). This same asymmetry present in production has also been replicated in comprehension, such that switches involving a masculine Spanish determiner and an English noun are less cognitively demanding (or more expected) than switches with a feminine Spanish determiner (see [Beatty-Martínez and Dussias 2017, 2019](#) and references therein). This same asymmetry surfaces in the present study with respect to the overall height differences when comparing LOLIs to the Unilingual condition. This is not to say, however, that masculine target nouns in the LOLI condition are processed identically to their unilingual counterparts; rather, these items elicited a more sustained pupillary response towards the middle of the epoch after the initial peak (approximately 1000–2000 ms post-stimulus onset). One reason for this sustained response may be due to the post-nominal Spanish adjective that occurred after the English noun, which may be unexpected. This unexpectedness may be driven by the frequency with which such language mixes occur, as we discuss below.

For the multi-word alternations, target noun/adjective pairs preceded by a masculine Spanish determiner elicit overall higher pupillary responses compared to the Unilingual condition. When preceded by a feminine Spanish determiner, there are no differences between multi-word alternations and either the Unilingual or LOLI conditions. Numerically, and as can be seen in [Figure 2](#), multi-word alternations preceded by feminine Spanish determiners do elicit marginally larger pupillary responses, however, and it is important to note that they do not differ from those preceded by masculine Spanish determiners. At first glance, it is unclear why multi-word alternations with masculine Spanish determiners would elicit larger pupillary responses given that this is the preferred strategy used in production. What differs, however, is the presence of the pre-nominal English adjective.

In one recent analysis of spontaneous Spanish-English bilingual production, [Torres Cacoullos and Vélez-Avilés \(forthcoming\)](#) examined language mixing in adjective/noun pairs in a Spanish-English bilingual community in northern New Mexico. The authors found that lone item incorporation of nouns (e.g., “*unas earrings plateadas*”) and adjectives (e.g., “*unas pantallas silver*”) was the most frequent form of language mixing, occurring 185 times in their sample of 527 language-mixed noun/adjective pairs. Of these, lone English-origin nouns with post-nominal Spanish adjectives (“*unas earrings plateadas*”, as in the present study) occurred 44 times. Alternations between a Spanish determiner and an English adjective phrase (e.g., “*unas silver earrings*”, as in the present study) were similarly as frequent as the lone English-origin nouns, occurring 38 times. Interestingly, the authors also found that alternations that occurred at the phrasal boundary were more frequent than both of these language mixes, occurring 157 times. In particular, utterances such as “*El novio le regaló some silver earrings*” were more frequent than both lone English nouns and multi-word alternations after the determiner, occurring 70 times. While this data ultimately comes from a different bilingual community than the one under study here, it nonetheless highlights the importance of understanding the relative frequencies of the different forms of language mixing that can occur between determiners, nouns, and adjectives. Indeed, that these Spanish-English bilinguals from Puerto Rico are behaving similarly to what was found in a different Spanish-English bilingual community highlights the importance of incorporating corpus data in the study of bilingual language processing to better understand cross-community generalizations that may exist.

With this in mind, it is plausible that the larger pupillary response observed for both multi-word alternations and LOLIs preceded by masculine determiners is a reflection of the relative frequencies with which these different types of language mixing occur in spontaneous production and *not* of their morphosyntactic characteristics. For MWAs, a phrase like “*el ladrón robó el shiny ring*” may be less expected compared to, for example, “*el ladrón robó the shiny ring*”, with the determiner also in English. Similarly, the post-nominal Spanish adjective in the LOLI condition may likewise be unexpected, with other

noun/adjective combinations being preferred. Corpus data from spontaneous bilingual Puerto Rican speech is necessary to support or refute this hypothesis. In the present study, however, it is not possible to directly test these hypotheses because at least one additional condition is needed: namely, where the switch occurs *at* the determiner rather than *after* it, which may reflect the most common type of language mixing that occurs with noun/adjective phrases in Spanish-English bilingual speech. In addition, this condition is also more unambiguously English than the current MWA condition, where the English adjective-noun pair is still headed by a Spanish determiner, which may lend some degree of morphosyntactic integration. Future research will add this condition in order to more directly test how the relative frequencies of these types of language mixing affect online processing and if there are differences in the ‘degree’ of morphosyntactic integration that results from the use of an English, rather than Spanish, determiner.

Is Morphosyntactic Integration Enough?

To conclude this discussion, let us return to one of the primary questions of the present study: does the morphosyntactic integration of LOLIs provide some sort of processing benefit when compared to multi-word alternations? The findings discussed above do not seem to support this hypothesis, given that reliable differences did not arise between the LOLI and MWA conditions. In addition, when preceded by feminine Spanish determiners, only LOLIs elicited significantly larger pupillary responses when compared to the Unilingual condition. Taken together, the results of the present study suggest that morphosyntactic integration is not the defining feature of LOLIs, at least with respect to online processing. Rather, there instead may be a complex interaction between various production strategies—e.g., the ‘masculine default strategy’, the phonetic realization of the other-language material (Fricke et al. 2016), and the relative frequencies of switches that occur at rather than after the determiner in noun/adjective language mixing—that produce a graded rather than absolute difference between the different types of language mixing explored in the present study. At least for comprehension, it does not appear that MWAs incur any additional processing demands as a result of “switching both lexicons and grammars” compared to LOLIs, as has been argued (e.g., Poplack and Dion 2012, p. 309). Rather, the comprehension system may ultimately treat these two forms of language mixing similarly. This suggests that the same underlying mechanism(s) may be involved for both MWAs and LOLIs, in line with similar proposals put forward for production (e.g., Backus 2014; Grimstad 2017).

That there were no reliable differences between the LOLI and MWA conditions may also speak to the control strategies that were employed by the listeners. Green (2018; see also Green and Wei 2014) posits that bilingual speakers deploy different control modes depending on how they use their two languages. When speaking one language, for example, the other must be inhibited, resulting in a competitive control mode. When engaging in language mixing, however, bilinguals may employ a coupled or cooperative control mode, whereby control may be passed back-and-forth between the two languages as needed to let elements from one language into an utterance of the other. Single-word insertions and multi-word alternations may thus both rely on a coupled control mode, which would result in similar processing.

Regardless of the underlying mechanisms, the findings of the present study prove problematic for accounts that posit that LOLIs arise from distinct processes (e.g., nonce borrowing) compared to multi-word alternations. One important caveat, however, is that the present study examine comprehension, while nonce borrowing has previously been studied only in production. Although we, and others, argue that the two are intimately linked, it is nonetheless plausible that different strategies may emerge that guide the production and comprehension of certain types of language mixing. Future research should continue to use converging methodologies, particularly the combination of psycholinguistic lab-based studies of language comprehension and sociolinguistic studies of spontaneous production,

to further our understanding of language mixing and the link between comprehension and production more generally.

6. Concluding Remarks

The goal of the present study was to begin assessing the role of morphosyntactic integration on the processing of MWAs and LOLIs, and whether such integration led to a facilitative effect in the processing of LOLIs given its frequent use in spontaneous bilingual speech. The results thus far suggest that morphosyntactic integration may not play as large a role in processing as some scholars have assumed (e.g., Poplack and Dion 2012). Nonetheless, there are some limitations in the present study that, if addressed, may lead to a more thorough understanding of the processing of MWAs and LOLIs. First, although Poplack and colleagues do not consider phonological integration a key component of ‘nonce borrowings’, previous literature has demonstrated that the phonetic characteristics of bilingual speech play a large role in shaping how it is processed (e.g., Fricke et al. 2016). The present study did not control for how the stimuli were produced by the speaker; as such, it is possible that the phonetic realization of the MWAs and LOLIs in the present study may have had an effect on how they were processed. In this same vein, while the stimuli were checked by three Puerto Rican Spanish-English bilingual codeswitchers (in addition to the speaker), the stimuli were not normed for naturalness which may also have affected the results.

A further limitation lies in the use of feminine determiner-English noun language mixes. Although these were congruent in gender (i.e., the Spanish equivalent of the English noun was also feminine), this particular combination is rare in spontaneous bilingual speech (e.g., Valdés Kroff 2016; though see Cruz 2021) but have not been shown to differ from masculine determiner-English noun language mixes in terms of processing (Beatty-Martínez and Dussias 2019). Despite this, that such an infrequent form of language mixing accounted for fully half of the stimuli in the present study may have likewise had an effect on how they were processed (see Johns et al. 2019 for a discussion).

Future research should seek to not only to address these issues but also to pursue them as research questions in and of themselves. For example, is the processing of MWAs and LOLIs affected by whether they adopt the phonotactics of the lexifier language or of the recipient language? Likewise, how are English nouns whose Spanish equivalent is feminine processed when preceded by a masculine, rather than a feminine, Spanish determiner? Ultimately, such questions should be guided by findings from spontaneous bilingual speech given the influential role of usage and experience on the processing of linguistic input (MacDonald 2013). By bringing together both sociolinguistic insights on language production and psycholinguistic evidence on language processing, we may show that the nature of contact between the two languages of a bilingual community and the creative abilities of individual language users together shape the patterns and processes of language mixing.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/languages7040267/s1>, Table S1: Experimental Stimuli; Table S2: Model summary with reference level of Unilingual, Masculine; Table S3: Model summary with reference level of Unilingual, Feminine; Table S4: Model summary with reference level of LOLI, Masculine; Table S5: Model summary with reference level of MWA, Feminine; Audio S1: 131_Unilingual_pantallas.wav (Unilingual, Feminine); Audio S2: 156_Unilingual_anillo.wav (Unilingual, Masculine); Audio S3: 221_LOLI_earrings.wav (LOLI, Feminine); Audio S4: 246_LOLI_ring.wav (LOLI, Masculine); Audio S5: 311_MWCS_silver_earrings.wav (MWA, Feminine); Audio S6: 336_MWCS_shiny_ring.wav (MWA, Masculine).

Author Contributions: Conceptualization, M.A.J. and P.E.D.; methodology, M.A.J. and P.E.D.; software, M.A.J.; validation, M.A.J. and P.E.D.; formal analysis, M.A.J.; investigation, M.A.J.; resources, M.A.J. and P.E.D.; data curation, M.A.J.; writing—original draft preparation, M.A.J.; writing—review and editing, M.A.J. and P.E.D.; visualization, M.A.J.; supervision, P.E.D.; project administration, M.A.J.

and P.E.D.; funding acquisition, M.A.J. and P.E.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Science Foundation grant number DGE1255832 to Michael A. Johns and grant number BCS1823634 to Paola E. Dussias. The APC was funded by the National Science Foundation grant number DGE1255832 to Michael A. Johns.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki and approved by the Institutional Review Board of The Pennsylvania State University (IRB 34810, approved on 10 February 2019).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: All data and R code supporting the reported results can be found at <https://osf.io/2cbmd/> (accessed on 20 August 2018).

Conflicts of Interest: The authors declare no conflict of interest.

Notes

- ¹ Stimuli contained 7–10 words, with an average of 7.58 words and was identical across conditions save for two stimuli where the English translation of the target Spanish noun was a two-word compound (*semáforo*, ‘stop sign’, and *helado*, ‘ice cream’). In addition, 15 of the nouns were plural and the remaining 105 were singular, and 15 of the adjectives had opaque gender (e.g., *débil*, ‘weak’) and the remaining 105 had transparent gender (e.g., *plateado/a*, ‘silver’).
- ² While this amount of excluded data may seem high compared to behavioral methods, where more than 10% of data excluded would be rare, this is not the case for pupillometry data. For example, Guasch et al. (2017, p. 52) report 7.43% of trials excluded during cleaning and Schmidtke (2014, p. 5) reports approximately 14% missing data, representing the range possible when using pupillometry.

References

- Aaron, Jessi E. 2015. Lone English-origin nouns in Spanish: The precedence of community norms. *International Journal of Bilingualism* 19: 459–80. [CrossRef]
- Adalar, Nevin, and Sali Tagliamonte. 1998. Borrowed Nouns; Bilingual People: The Case of the “Londrali” in Northern Cyprus. *International Journal of Bilingualism* 2: 139–59. [CrossRef]
- Alnæs, Dag, Markus H. Sneve, Thomas Espeseth, Tor Endestad, Steven H. P. Van de Pavert, and Bruno Laeng. 2014. Pupil size signals mental effort deployed during multiple object tracking and predicts brain activity in the dorsal attention network and the locus coeruleus. *Journal of Vision* 14: 1–20. [CrossRef] [PubMed]
- Aston-Jones, Gary, and Jonathan D. Cohen. 2005. An integrative theory of locus coeruleus-norepinephrine function: Adaptive gain and optimal performance. *The Annual Review of Neuroscience* 28: 403–50. [CrossRef]
- Attar, Nada, Matthew Schneps, and Marc Pomplun. 2013. Pupil size as a measure of working memory load during a complex visual search task. *Journal of Vision* 13: 160. [CrossRef]
- Baayen, Harald R., Matteo Fasiolo, Simon N. Wood, and Yu-Ying Chuang. 2022. A note on the modeling of the effects of experimental time in psycholinguistic experiments. *The Mental Lexicon, Online-First Article*. Published electronically 4 April 2022. [CrossRef]
- Backus, Ad. 2014. A usage-based approach to borrowability. In *New Perspectives on Lexical Borrowing*. Edited by Eline Zenner and Gitte Kristiansen. Berlin: Mouton de Gruyter, pp. 19–39. [CrossRef]
- Beatty, Jackson, and Daniel Kahneman. 1966. Pupillary changes in two memory tasks. *Psychonomic Science* 5: 371–72. [CrossRef]
- Beatty-Martínez, Anne L., and Paola E. Dussias. 2017. Bilingual experience shapes language processing: Evidence from codeswitching. *Journal of Memory and Language* 95: 173–89. [CrossRef]
- Beatty-Martínez, Anne L., and Paola E. Dussias. 2019. Revisiting Masculine and Feminine Gender in Spanish: Linguistic, Psycholinguistic, and Neurolinguistics Evidence. *Frontiers in Psychology* 10: 751. [CrossRef]
- Beatty-Martínez, Anne L., Christian A. Navarro-Torres, Paola E. Dussias, M. Teresa Bajo, Rosa E. Guzzardo Tamargo, and Judith F. Kroll. 2020. Interactional context mediates the consequences of bilingualism for language and cognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 46: 1022–47. [CrossRef]
- Bhatt, Rakesh M., and Agnes Bolonyai. 2011. Code-switching and the optimal grammar of bilingual language use. *Bilingualism: Language and Cognition* 14: 522–46. [CrossRef]
- Blanco-Elorrieta, Esti, and Liina Pykkänen. 2017. Bilingual language switching in the lab vs. in the wild: The spatio-temporal dynamics of adaptive language control. *Journal of Neuroscience* 37: 9022–36. [CrossRef] [PubMed]
- Bobb, Susan C., and Zofia Wodniecka. 2013. Language switching in picture naming: What asymmetric switch costs (do not) tell us about inhibition in bilingual speech planning. *Journal of Cognitive Psychology* 25: 568–85. [CrossRef]
- Boersma, Paul, and David Weenink. 2022. Praat: Doing Phonetics by Computer [Computer Program]. V. 6.1.09. Available online: <http://www.praat.org/> (accessed on 20 August 2018).

- Bradshaw, John L. 1967. Pupil size as a measure of arousal during information processing. *Nature* 216: 515–16. [CrossRef] [PubMed]
- Budzhak-Jones, Svitlana. 1998. Against Word-internal Codeswitching: Evidence from Ukrainian-English Bilingualism. *International Journal of Bilingualism* 2: 161–82. [CrossRef]
- Bultena, Sybrine, Ton Dijkstra, and Janet G. Van Hell. 2015a. Language switch costs in sentence comprehension depend on language dominance: Evidence from self-paced reading. *Bilingualism: Language and Cognition* 18: 453–69. [CrossRef]
- Bultena, Sybrine, Ton Dijkstra, and Janet G. Van Hell. 2015b. Switch cost modulations in bilingual sentence processing: Evidence from shadowing. *Language, Cognition and Neuroscience* 30: 586–605. [CrossRef]
- Byers-Heinlein, Krista, Elizabeth Morin-Lessard, and Casey Lew-Williams. 2017. Bilingual infants control their languages as they listen. *Proceedings of the National Academy of Sciences* 114: 9032–37. [CrossRef]
- Costa, Albert, and Mikel Santesteban. 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. [CrossRef]
- Cruz, Abel. 2021. A syntactic approach to gender assignment in Spanish-English bilingual speech. *Glossa: A Journal of General Linguistics* 6: 1–40. [CrossRef]
- Davies, Mark. 2016. Corpus del Español: Web/Dialects. Available online: www.corpusdelespanol.org (accessed on 20 August 2018).
- Dell, Gary S., and Franklin Chang. 2014. The P-chain: Relating sentence production and its disorders to comprehension and acquisition. *Philosophical Transactions of the Royal Society B* 369: 20120394. [CrossRef] [PubMed]
- Deuchar, Margaret, Peredur Davies, Jon Russell Herring, M. Carmen Parafita Couto, and Diana Carter. 2014. Building bilingual corpora. In *Advances in the Study of Bilingualism*. Edited by Enlli Môn Thomas and Ineke Menne. Bristol: Blue Ridge Summit: Multilingual Matters, pp. 93–111. ISBN 978-1-78309-170-6.
- Dussias, Paola E. 1997. Switching at No Cost: Exploring Spanish-English Codeswitching Using the Response-Contingent Sentence Matching Task. Doctoral Dissertation, University of Arizona, Tucson, AZ, USA.
- Eze, Ejike. 1998. Lending Credence to a Borrowing Analysis: Lone English-Origin Incorporations in Igbo Discourse. *International Journal of Bilingualism* 2: 183–201. [CrossRef]
- Fasiolo, Matteo, Raphaël Nedellec, Yannig Goude, and Simon N. Wood. 2020. Scalable visualisation methods for modern Generalized Additive Models. *Journal of Computational and Graphical Statistics* 29: 78–86. [CrossRef]
- Foote, Stephen L., and John H. Morrison. 1987. Extrathalamic modulation of cortical function. *Annual Review of Neuroscience* 10: 67–95. [CrossRef]
- Fricke, Melinda, Judith F. Kroll, and Paola E. Dussias. 2016. Phonetic variation in bilingual speech: A lens for studying the production-comprehension link. *Journal of Memory and Language* 89: 110–37. [CrossRef]
- Gabay, Shai, Yoni Pertzov, and Avishai Henik. 2011. Orienting of attention, pupil size, and the norepinephrine system. *Attention, Perception, and Psychophysics* 73: 123–29. [CrossRef]
- Gagl, Benjamin, Stefan Hawelka, and Florian Hutzler. 2011. Systematic influence of gaze position on pupil size measurement: Analysis and correction. *Behavior Research Methods* 43: 1171–81. [CrossRef]
- Goldwater, Bram C. 1972. Psychological significance of pupillary movements. *Psychological Bulletin* 77: 340–55. [CrossRef]
- Gollan, Tamar H., and Victor S. Ferreira. 2009. Should I stay or should I switch? A cost-benefit analysis of voluntary language switching in young and aging bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 35: 640–65. [CrossRef]
- Gosselin, Leah, and Laura Sabourin. 2021. Lexical-semantic processing costs are not inherent to intra-sentential code-switching: The role of switching habits. *Neuropsychologia* 159: 107922. [CrossRef]
- Green, David W. 2018. Language Control and Code-Switching. *Languages* 3: 8. [CrossRef]
- Green, David W., and Jubin Abutalebi. 2013. Language control in bilinguals: The adaptive control hypothesis. *Journal of Cognitive Psychology* 25: 515–30. [CrossRef] [PubMed]
- Green, David W., and Li Wei. 2014. A control process model of code-switching. *Language, Cognition, and Neuroscience* 29: 499–511. [CrossRef]
- Grimstad, Maren B. 2017. The code-switching/borrowing debate: Evidence from English-origin verbs in American Norwegian. *Lingue e linguaggio, Rivista semestrale* 1: 3–34. [CrossRef]
- Guasch, Marc, Pilar Ferré, and Juan Haro. 2017. Pupil dilation is sensitive to the cognate status of words: Further evidence for non-selectivity in bilingual lexical access. *Bilingualism: Language and Cognition* 20: 49–54. [CrossRef]
- Gullifer, Jason W., Judith F. Kroll, and Paola E. Dussias. 2013. When language switching has no apparent cost: Lexical access in sentence context. *Frontiers in Psychology* 4: 278. [CrossRef]
- Guzmán, Gualberto, Joseph Ricard, Jacqueline Serigos, Barbara E. Bullock, and Almeida J. Toribio. 2017. Metrics for Modeling Code-Switching Across Corpora. *Proc. Interspeech* 2017: 67–71. [CrossRef]
- Guzzardo Tamargo, Rosa E., Jorge R. Valdés Kroff, and Paola E. Dussias. 2016. Examining the relationship between comprehension and production processes in code-switched language. *Journal of Memory and Language* 89: 138–61. [CrossRef]
- Han, Xuran, Wei Li, and Roberto Filippi. 2022. The effects of habitual code-switching in bilingual language production on cognitive control. *Bilingualism: Language and Cognition* 2022: 1–21. [CrossRef]
- Hernandez, Arturo E., Mirella Dapretto, John Mazziotta, and Susan Bookheimer. 2001. Language switching and language representation in Spanish-English bilinguals: An fMRI study. *NeuroImage* 14: 510–20. [CrossRef]

- Herring, Jon R., Margaret Deuchar, M. Carmen Parafita Couto, and Mónica M. Quintanilla. 2010. 'I saw the madre': Evaluating predictions about codeswitched determiner-noun sequences using Spanish-English and Welsh-English data. *International Journal of Bilingual Education and Bilingualism* 13: 553–73. [CrossRef]
- Hofweber, Julia, Theodoros Marinis, and Jeanine Treffers-Daller. 2016. Effects of dense code-switching on executive control. *Linguistic Approaches to Bilingualism* 6: 648–68. [CrossRef]
- Johns, Michael A., and Jonathan Steuck. 2021. Is codeswitching easy or difficult? Testing processing cost through the prosodic structure of bilingual speech. *Cognition* 211: 104634. [CrossRef]
- Johns, Michael A., and Paola E. Dussias. 2021. Of revistas and magazines: Lexical competition in the online processing of established loanwords. *Journal of Second Language Studies* 4: 375–411. [CrossRef]
- Johns, Michael A., Jorge R. Valdés Kroff, and Paola E. Dussias. 2019. Mixing things up: How blocking and mixing affect the processing of codemixed sentences. *International Journal of Bilingualism* 23: 584–611. [CrossRef]
- Keuleers, Emmanuel, and Marc Brysbaert. 2010. Wuggy: A multilingual pseudoword generator. *Behavior Research Methods* 42: 627–633. [CrossRef]
- Kheder, Souad, and Edith Kaan. 2021. Cognitive control in bilinguals: Proficiency and code-switching both matter. *Cognition* 209: 104575. [CrossRef]
- Kroll, Judith F., Andrea Takahesu Tabori, and Christian Navarro-Torres. 2021. Capturing the variation in language experience to understand language processing and learning. *Language, Interaction, and Acquisition* 12: 82–109. [CrossRef]
- Kuchinsky, Stefanie E., Jayne B. Ahlstrom, Kenneth I. Vaden Jr., Stephanie L. Cute, Larry E. Humes, Judy R. Dubno, and Mark A. Eckert. 2013. Pupil size varies with word listening and response selection difficulty in older adults with hearing loss. *Psychophysiology* 50: 23–34. [CrossRef]
- MacDonald, Maryellen C. 2013. How language production shapes language form and comprehension. *Frontiers in Psychology* 4: 226. [CrossRef]
- Makowski, Dominique. 2018. The Psycho Package: An Efficient and Publishing-Oriented Workflow for Psychological Science. *Journal of Open Source Software* 3: 470. Available online: <https://github.com/neuropsychology/psycho.R> (accessed on 1 May 2022). [CrossRef]
- Meuter, Renata F., and Alan Allport. 1999. Bilingual language switching in naming: Asymmetrical costs of language selection. *Journal of Memory and Language* 40: 25–40. [CrossRef]
- Muysken, Pieter. 2000. *Bilingual Speech: A Typology of Code-Mixing*. New York: Cambridge University Press.
- Muysken, Pieter. 2013. Language contact outcomes as the result of bilingual optimization strategies. *Bilingualism: Language and Cognition* 16: 709–30. [CrossRef]
- Myers-Scotton, Carol. 1993. *Dueling Languages*. New York: Oxford University Press.
- Poplack, Shana, and David Sankoff. 1984. Borrowing: The synchrony of integration. *Linguistics* 22: 99–135. [CrossRef]
- Poplack, Shana, and Natalie Dion. 2012. Myths and facts about loanword development. *Language Variation and Change* 24: 279–315. [CrossRef]
- Poplack, Shana, David Sankoff, and Christopher Miller. 1988. The social correlates and linguistic processes of lexical borrowing and assimilation. *Linguistics* 26: 4–104. [CrossRef]
- Poplack, Shana. 1993. Variation theory and language contact. In *American Dialect Research*. Edited by D. Preston. Amsterdam: Benjamins, pp. 251–86.
- Poplack, Shana. 2012. What does the Nonce Borrowing Hypothesis hypothesize? *Bilingualism: Language and Cognition* 15: 644–48. [CrossRef]
- Poplack, Shana. 2018. *Borrowing: Loanwords in the Speech Community and in the Grammar*. New York: Oxford University Press.
- R Core Team. 2022. R: A language and environment for statistical computing. In *R Foundation for Statistical Computing*. Vienna: R Core Team. Available online: <https://www.R-project.org/> (accessed on 1 May 2022).
- Samar, Reza G., and Marjory Meechan. 1998. The Null Theory of Code-Switching versus the Nonce Borrowing Hypothesis: Testing the Fit in Persian-English Bilingual Discourse. *International Journal of Bilingualism* 2: 203–19. [CrossRef]
- Samuels, Ebony R., and Elemer Szabadi. 2008. Functional neuroanatomy of the noradrenergic locus coeruleus: Its role in the regulation of arousal and autonomic function part I: Principles of functional organisation. *Current Neuropharmacology* 6: 235–53. [CrossRef]
- Sankoff, David, Shana Poplack, and Swathi Vanniarajan. 1990. The case of the nonce loan in Tamil. *Language Variation and Change* 2: 71–101. [CrossRef]
- Schmidtke, Jens. 2014. Second language experience modulates word retrieval effort in bilinguals: Evidence from pupillometry. *Frontiers in Psychology* 5: 137. [CrossRef] [PubMed]
- Schmidtke, Jens. 2018. Pupillometry in Linguistic Research: An introduction and review for second language researchers. *Studies in Second Language Acquisition* 40: 529–49. [CrossRef]
- Shin, Sarah J. 2002. Differentiating language contact phenomena: Evidence from Korean-English bilingual children. *Applied Psycholinguistics* 23: 337–60. [CrossRef]
- Sirois, Sylvain, and Julie Brisson. 2014. Pupillometry. *WIREs Cognitive Science* 5: 679–92. [CrossRef] [PubMed]
- Sóskuthy, Márton. 2021. Evaluating generalised additive mixed modelling strategies for dynamic speech analysis. *Journal of Phonetics* 84: 101017. [CrossRef]

- Stammers, Jonathan R., and Margaret Deuchar. 2012. Testing the nonce borrowing hypothesis: Counter-evidence from English-origin verbs in Welsh. *Bilingualism: Language and Cognition* 15: 630–43. [CrossRef]
- Tomić, Aleksandra, and Edith Kaan. 2022. Oscillatory brain responses to processing code-switches in the presence of others. *Brain and Language* 231: 105139. [CrossRef]
- Torres Cacoulllos, Rena, and Catherine E. Travis. 2018. *Bilingualism in the Community: Codeswitching and Grammars in Contact*. Cambridge: Cambridge University Press.
- Torres Cacoulllos, Rena, and Jessi E. Aaron. 2003. Bare English-origin nouns in Spanish: Rates, constraints, and discourse functions. *Language Variation and Change* 15: 289–328. [CrossRef]
- Torres Cacoulllos, Rena, and Jessica Vélez-Avilés. forthcoming. Mixing Adjectives: Bilingual Strategies for Variable Word Order. *Linguistic Approaches to Bilingualism*. article under review, last saved 24 June 2022.
- Traxler, Matthew J., Robin K. Morris, and Rachel E. Seely. 2002. Processing subject and object relative clauses: Evidence from eye movements. *Journal of Memory and Language* 47: 69–90. [CrossRef]
- Tsoukala, Chara, Mirjam Broersma, Antal van den Bosch, and Stefan L. Frank. 2021. Simulating Code-switching Using a Neural Network Model of Bilingual Sentence Production. *Computational Brain and Behavior* 4: 87–100. [CrossRef]
- Turpin, Danielle. 1998. 'Le français, c'est le last frontier': The Status of English-origin Nouns in Acadian French. *International Journal of Bilingualism* 2: 221–33. [CrossRef]
- Valdés Kroff, Jorge R., Paola E. Dussias, Chip Gerfen, Lauren Perrotti, and M. Teresa Bajo. 2017. Experience with code-switching modulates the use of grammatical gender during sentence processing. *Linguistic Approaches to Bilingualism* 7: 163–98. [CrossRef]
- Valdés Kroff, Jorge R. 2016. Mixed NPs in Spanish-English bilingual speech: Using a corpus-based approach to inform models of sentence processing. In *Spanish-English Codeswitching in the Caribbean and the US*. Edited by Rosa E. Guzzardo Tamargo, Catherine M. Mazak and M. Carmen Parafita Couto. Amsterdam and Philadelphia: John Benjamins, pp. 281–300. [CrossRef]
- van Rij, Jacolien, Martijn Wieling, Harald R. Baayen, and Hedderik van Rijn. 2020. itsadug: Interpreting Time Series and Autocorrelated Data Using GAMMs [R package]. V. 2.4.1. Comprehensive R Archive Network (CRAN). R. Available online: <https://cran.r-project.org/web/packages/itsadug/index.html> (accessed on 1 May 2022).
- van Rij, Jacolien, P. Hendriks, Hedderik van Rijn, Harald R. Baayen, and Simon N. Wood. 2019. Analyzing the Time Course of Pupillometric Data. *Trends in Hearing* 23: 1–22. [CrossRef] [PubMed]
- Wells, Justine B., Morten H. Christiansen, David S. Race, Daniel J. Acheson, and Maryellen C. MacDonald. 2009. Experience and sentence processing: Statistical learning and relative clause comprehension. *Cognitive Psychology* 58: 250–71. [CrossRef]
- Wieling, Martijn. 2018. Analyzing dynamic phonetic data using generalized additive mixed modeling: A tutorial focusing on articulatory differences between L1 and L2 speakers of English. *Journal of Phonetics* 70: 86–116. [CrossRef]
- Wood, Simon N. 2011. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 73: 3–36. [CrossRef]
- Wood, Simon N. 2017. *Generalized Additive Models: An Introduction with R*. Boca Raton: Chapman and Hall/CRC.