

Article

# Language Mixing in the Nominal Phrase: Implications of a Distributed Morphology Perspective <sup>†</sup>

Michèle Burkholder

Department of Linguistics, University of Ottawa, 70 Laurier Ave. E., Ottawa, ON K1N 6N5, Canada; mburk048@uottawa.ca

<sup>†</sup> Some of the ideas and data that are discussed here were presented at the 14th Generative Approaches to Second Language Acquisition (GASLA) Conference, Southampton, UK, 7–9 April 2017.

Received: 14 January 2018; Accepted: 9 April 2018; Published: 12 April 2018



**Abstract:** This paper investigates a pattern found in Spanish–English mixed language corpora whereby it is common to switch from a Spanish determiner to an English noun (e.g., *la house*, ‘the house’), but rare to switch from an English determiner to a Spanish noun (e.g., *the casa*, ‘the house’). Unlike previous theoretical accounts of this asymmetry, that which is proposed here follows assumptions of the Distributed Morphology (DM) framework, specifically those regarding the relationship between grammatical gender and nominal declension class in Spanish. Crucially, and again in contrast to previous accounts, it is demonstrated that this approach predicts no such asymmetry for French–English. This hypothesis is tested experimentally using an acceptability judgment task with self-paced reading, and as expected, no evidence is found for an asymmetry. This experiment is also used to test predictions regarding how English nominal roots in mixed nominal phrases are assigned grammatical gender, and the impact of language background factors such as age of acquisition. Evidence is found that bilinguals attempt to assign analogical gender if possible, but that late sequential bilinguals have a stronger preference for this option than do simultaneous bilinguals.

**Keywords:** Distributed Morphology; language mixing; nominal phrase; Spanish–English; French–English; analogical criterion; age of acquisition; language dominance; bilingualism

## 1. Introduction

Current mainstream linguistic theory has been built upon language data obtained from native speakers of a large sample of the world’s diverse languages. What is often overlooked, however, is that most people in the world speak more than one language, and that such speakers often combine elements from both of their languages within a single utterance, and sometimes even within a single word. This type of language mixing is spontaneous in that it is not possible to predict when a switch or insertion will occur; however, it is also highly constrained (Lipski 1978; Pfaff 1979; Poplack 1980; Woolford 1983).<sup>1</sup> Considerable research has consequently been dedicated to understanding the formal linguistic principles that underlie such language mixing (e.g., Poplack 1980, 1981; Sankoff and Poplack 1981; Joshi 1985; Di Sciullo et al. 1986; Mahootian 1993; Belazi et al. 1994; MacSwan 2000; Liceras et al. 2008).

<sup>1</sup> The term “language mixing” will be used to refer to a phenomenon termed “intra-sentential code-switching” by some (e.g., MacSwan 2005), and “nonce borrowing” by others (e.g., Poplack and Dion 2012).

Within the framework of the minimalist program (Chomsky 1993, 1995, 2005), a productive line of research has assumed that the principles that govern bilingual language mixing are the same as those that govern monolingual unmixed language; that there should be no specific constraints that apply only to language mixing (MacSwan 2000). For example, if linguistic competence involves choosing and combining lexical items via operations such as select and merge to form convergent syntactic derivations, then the simple difference between unilingual unmixed language and bilingual mixed language is that the latter operates over lexical items from two languages rather than one. Consequently, mixed-language data can provide a unique testing ground for linguistic theory and, in turn, it can offer important insights into the nature of the underlying linguistic system(s) a bilingual has acquired.

While investigations within this line of research have often taken a lexicalist view of morphosyntax (e.g., Cantone and Müller 2008; Van Gelderen and MacSwan 2008; Van Dulm 2009; Moro Quintanilla 2014), other researchers (e.g., González-Vilbazo and López 2011; Grimstad et al. 2014; Alexiadou et al. 2015; Riksem 2017) have also been exploring the predictions and explanatory power of non-lexicalist frameworks such as Distributed Morphology (DM; Halle and Marantz 1993; Marantz 1997; Embick and Noyer 2007). As such, the first goal of this paper is to focus on a particular language mixing pattern that has been reported in the literature, and to re-examine it within the theoretical framework of DM. Specifically, the object of investigation is a fairly robust asymmetry whereby switches between a Spanish determiner (D) and an English noun (N), such as (1a), are considered to be grammatical, but those between an English D and a Spanish N, as in (1b), are not (Jake et al. 2005; MacSwan 2005; Licerias et al. 2008, 2016).<sup>2</sup>

- |    |    |             |             |
|----|----|-------------|-------------|
| 1. | a. | <i>la</i>   | house       |
|    |    | the.DF.SG.F | house       |
|    | b. | * the       | <i>casa</i> |
|    |    | the.DF      | house       |

In previous minimalist accounts of this asymmetry, some have argued that it is a consequence of an underlying asymmetry in the morphosyntactic features of the two languages; specifically, that Spanish has grammatical gender features whereas English does not (e.g., MacSwan 2005; Licerias et al. 2008; Moro Quintanilla 2014). While the account that will be proposed here also attributes a central role to the gender feature asymmetry, it will be argued that it is the relationship between gender and nominal declension classes in Spanish that is the true source. Interestingly, the proposed theoretical account makes the prediction that not all language pairs with a gender feature asymmetry will exhibit a mixed determiner phrase (DP) asymmetry. It will be argued that this is the case for French–English.

It is important to note that this account will assume that the bilinguals in question have the same underlying linguistic system in their two languages as do monolingual native speakers of each of their respective languages. Indeed, this may be the case for individuals who acquired both of their languages from birth, and who are perfectly balanced with respect to proficiency and frequency of use.<sup>3</sup> An important issue with this assumption, however, is that most bilinguals do not fit these “ideal” characteristics. As such, the theoretical proposal will also be discussed with respect to the different predictions it makes for bilinguals who have acquired their two languages under different learning conditions (e.g., simultaneous versus late sequential bilinguals). The second goal of this paper is then to present an experiment investigating some of these predictions for French–English mixed

<sup>2</sup> The following abbreviations for features will be used throughout the text: plural = [PL]; singular = [SG]; definite = [DF]; feminine = [F] or [+FEM]; masculine = [M] or [-FEM]; neuter = [N].

<sup>3</sup> Although it is often argued that even such bilinguals are not necessarily “two monolinguals in one” (e.g., González-Vilbazo and López 2011), this can be an adequate initial hypothesis.

DPs. Experimental data, however, are not directly reflective of a bilingual's underlying linguistic competence, but may also reflect performance factors related to language processing. For bilinguals, the relative ease with which they process and access words from their two languages can depend on external factors such as language proficiency and dominance (i.e., whether they use one language preferentially). As such, the impact of these factors will also be taken into consideration when forming hypotheses.

The remainder of this paper will be organized as follows. Section 2 will provide the necessary background, including a more detailed look at the mixed DP asymmetry in Spanish–English (2.1), a brief overview of the DM framework (2.2), and the particular view on gender that will be adopted (2.3). This view will be demonstrated in the context of unmixed Spanish DPs. In Section 3, this framework will be extended to account for Spanish–English mixed DPs (3.1), and to make predictions for French–English mixed DPs (3.2). This proposal will then be discussed with respect to the predictions that such a model makes for bilinguals with a late age of acquisition (3.3), and with respect to how language proficiency and language dominance might influence the patterns observed in actual mixed-language data (3.4). Section 4 will then present the details of an experimental study designed to test the predictions made for different types of French–English bilinguals. Results will be presented in Section 5 and discussed in Section 6. Section 7 will conclude.

## 2. Background

### 2.1. The Mixed Determiner Phrase Asymmetry in Spanish–English

The claim that it is ungrammatical to switch between an English D and a Spanish N is based on the observation that such constituents are quite rare in the speech of Spanish–English bilinguals. For example, in Licerias et al.'s (2006) analysis of the Moyer's (1993) Spanish–English corpus, they found that only two of the 215 mixed DPs consisted of an English D and a Spanish N; and in the corpus reported on in Jake et al. (2002), there are no reported instances of such mixed DPs (Jake et al. 2005).<sup>4</sup> Further, Licerias et al. (2016) report a similar pattern found in the spontaneous productions of simultaneous bilingual children, based on corpus data from the FerFuLice corpus (Fernández Fuertes and Licerias 2009) available in CHILDES MacWhinney (2000).

MacSwan (2005) accounts for this asymmetry based on Moro's (2001) analysis, which takes a minimalist point of view with lexicalist assumptions. They propose that mixed DPs such as *the casa* crash because a Spanish N cannot successfully value and delete the uninterpretable features of an English D “in one fell swoop”. This is because the Spanish N has PERSON, NUMBER and GENDER features, but there is no corresponding GENDER feature on an English D. In contrast, mixed DPs such as *la house* converge because the English N's features, PERSON and NUMBER, are a subset of those of the Spanish D. These can therefore be valued and deleted, leading to convergence (MacSwan 2005).

Licerias et al. (2008, 2016), however, argue that this proposal is too categorical, and as such, it is not able to account for the fact that this asymmetry has not been replicated in acceptability judgment tasks. For example, they present data showing that adult L1-Spanish L2-English sequential bilinguals (Licerias et al. 2008; Licerias 2014), as well as both sequential and simultaneous bilingual children (Fernández Fuertes et al. 2011), evaluate mixed DPs with an English D to be as acceptable as those with a Spanish D, as long as the Spanish D matches in gender with the English N's translation equivalent.<sup>5</sup> In order to account for this data as well as the patterns found in the corpus data, they propose the grammatical features spell-out hypothesis. This hypothesis is based on the theoretical assumptions of Pesetsky and Torrego (2001) regarding feature checking, and those of Kihm (2005)

<sup>4</sup> Note that these researchers expect that mixed DPs with English Ds are possible and would be more frequent in a corpus where the speakers' dominant language is English, as there would be a greater tendency towards English providing the morphosyntactic frame (i.e., the “matrix language”), which they argue is the source of functional items such as determiners.

<sup>5</sup> Note that “L1” refers to an individual's native language, and “L2” to their second language.

regarding gender as a feature on the nominalizing head, *n*. Importantly, it provides a solution to the conflicting experimental and naturalistic data by suggesting that in processing, but not in production, the gender valuation mechanism between an N and a D can be cancelled in order to be more economical with processing resources.

What these two accounts have in common, however, is that they predict a similar mixed DP asymmetry for any language pair that has an underlying gender feature asymmetry. Indeed, this asymmetry has been reported for other language pairs, such as Welsh–English (Deuchar 2006) and German–English (Jorschick et al. 2011). On the other hand, there is evidence from a Norwegian–English mixed language corpus that switches in both directions are attested (Alexiadou et al. 2015; Riksem 2017). Further, there is little in the literature reporting on French–English mixed DPs, and it does not appear that the mixed DP asymmetry has been investigated here. As such, it is unclear if it exists for this language pair. The DM-based theoretical account that will be proposed aims to reconcile these conflicting observations. Crucially and in contrast with previous accounts, it does not predict such an asymmetry for French–English.

## 2.2. The Distributed Morphology Framework

From a traditional lexicalist perspective of minimalist syntax, the lexicon includes items that are specified for phonological form, word category information, and morphosyntactic features such as PERSON, NUMBER and GENDER (e.g., Lieber 1992; Chomsky 1995; Müller and Wechsler 2014). Morphologically complex words are derived in the lexicon and enter the derivation with their features. Syntactic structure is then projected from those features. This “lexicalist” approach has been extended to account for bilingual language mixing, where it has been proposed to operate in exactly the same way, the only difference being that items from either of a bilingual’s two lexicons can be selected into the numeration (MacSwan 1999, 2000). MacSwan (1999, p. 235) explicitly makes reference to such lexicalist assumptions when he defines borrowing from one language into another as “an operation whereby a new stem is introduced into a specific lexicon where morphologically complex items are formed before entering the numeration”.

In contrast, the DM framework (Halle and Marantz 1993; Marantz 1997; Embick and Noyer 2007), as a non-lexicalist approach, assumes not a single lexicon, but one that is divided up into three separate “lists”. Each of these lists is accessed at different points in the derivation. List 1 consists of bare lexical roots and abstract morphemes, which are the elements that can be selected to form a numeration. Roots (e.g.,  $\sqrt{\text{HOUSE}}$ ,  $\sqrt{\text{BOOK}}$ ) have no morphosyntactic features and they are not specified for word categories. It has also been argued that they are not specified for phonological forms (e.g., Siddiqi 2009), nor even for semantics (e.g., Harley 2014). Abstract morphemes, on the other hand, consist of morphosyntactic features (e.g., [SG], [DF],  $n_{[+FEM]}$ ). List 1 elements that are selected into the numeration combine with each other in the computational syntactic component to form morphologically complex words via the same operations as those which combine words to form syntactic structures (i.e., merge, agree). For example, bare roots can become nouns by combining with the nominalizing head, *n*.

Syntactic terminal nodes at spell-out therefore consist of roots fused together with the abstract morphemes that have merged above them (Siddiqi 2009), and still lack phonological and semantic information. This brings us to the vocabulary items (VIs) of List 2, which are the mappings of these root and feature bundles to their corresponding phonological forms. List 3 items are the mapping of terminal nodes to their semantic exponents.<sup>6</sup> Crucially, VIs are morphologically underspecified and compete for insertion into the fully specified syntactic terminal nodes, such that the exponent with the most compatible features (but no incompatible features) is selected. This selection process is known as

<sup>6</sup> This is somewhat of a simplification; however, because List 3 is not relevant to the topic of this paper, it will not be discussed further.

the subset principle (Halle 1997). For example, of the Spanish VIs that compete at a D terminal node with [M], [SG] and [DF] features, (2d) provides the best match and so is inserted.

2. List 2 Spanish VIs competing for D[M.SG.DF]
  - a. F.PL.D → las
  - b. PL.DF → los
  - c. F.DF → la
  - d. **M.DF** → **el**
  - e. DF → lo

### 2.3. A Distributed Morphology Account of Unmixed Spanish Determiner Phrases

Inanimate nouns in Spanish are arbitrarily assigned one of two grammatical genders, feminine and or masculine, and typically fall into one of three declension classes (I, II, III), which are each associated with a particular declension class suffix (*-o*, *-a*, or *-e*, respectively). However, while declension classes are strongly correlated with gender, such as in (3a) and (3b), they are indeed distinct from gender as there are exceptions, for example (3c). Also, while some nouns take no overt suffix, such as established loanwords, for example (3d), nouns that do take a declension class suffix never surface bare; they are bound roots.

3.
  - a. la cas-a  
the.F house.F-II
  - b. el libr-o  
the.M book.M-I
  - c. la man-o  
the.F hand.F-I
  - d. el béicon  
the.M bacon.M

Turning now to how unmixed Spanish DPs are derived from a DM perspective (see Figure 1a), it will be assumed that bare lexical roots must merge with a nominalizing head, *n*, in order to surface as a noun. In languages without grammatical gender, such as English, *n* is plain; however, in Spanish, *n* is bundled together with either an uninterpretable *u*[+FEM] or *u*[-FEM] feature (Kihm 2005; Kramer 2015).<sup>7</sup> Spanish roots are licensed under one or the other, and this is what ensures that nouns consistently surface with the correct gender, feminine or masculine, respectively.<sup>8</sup> Importantly, Kramer (2015) proposes that, for Spanish, there are also “theme nodes” which are adjoined to *n* after spell-out. This is done in accordance with theme node insertion rules, such as those in (4). The rules in (4a) capture all the listed exceptions, such as (3c) above, and those in (4b) and (4c) ensure that, aside from those exceptions, declension class is indeed dependent on the underlying gender feature.

<sup>7</sup> See Kramer (2015) and also Kučerová (2017) for a more complete view on the different types of gender that are involved, for example, in the derivation of animate nouns with biological gender.

<sup>8</sup> Nouns that surface with masculine gender could alternatively be licensed under a plain *n* as default, as assumed by Kramer (2015); however, it will be assumed here that there is an uninterpretable [-FEM] for inanimate nouns just as there is an interpretable [-FEM] feature for animate nouns with male biological gender.



5. List 2 vocabulary insertion rules for theme nodes
  - a. [THEME, III] → -e/√X- \_\_ Num
  - b. [THEME, II] → -a/√X- \_\_ Num
  - c. [THEME, I] → -o/√X- \_\_ Num
  - d. [THEME] → ∅

### 3. Proposed Distributed Morphology Account of Mixed Determiner Phrases

Extending the DM model to language mixing in general, we can make the following two assumptions. First, it will be assumed that bilinguals have functionally separate distributed lexicons, such that in unilingual language mode they can access the set of lists associated to one language with limited interference from the other.<sup>10</sup> Second, during language mixing, neither language is inhibited. As such, List 1 elements from either language can be selected into the numeration, and List 2 VIs from both languages can compete for insertion after spell-out.

By applying these assumptions to the above account for unmixed Spanish DPs, an analysis for Spanish–English mixed DPs can now be proposed. This will then be extended to make predictions for French–English. As mentioned in the introduction, this proposal will assume that the bilinguals in question have “perfect” competence in each of their two languages. It will also assume that their innate knowledge is reflected in mixed language data. As such, the discussion of the theoretical proposal will focus on the impact that age of acquisition (AoA) and factors related to language processing may have on the expected patterns.

#### 3.1. Analysis for Spanish–English

With respect to the specific structure involved in DPs, it will be assumed that English differs from Spanish in two relevant ways. First, because English has no grammatical gender system, it has only a plain *n*, and the definite determiner, *the*, has only a [DF] feature. Second, like established loanwords in Spanish, it will be assumed that English nominal roots are free, not bound.

##### 3.1.1. The Mixed Determiner Phrase Asymmetry

Looking first at how mixed DPs such as *la house* are derived, Figure 1b shows that when an English root merges with a Spanish *n*, this operates in exactly the same way as with unmixed Spanish DPs. The crucial difference occurs during vocabulary insertion. Because the theme node is not right adjacent to a bound root in this case, rule (5d) applies, and no theme vowel is inserted. Because D values its uninterpretable  $\phi$ -features (PERSON, NUMBER and GENDER) against a feature bundle containing a feminine gender feature, the Spanish determiner *la* (6c) wins the competition for insertion, resulting in *la house*. Note that an important question in this case is whether an English root will merge an *n* with the *u*[+FEM] feature or that with the *u*[-FEM] feature. This will be addressed below in Section 3.1.2.

We turn now to the case of *the casa*, where a Spanish root merges with an English *n*. Because *n* has no gender feature, theme node insertion rule (4d) applies and so nothing is inserted. As such, the bound root cannot get a theme vowel to complete its phonological form. In order to get a theme node inserted and thus a theme vowel, the Spanish root has to merge with a Spanish *n* with a gender feature; however, if that is the case, then the D terminal node will also have a feminine feature by virtue of agree, and so the feminine Spanish determiner *la* (6c) is the VI that

<sup>10</sup> It is assumed that language membership is a psychologically emergent property that allows for the functional separation of a bilingual’s two languages, such that activation (of one, the other, or both) is managed via cognitive control mechanisms. Importantly, the computational system does not see language membership as a formal feature.

wins the competition for insertion, not *the* (6f). As such, *the casa* is blocked via the subset principle.

6. List 2 VIs for Spanish (a–e) and English (f) competing for D[DF.SG.F]
  - a. F.PL.DF → las
  - b. PL.DF → los
  - c. **F.DF** → **la**
  - d. M.DF → el
  - e. DF → lo
  - f. DF → the

It is therefore not so much the case that mixed DPs consisting of an English D and a Spanish N violate any morphosyntactic requirements of the computational system. Rather, they are blocked by the availability of a more highly specified Spanish VI. Interestingly, such an account is not incompatible with the grammatical features spell-out hypothesis of [Liceras et al. \(2008\)](#), which also assigns a preference to the more grammaticalized functional element, but leaves room for the acceptability of the less grammaticalized option (the English D) in situations where it is more efficient with respect to processing resources.

### 3.1.2. The Analogical Criterion

As it is assumed that Spanish has two possible *ns*, ( $n_{u[-FEM]}/n_{u[+FEM]}$ ), there is an option with respect to which Spanish *n* an English root will be licensed under in mixed DPs. Indeed, there are three possible outcomes for root licensing. First, it could be entirely arbitrary; because such English roots are combined spontaneously with Spanish elements and not repeatedly, it is unlikely that licensing conditions exist for these roots in the same way that they do for Spanish roots, whose licensing conditions would have been established during lexical acquisition via cues in the input. This would predict that all the options presented in (7) would be equally acceptable. Second, there could be a preference for licensing all English roots under the  $n_{u[-FEM]}$  by default, leading to preferences for (7b) and (7d). Finally, there could be a preference for licensing English roots under the *n* of their translation equivalent, leading to preferences for (7a) and (7d).

7.
  - a. *la* house  
the.DF.F house (c.f. house = casa.F)
  - b. *el* house  
the.DF.M house
  - c. *la* book  
the.DF.F book (c.f. book = libro.M)
  - d. *el* book  
the.DF.M book

This third option is known as the “analogical criterion”, and [Liceras et al. \(2008\)](#) found strong evidence for this with the L1-Spanish L2-English sequential bilinguals that they tested using an acceptability judgment task. Data from the corpora, however, suggest that the analogical criterion may be more of a tendency than a rule. For example, in the Spanish–English corpus-based studies presented in Table 1, it is clear that an English N with a masculine translation equivalent almost categorically appeared with a masculine Spanish D; whereas an English N with a feminine translation equivalent had considerably more optionality in terms of the gender of the Spanish D with which it appeared.

This suggests that, in production, analogical gender is applied to all English Ns if it can be easily retrieved from memory. If it is not retrieved, then masculine gender is applied as default.<sup>11</sup>

**Table 1.** The analogical criterion in Spanish–English corpora.

Study	Gender of Translation = M		Gender of Translation = F	
	Matching D	Mismatching D	Matching D	Mismatching D
	e.g., <i>el book</i>	e.g., <i>la book</i>	e.g., <i>la house</i>	e.g., <i>el house</i>
Poplack et al. (1982)	97%	3%	78%	22%
Clegg and Waltermire (2009)	96%	4%	21%	79%
Jake et al. (2002)	97%	3%	19%	81%

D: determiner.

Experimental results from a study employing eye-tracking with the visual world paradigm support this pattern. Kroff et al. (2016) found that lexical access to English Ns with a feminine Spanish translation was facilitated by the presence of a feminine D, indicating that this was a useful cue during comprehension. This could be because feminine Ds rarely co-occur in the input with English Ns that have a masculine translation (i.e., only about 3–4%). On the other hand, they found that lexical access to English Ns that have a masculine translation was not facilitated by the presence of a masculine D, suggesting that the masculine D is not a useful cue. Again, this is perhaps because masculine Ds co-occur in the input with English Ns that have either a masculine translation (96–97%) or a feminine translation (22–81%), which appears to be due to the default masculine status.

### 3.2. Predictions for French–English

French and Spanish have very similar gender systems, with two grammatical genders, masculine and feminine, that are assigned arbitrarily to inanimate nouns. It will therefore be assumed that inanimate roots in French are also licensed under either  $n_{u[+FEM]}$  or  $n_{u[-FEM]}$ . Unlike Spanish, however, French has no declension class suffixes. As such, roots in French are typically free as opposed to bound. While French has both feminine and masculine singular definite determiners, as does Spanish, there are some differences with respect to the VIs that French–English bilinguals have available for insertion into D, as shown in (8). Importantly, the feature specification of *le* (8c) is the same as that of *the* (8d), because unlike in Spanish, there is no other gender-unspecified determiner (i.e., Spanish *lo*). These properties of French sometimes lead to different predictions with respect to the mixed DP asymmetry and the analogical criterion.

8. List 2 VIs for French (a–c) and English (d) competing for D[DF.SG.F]
  - a. PL.DF → les
  - b. F.DF → la
  - c. DF → le
  - d. DF → the

#### 3.2.1. The Mixed Determiner Phrase Asymmetry

Since French nouns do not have obligatory affixes that are dependent on a gender feature (c.f. declension class suffixes in Spanish), this means that there is nothing prohibiting French roots from combining with a genderless English *n*. As such, one of the main predicted differences between French–English and Spanish–English mixed DPs is that combinations of an English D and a French N,

<sup>11</sup> Note that, of the three groups, the bilinguals in Poplack et al.’s (1982) corpus appear to be much better at retrieving the analogical gender of English Ns, as they do so for 78% of those with a feminine translation. It can be assumed that they do so at a similar rate for those with a masculine translation, though this is obscured by the correspondence between analogical gender and the default gender.

as in (9a), are expected to be grammatical, as well as those involving a French D and an English N, as in (9b). In other words, no mixed DP asymmetry is expected.

In support of this prediction is data from a Norwegian–English mixed language corpus (Alexiadou et al. 2015; Riksem 2017), which suggest that the mixed DP asymmetry is not present in this language pair despite their having a similar underlying gender feature asymmetry; both DPs in (10) are attested (see (Burkholder et al. 2017) for more details). Importantly, Norwegian roots, like French roots and unlike Spanish roots, are free.<sup>12</sup>

- |     |    |            |               |                |
|-----|----|------------|---------------|----------------|
| 9.  | a. | the        | <i>maison</i> |                |
|     |    | the.DF     | house         |                |
|     | b. | <i>la</i>  | house         |                |
|     |    | the.DF.F   | house         |                |
| 10. | a. | <i>det</i> | <i>andre</i>  | <i>crew-et</i> |
|     |    | the.DF.N   | other         | crew-DF.SG.N   |
|     | b. | the        | <i>by</i>     |                |
|     |    | the.DF     | city          |                |

### 3.2.2. The Analogical Criterion

Because it is assumed that French and Spanish have *ns* with the same gender features ( $n_{u[-FEM]}/n_{u[+FEM]}$ ), this leads to the prediction that French–English, like Spanish–English, has an option with respect to which French *n* an English root will be licensed under in mixed DPs. There is indeed evidence of this from French–English corpus data (Poplack et al. 1982). After controlling for biological gender, it was determined that 78% of English Ns with masculine translation equivalents took analogical gender (e.g., 11d), compared to only 42% of those with feminine translation equivalents (e.g., 11a). While this pattern is very similar to what they found in the Spanish–English corpora (see Table 1), it is notable that the application of analogical gender for masculine Ns was not as categorical in this French–English corpus. The authors note that this increased variability could be due to a higher number of cases where it was unclear to them which analogical gender should be assigned, which may have been partly related to their own lesser familiarity with French compared to Spanish.

- |     |    |           |                               |  |
|-----|----|-----------|-------------------------------|--|
| 11. | a. | <i>la</i> | house                         |  |
|     |    | the.DF.F  | house (c.f. house = maison.F) |  |
|     | b. | <i>le</i> | house                         |  |
|     |    | the.DF.M  | house                         |  |
|     | c. | <i>la</i> | sun                           |  |
|     |    | the.DF.F  | sun (c.f. sun = soleil.M)     |  |
|     | d. | <i>le</i> | sun                           |  |
|     |    | the.DF.M  | sun                           |  |

<sup>12</sup> Note that even though Norwegian nouns can take definite suffixes that have been argued to mark declension class rather than gender (Lohndal and Westergaard 2016), these nouns can surface without any suffixes, such as in indefinite constructions. This is evidence that Norwegian nouns involve free roots. It should also be noted that overt definite determiners are only present under certain conditions, such as when there is an adjective (e.g., see Julien 2005).

There is also evidence that L1-French speakers may apply different strategies than L1-Spanish speakers with respect to the analogical criterion. While Licerias et al. (2008) found strong evidence for the analogical criterion with their L1-Spanish L2-English sequential bilinguals, they found that L1-French L2-Spanish bilinguals (with English as a second L1) tended to prefer applying masculine as default, at least in Spanish–English mixed DPs.

### 3.3. Impact of Age of Acquisition

The DM-based analysis suggests that there may be different mixed DP patterns for simultaneous versus late sequential bilingual acquisition. For bilinguals acquiring two L1s simultaneously, it could be expected that new Spanish/French roots encountered in nominal contexts would become licensed under the appropriate gender-bearing *n* (either  $n_{u[+FEM]}$  or  $n_{u[-FEM]}$ ), whereas new English roots would become licensed under the plain *n*. As such, in language mixing contexts, both gender-bearing *ns* and the plain *n* would be available to these bilinguals when selecting the numeration. This is what has been assumed in the above theoretical analyses. In contrast, if it is assumed that after a particular age, functional features not present in the L1 are no longer available to L2 learners (e.g., Hawkins and Chan 1997), then it is possible that late L2 learners may not have all three *ns* in their repertoire.

For late L1-Spanish/L1-French L2-English bilinguals, this may mean that they do not have a plain *n*. If so, this predicts that both English and French/Spanish roots can only merge with a gender-bearing *n*. This has several possible consequences.

First, in Spanish–English mixed DPs, when Spanish roots merge with a plain English *n*, this already creates issues due to declension classes. As such, not having an English *n* in the repertoire of abstract morphemes does not change anything for Spanish roots. On the other hand, this possibility would mean that English roots (e.g.,  $\sqrt{\text{HOUSE}}$ ) appearing in Spanish–English mixed language contexts would always have to merge with either  $n_{u[+FEM]}$  or  $n_{u[-FEM]}$ . This predicts that such English Ns should only appear with Spanish Ds. This is because *el* and *la* are both more specified than *the*, meaning that they would be preferentially chosen at vocabulary insertion if available, as per the subset principle.<sup>13</sup> There is actually good evidence for this prediction. In the Spanish–English corpora presented in Table 1, full English DPs (e.g., *the book*) are extremely rare compared to those consisting of a Spanish D and an English N (e.g., *el book*).

Second, for DPs with French roots appearing in mixed language contexts, not having a plain English *n* could mean that D will always have a gender feature obtained from agreeing with a gender-bearing French *n*. As such, mixed DPs with a feminine noun, such as *the maison*, should be blocked in favour of unmixed DPs, such as *la maison*, because the more specified French VI *la* (8b) will always win over the English *the* (8d). However, because *le* (8c) has the same specification as *the* (8d), optionality is expected for mixed DPs such as *the soleil* and unmixed DPs such as *le soleil* (11a). In other words, a partial mixed DP asymmetry would actually be expected for a group of late L1-French L2-English bilinguals, whereby only feminine French Ns cannot appear with English Ds.

Finally, for English roots appearing in French–English mixed language contexts, no preference would be expected for a French D to be inserted. Consequently, full English DPs would be considered acceptable, unlike in Spanish–English. This is because *the* is assumed to have the same feature specification as *le*, and because English roots can be merged with  $n_{u[-FEM]}$  by default regardless of the gender of its translation.

With respect to late L1-English L2-French/L2-Spanish bilinguals, they may not have gendered *ns* in their repertoire of functional features. These bilinguals would also not likely have any gender features specified on the VIs for determiners, and therefore Spanish *la* and *el*, as well as French *la* and

<sup>13</sup> Note that *el/la* would not be available in unilingual English mode, allowing for full English DPs in that case.

*le*, would all have the same feature specifications as English *the*.<sup>14</sup> This would predict that, in mixed language contexts, English roots would be acceptable with all three of the available determiners, and as such, there would be considerable optionality. This prediction is consistent with the acceptability judgment data for L1-English L2-Spanish bilinguals presented by Licerias et al. (2008), which indicated that English nouns were found to be equally acceptable with either Spanish determiner, regardless of the gender of the noun's Spanish translation equivalent.

For such late L1-English bilinguals, it would also be expected that they would not represent Spanish declension classes in the same way as L1-Spanish speakers, because they would not have the gendered *ns* on which declension class suffixes are dependent. As such, this also predicts that such bilinguals would not display mixed DP asymmetry. Indeed, the L1-English bilinguals in the Licerias et al. (2008) study show a strong preference for mixed DPs consisting of an English D and a Spanish N (e.g., *the casa*) over those consisting of a Spanish D and an English N (e.g., *la house*), and rated them considerably higher than the L1-Spanish group.

### 3.4. Effects of Language Dominance and Proficiency on Processing

Experimental data collected using methods other than acceptability judgment tasks have actually found that there may be greater language switch costs associated with mixed DPs such as *la house*, for example using a picture naming task (Fairchild and Van Hell 2017), self-paced reading (Litcofsky and Hell 2017), as well as event-related potentials combined with rapid serial visual presentation (idem.). In Fairchild and Van Hell's picture naming task, it was proposed that this cost was associated with difficulty retrieving the appropriate gender-marked Spanish D rather than difficulty retrieving the English N, as even though participants were L1-Spanish, they were currently dominant in English. They hypothesize that balanced or Spanish-dominant bilinguals might not exhibit the same processing asymmetry. However, in the experiments conducted by Litcofsky and Hell (2017), also with L1-Spanish L2-English bilinguals, they found that that switch costs were related to language dominance such that they were greater in the dominant-to-weak direction, regardless of the L1. In other words, it was Spanish-dominant bilinguals who had greater processing costs when reading mixed DPs such as *la house*. This is what was found in the Fairchild and Van Hell study for English-dominant bilinguals. These different results could reflect the fact that the former study tested production, while the latter tested reading.

Relative language proficiency may also impact the extent to which the analogical criterion is applied in language use. In order for an English noun to surface with the gender of its translation equivalent, the mental representation of that translation must be accessed from memory. The ease with which this is done may depend on the way in which translations are stored, which may in turn depend on the bilingual's language proficiency. For example, one of the most prominent models of the bilingual mental lexicon, Kroll and Stewart (1994) revised hierarchical model (RHM), suggests that L1 words are learned in direct association with their concepts, whereas L2 words (i.e., in sequential L2 acquisition) tend to be learned by directly associating them to their L1 translation. As such, L2 links to concepts are only indirect, mediated through the L1. As learners become more proficient in their L2, the direct links between L2 words and concepts become stronger. Assuming this type of psycholinguistic model, it is possible that unbalanced sequential bilinguals who have these strong lexical links between L2 and L1 words might have a stronger tendency to transfer the licensing conditions of the L1 (French or Spanish) root to that of the L2 (English) root than would balanced simultaneous bilinguals who do not have such strong associations between L2 and L1 lexical representations.

---

<sup>14</sup> That which would ensure that French/Spanish nouns typically appear with the appropriate gender-marked determiner is not licensing conditions under the appropriate *n*, but proceduralized lexical associations between nouns and the particular determiners with which they frequently co-occur in the input.

The following section presents a study designed to test some of the predictions and issues presented in the previous sections. Two primary research questions are addressed: first, whether there is evidence of a French–English mixed DP asymmetry; and second, how the analogical criterion is applied in mixed DPs with an English lexical root. In particular, the impact of language background factors such as AoA and language dominance will be investigated by comparing the results of simultaneous bilinguals to those of late sequential L1-French L2-English bilinguals. The specific hypotheses will be presented in Section 4.4.

#### 4. Methods and Materials

The experiment combines two methodologies: an acceptability judgment task (AJT) and self-paced reading (SPR). The SPR paradigm was chosen as it was expected to elicit more implicit grammaticality assessments of targeted structures. This paradigm is typically used to measure the processing load of words during reading comprehension, where reading times (RTs) are correlated with lexical properties such as frequency and word length (Haberlandt 1994). Importantly, the processing load can also be increased due to the presence of a morphosyntactic violation (e.g., subject–verb agreement error). Indeed, previous studies using this methodology have demonstrated that morphosyntactic violations lead to increased RTs at the point of the anomaly (e.g., De Vincenzi et al. 2003). This effect has also been correlated with the P600 event-related potential (ERP) component, which indexes morphosyntactic violations (Ditman et al. 2007). In the context of language mixing, it is expected that RTs in an SPR will be longer at points of illicit or less preferred switches between languages. Increased RTs, however, may also reflect increased processing effort due to factors other than the implicit detection of grammatical anomalies. For example, it could be more costly to switch to a less activated language (Litcofsky and Hell 2017). Having participants also judge the acceptability of each sentence allows for the results of two dependent measures to be directly compared, and for the results of this study to be more easily related to those of previous studies that employed an AJT.

##### 4.1. Participants

A total of 49 French–English bilinguals participated in this study. Of these, 10 were excluded for the following reasons: (1) not being a native speaker of just French, or French and English; (2) having early exposure to a language other than French or English; (3) having a high level of proficiency in a language other than French or English; and (4) not being a self-described language mixer.<sup>15</sup> This was determined by their responses to a detailed language background questionnaire (Sabourin et al. 2016) and a language-mixing questionnaire (adapted from (Byers-Heinlein 2013)). One additional participant was later excluded for rating sentences significantly lower than all other participants (see Section 5). The remaining 38 participants (30 female) were all between the ages of 18 and 35 (mean = 20.5, SD = 4.1). They were divided into groups based on their age of immersion (AoI) in English, which was operationalized at the age at which their daily exposure to English reached 20%. Table 2 shows the linguistic profiles of each of the three resulting groups: simultaneous bilinguals (AoI = 0), early L2 learners (AoI < 7), and late L2 learners (AoI > 7).

---

<sup>15</sup> Being a native speaker was defined as having acquired that language from a primary caregiver during infancy (i.e., between birth and the age of two) in an immersive environment (daily exposure > 20%).

**Table 2.** Language background of participants. Means are shown with standard deviations (SD).

	Simultaneous (n = 18)	Early L2 English (n = 7)	Late L2 English (n = 13)
<b>1. Age of Acquisition (AoA)</b>			
Age of first exposure to French	0 (0)	0 (0)	0 (0)
Age of Immersion (AoI) in French	0 (0)	0 (0)	0 (0)
Age of first exposure to English	0 (0)	1.3 (1.3)	6.9 (3.3)
Age of Immersion (AoI) in English	0 (0)	3.0 (1.4)	14.6 (6.2)
<b>2. Language exposure &lt; age 2</b>			
French	55% (22%)	87% (6%)	91% (11%)
English	47% (24%)	13% (5%)	7% (11%)
<b>3. Current self-reported language proficiency<sup>16</sup></b>			
French Total Proficiency/29	27.8 (1.7)	28.4 (1.1)	27.5 (4.0)
English Total Proficiency/29	28.0 (2.5)	27.7 (1.1)	19.3 (6.7)
French Reading Proficiency/6	5.9 (0.2)	5.9 (0.4)	5.7 (0.6)
English Reading Proficiency/6	5.8 (0.3)	6.0 (0.0)	4.5 (1.4)
<b>4. Frequency of Language Mixing<sup>17</sup></b>			
General mixing	2.5 (1.4)	1.9 (1.1)	2.6 (1.6)
English to French switches	4.1 (1.6)	3.1 (2.2)	3.8 (1.5)
French to English switches	3.3 (1.3)	2.1 (1.1)	3.1 (1.6)
English borrowing	2.4 (1.2)	2.1 (1.1)	2.0 (1.3)
French borrowing	2.6 (1.5)	1.6 (0.5)	3.3 (1.9)

## 4.2. Stimuli

### 4.2.1. Critical Stimuli

Forty-eight nouns were selected to form part of the critical DPs. All were mono-morphemic and commonly known non-cognates, and approximately half had feminine gender in French (23/48). Nouns beginning with vowels were generally excluded in order to avoid élision between the definite determiner and the noun, which neutralizes the overtly marked gender of the determiner. No established loan words (such as the English words *job*, *fun*, which are commonly used in French) were included. The properties of these critical nouns in both languages are compared in Table 3.

A two-way repeated measures analysis of variance (ANOVA) was conducted for both word frequency and word length, using Language (English vs. French) as the within-item independent variable, and Gender (masculine vs. feminine) as the between-item independent variable. For word frequency, there were no significant main effects or interaction, indicating that word frequency was matched across language and gender. In terms of word length, neither Gender nor the interaction between Language and Gender were determined to be significant; however, the main effect of Language was determined to be significant ( $F(1,46) = 10.497$ ,  $MSE = 10.109$ ,  $p = 0.002$ ), indicating that French nouns were significantly longer on average than English nouns overall. Little impact, however, is expected, as the magnitude of the difference is quite small (0.7 graphemes), and the words in French and English are all relatively short (both have a maximum of seven graphemes).

<sup>17</sup> On a scale of 1–7, where 1 means ‘very true’ and 7 means ‘not true at all’ with respect to statements such as “In general, I often mix English and French” and “I often start a sentence in English then switch to French”.

<sup>16</sup> Measured with respect to five skill areas—reading, writing, listening, speaking, and pronunciation—with the first four given on a scale of 1–6 (very low to native) and the last on a scale of 1–5 (very accented to native).

**Table 3.** Lexical properties of critical nouns.

		Mean	SD	Min.	Max.
<b>Word Frequency</b> <sup>18</sup> (words per million)	<b>French</b>	72.0	123.0	0.7	570.3
	<b>English</b>	65.7	87.8	2.2	514.0
<b>Word Length</b> (graphemes)	<b>French</b>	5.4	1.1	4	7
	<b>English</b>	4.7	1.0	3	7

To create the critical DPs, nearly all of these nouns were paired with a singular definite determiner (43/48). The remaining five appeared with a singular indefinite determiner, either due to a better fit with the sentence context, or to avoid élision, as mentioned above. No differences were expected between the results for definite and indefinite determiners, as their features are comparable in both languages. One sentence was created for each of the 48 critical DPs, and three dependent variables were manipulated in order to produce six conditions for each sentence (see Table 4 for an example and the Supplementary Materials for a complete list). First, the language of the determiner in the critical DP was either French or English. Second, the language of the noun was either the same as the determiner (unmixed DP), or the opposite (mixed DP). Finally, for conditions with French determiners, the gender of the determiner either matched or mismatched the gender of the noun or its translation equivalent. Sentences were also coded for the gender of the noun or its translation equivalent.

The critical DP was always the third DP of the sentence, where the first was a language neutral proper name, and the second was a fully switched DP. This was done to establish the language mixing of the sentence before the critical DP was encountered, potentially reducing the processing cost of the critical language switch by priming it with a previous language switch. Also, the words preceding the critical determiner were always in the same language as that determiner, and the words following the critical noun were always in the same language as that noun. This was done in order to ensure that the language switch between the determiner and the noun was more along the lines of a multi-word “code-switch” as opposed to a single word “nonce borrowing” or insertion. Finally, the critical noun was never the last constituent in the sentence in order to avoid any sentence wrap-up effects.

**Table 4.** Conditions for critical stimuli (critical Determiner Phrase (DP) bolded). \*

D	DP Type	Congruence	Example
FR	Mixed	Matched Mismatched	Paul bought trois livres pendant <b>le month</b> of May. Paul bought trois livres pendant <b>la month</b> of May.
	Unmixed	Matched Mismatched	Paul bought trois livres pendant <b>le mois</b> de mai. Paul bought trois livres pendant <b>la mois</b> de mai.
EN	Mixed	N/A	Paul a acheté three books during <b>the mois</b> de mai.
	Unmixed	N/A	Paul a acheté three books during <b>the month</b> of May.

\* All sentences mean ‘Paul bought three books during the month of May’.

#### 4.2.2. Filler Items

Three types of filler sentences were included as stimuli (see Table S1). First, 15 sentences were created such that they contained a morphosyntactic mismatch between constituents belonging to the same language, such as number and verb tense violations. A control condition was created for these sentences by repairing the mismatch. This type of filler was included in order to determine whether both measures of the experimental task, ratings and RTs, were able to elicit the expected distinctions between grammatical and ungrammatical sentences. Similarly, another 15 sentences were created

<sup>18</sup> French and English word frequencies were determined using the Lexique 2 database (New et al. 2004) and the SUBTlexUS database (Brysbaert and New 2009), respectively. As they are very similar in nature—both based on subtitle corpora—it is possible to compare frequencies across languages.

such that they contained a type of language switch that is commonly considered in the literature to be ungrammatical, such as between an auxiliary and a main verb, between a pronoun and a verb, or between a noun and a post-posed adjective (which produces a word-order conflict). Again, a control condition was created for these sentences by repairing the mismatch. This type of filler was included in order to determine whether language switch violations were treated in the same way as the within-language morphosyntactic violations with respect to both dependent measures. Finally, 10 mixed-language sentences were created that were completely grammatical, with language switches occurring at uncontroversial sites, such as between a noun and a verb, or between a noun and a prepositional phrase.

#### 4.3. Procedure

All participants were recruited through the University of Ottawa's Integrated System of Participation in Research (ISPR), and were awarded one point towards their course grade. All subjects gave their informed consent for inclusion before they participated in the study. The study was conducted in accordance with the Declaration of Helsinki, and the protocol was approved by the Social Sciences and Humanities Research Ethics Board of the University of Ottawa (REB# 02-16-26). The entire testing session lasted approximately 60 min for each participant. Participants were first asked to complete the consent form, language background questionnaire and language-mixing questionnaire. They were then seated in a sound-attenuated room in front of a computer screen for the experimental task.

Participants were instructed that they would be reading mixed-language sentences one word at a time on the computer screen. They were told that the words of each sentence would be first covered in dashes, and to see each successive word, they were to press the space bar (as per the "moving window paradigm" of SPR). After each sentence, they would be asked to rate how acceptable it was to them on a scale from 1 to 4 by pressing the corresponding number key on the keyboard. They were instructed that a rating of "1" meant that the sentence "sounded terrible, it was not acceptable"; a rating of "2" meant that it "sounded pretty bad"; a rating of "3" meant that it "sounded ok, but not very natural"; and a rating of "4" meant that it "sounded good, it was perfectly acceptable". This 4-point scale was visible on the computer screen each time they were asked to make a rating, with the "1" marked as "Terrible" and the "4" marked as "Fine".

Each participant saw a total of 88 experimental sentences, presented using the Linger software program (version 2.94, <http://tedlab.mit.edu/~dr/Linger>), plus eight practice sentences. All words were displayed in black 20-point font on a light grey background. Spaces between words were not covered in dashes, and a period marked the end of each sentence. They saw each critical ( $n = 48$ ) and filler ( $n = 40$ ) sentence once, counterbalanced across conditions such that they saw an equal number of sentences in each condition. Sentences were presented in a semi-randomized order, and separated into six blocks such that each block began with a filler sentence and no more than two sentences of the same condition appeared consecutively. Participants were given a self-regulated break between each block. One third of both critical and filler sentences were followed by a yes/no question that was entirely in English.

#### 4.4. Hypotheses

The first set of hypotheses that were tested pertained to the mixed DP asymmetry. Based on the arguments presented in Section 3.2.1, it was expected that there should be no directional asymmetry for language switches between English and French determiners and nouns. This was expected to be the case at least for the simultaneous bilinguals. In contrast, and as discussed in Section 3.3, late L1-French L2-English bilinguals were expected to possibly show a partially-mixed DP asymmetry; DPs such as *the soleil<sub>M</sub>* should be acceptable, but not those such as *the maison<sub>F</sub>*. Finally, this group of late bilinguals may also show reduced processing costs for mixed DPs consisting of an English D and

a French N, because this language switch goes in the direction of the weak to the dominant language (see Section 3.4).

The second set of hypotheses tested predictions regarding gender assignment for English roots in mixed DPs, and the extent to which the analogical criterion applied. Based on the arguments presented in Section 3.2.2, it was expected that there would be considerable variability for English roots whose translation equivalents are licensed under  $n_{u[+FEM]}$  (i.e., sentences containing *le house* and *la house* should have similar acceptability ratings and similar RTs for the critical nouns), but no variability for roots whose translation equivalents are licensed under  $n_{u[-FEM]}$  (i.e., only *le soleil* should be acceptable, not *la soleil*).

It was further expected that the late bilinguals would have a greater tendency to apply analogical gender to English nouns whereas the simultaneous bilinguals were expected to possibly show a greater tendency to apply the masculine gender as default. This is because the direct lexical connections between translations in the L2-to-L1 direction translation equivalents were expected to be stronger for the late bilinguals, as per the RHM, and thus easier to activate (see Section 3.4).

## 5. Results

The accuracy of participants' responses to the comprehension questions was first calculated in order to determine whether they were paying sufficient attention to the reading task. It was determined that all participants responded with high accuracy to the comprehension questions, with a mean of 95.9% (SD = 4.8%), and a range of 77.4–100%.

Acceptability ratings were collected for each sentence, and the mean and standard deviation were calculated for each participant across all conditions in order to identify any outliers. One participant's mean rating was indeed more than 2.5 standard deviations from the group's overall mean, and so was excluded from the data analysis.<sup>19</sup> For the 38 remaining participants, mean ratings were calculated for each participant by condition, and this was used as the first dependent variable.

RTs were time-locked to participants' button press on the critical word in each sentence, that is, the word where the language switch or morphosyntactic violation became apparent (or not, for their respective control conditions). Mean RTs were calculated for each participant, and data points that were 2.5 standard deviations above (4.9% of the data) or below (none) each participant's mean were replaced by their mean value. Mean RTs were then re-calculated based on this trimmed data for each participant across conditions in order to identify outliers; one participant's mean was greater than 2.5 standard deviations from the group mean (a simultaneous bilingual), and so this participant's RT data was excluded from the statistical analyses; however, their acceptability judgment data were still analyzed. For each of the remaining 37 participants, mean RTs were calculated by condition, and this was used as the second dependent variable.

### 5.1. Grammaticality Violations in Filler Items

The two types of filler items involving within-language morphosyntactic violations and uncontroversial language switch violations were analyzed using a three-way repeated measures ANOVA for each dependent variable (ratings and RTs). Filler Type (language switch violation vs. morphosyntactic violation), Grammaticality (violation vs. control), and Group (simultaneous vs. early vs. late) were the independent variables.

Results for the ratings revealed a significant main effect of Grammaticality ( $F(1,35) = 188.437$ ,  $MSE = 32.352$ ,  $p < 0.001$ ), indicating that both types of violations received significantly lower ratings than their grammatical counterparts. There was also a significant main effect of Filler Type

<sup>19</sup> This participant was the oldest of the group (age at testing = 35), and was also one of the less frequent language mixers (her reported overall tendency to mix language was rated as 5/7, where 7 is "never"). Her low ratings in the experiment (mean = 1.31/4, SD = 0.63) suggest that she did not find the mixed language sentences natural, something that was not typical of the rest of the group.

( $F(1,35) = 6.548$ ,  $MSE = 1.480$ ,  $p = 0.015$ ), which was likely driven by the significant interaction between Filler Type and Grammaticality ( $F(1,35) = 7.772$ ,  $MSE = 1.585$ ,  $p = 0.009$ ). This interaction indicates that, while both types of violations were rated significantly lower than their grammatical counterparts, as confirmed by planned paired-sample t-tests ( $t(37) = 11.963$ ,  $p < 0.001$ ; and  $t(37) = 7.969$ ,  $p < 0.001$ , respectively), the magnitude of this difference was significantly larger for the within-language morphosyntactic violations (mean difference = 1.25 compared to 0.73). This likely reflects the fact that the participants have more metalinguistic knowledge of grammatical errors than they do of language switch violations. The main effect of Group did not reach significance ( $p = 0.897$ ), nor was its three-way interaction with Filler Type and Grammaticality, suggesting that all groups showed similar effects.

Results for RTs mirrored those of the ratings. There was a significant main effect of Grammaticality ( $F(1,34) = 22.750$ ,  $MSE = 527,233.974$ ,  $p < 0.001$ ), indicating that both language switch violations and morphosyntactic violations were read significantly more slowly than their grammatical counterparts. The main effect of *filler type* was only marginally significant ( $p = 0.067$ ), as was the interaction between Filler Type and Grammaticality ( $p = 0.093$ ). This interaction suggests that, while both within-language morphosyntactic violations and language switching violations were read significantly more slowly than their grammatical counterparts, as confirmed by planned paired-sample t-tests ( $t(36) = 5.590$ ,  $p < 0.001$ ; and  $t(36) = 3.685$ ,  $p = 0.001$ , respectively), the magnitude of this difference was marginally larger for the within-language morphosyntactic violations. There was also a significant main effect of Group ( $F(2,34) = 5.550$ ,  $MSE = 787,180.376$ ,  $p = 0.008$ ), indicating that the simultaneous bilinguals were overall slower than the early and late groups ( $p = 0.027$  and  $p = 0.032$ , respectively).<sup>20</sup> There was, however, no three-way interaction between Group, Filler Type, and Grammaticality ( $p = 0.596$ ), suggesting again that all groups showed similar effects of grammaticality on both types of filler sentence.

## 5.2. The Mixed Determiner Phrase Asymmetry

A three-way repeated measures ANOVA was conducted for each dependent variable (ratings and RTs) with Determiner Language (French<sup>21</sup> vs. English), DP Type (unmixed vs. mixed) and Group (simultaneous vs. early vs. late) as independent variables. Results for ratings (Figure 2a) revealed a significant main effect of DP Type ( $F(1,35) = 62.703$ ,  $MSE = 17.851$ ,  $p < 0.001$ ), indicating that participants preferred unmixed over mixed DPs overall (mean difference = 0.69).<sup>22</sup> The main effect of Determiner Language was not significant ( $p = 0.343$ ), and crucially, neither was the interaction between Determiner Language and DP Type ( $p = 0.457$ ), indicating that there was indeed no preference for mixed DPs with French determiners over English determiners. The main effect of Group was not determined to be significant ( $p = 0.897$ ), nor did it significantly interact with any other variable, suggesting that all groups shared the same pattern.

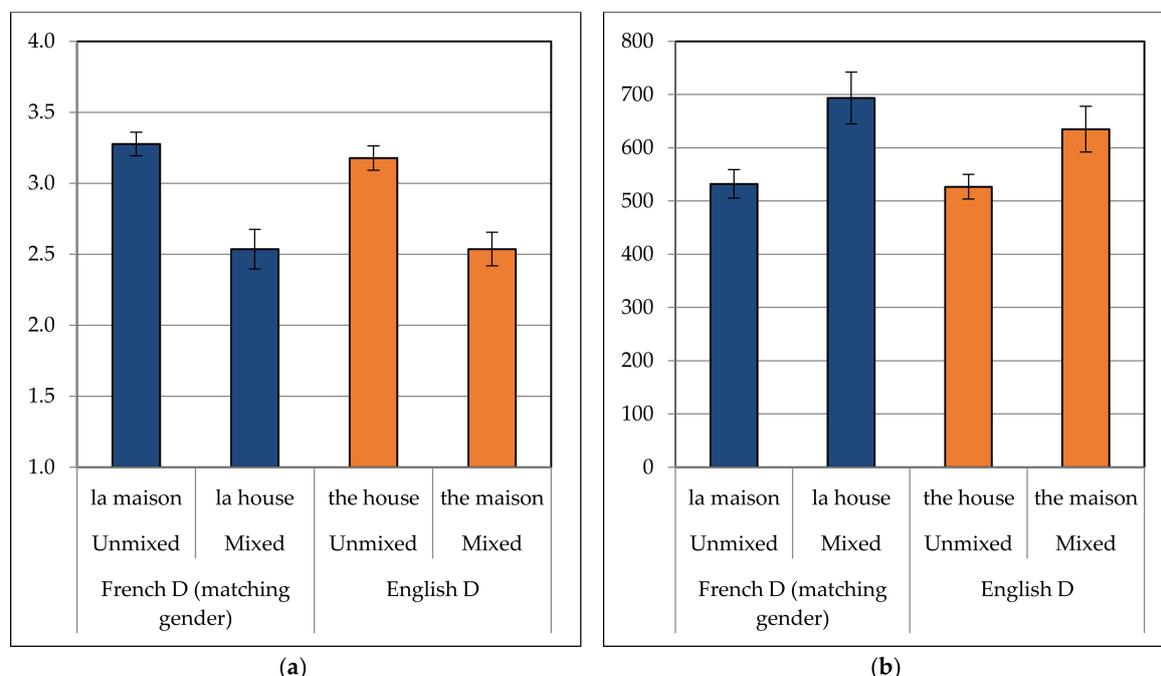
The patterns for ratings are mirrored in Figure 2b, which shows the mean RTs for critical nouns in the same four conditions. Indeed, statistical analyses confirm these patterns. There was a significant main effect of DP Type ( $F(1,34) = 22.799$ ,  $MSE = 486,233.390$ ,  $p < 0.001$ ), indicating that participants preferred unmixed over mixed DPs (mean difference = 135 ms). The main effect of Determiner Language was not significant ( $p = 0.519$ ), and again, neither was the interaction between Determiner Language and DP Type ( $p = 0.318$ ). Interestingly, the numerical trend was in the opposite direction than would be expected for the Spanish–English mixed DP asymmetry, with the English nouns following French determiners (e.g., *la house*) being read more slowly than French nouns following English

<sup>20</sup> A similar effect was also found in another study conducted with the same population (Sabourin et al. 2014). The authors attributed this effect to the idea that simultaneous bilinguals have a larger, more intimately integrated bilingual mental lexicon from which to activate and retrieve words during processing.

<sup>21</sup> These comparisons were only made with the gender congruent conditions with French determiners. This decision was based on previous results (e.g., Licerias et al. 2008) that indicated that this is the likely overall preference. Results in Section 5.3 support this decision.

<sup>22</sup> This preference could be a consequence of the fact that sentences with mixed DPs had two language switch points, compared to sentences with unmixed DPs, which only had one language switch point.

determiners (e.g., *the maison*; mean difference = 59 ms). There was a significant main effect of Group ( $F(2,34) = 3.694$ ,  $MSE = 422,112.890$ ,  $p = 0.035$ ) suggesting that the simultaneous bilinguals were again the slowest overall. The factor Group did not interact with any other factor, again suggesting that the patterns were similar across all three participant groups. Because, however, the hypotheses were made a priori regarding differences between the simultaneous and late bilingual groups, these were further investigated in order to identify any potential patterns or tendencies.



**Figure 2.** Results for the mixed determiner (D) phrase asymmetry: (a) Mean global acceptability ratings ( $N = 38$ ); (b) Mean reading times of the critical noun in milliseconds ( $N = 37$ ).

### 5.2.1. Between-Group Differences: *The Maison* vs. *The Soleil*

It was hypothesized that there might be a partial mixed DP asymmetry for the late bilinguals, whereby they might have only a dispreference for mixed DPs consisting of an English D paired with a feminine French N (e.g., *the maison<sub>F</sub>*) and not one for those paired with a masculine French N (e.g., *the soleil<sub>M</sub>*). In order to further investigate this possibility, a closer look was taken at the conditions including an English D. The ratings data are shown in Table 5.<sup>23</sup> The raw means alone suggest that both groups seem to have a slight preference for mixed DPs with a masculine N over those with a feminine N, and that the preference is greater for the late bilinguals (0.19 compared to 0.07). This pattern is consistent with the hypothesis, but of course the magnitudes of the differences are very small (though not nearly as small as for their unmixed counterparts, where the differences are very near null).

<sup>23</sup> Note that inspection of the RT data in the same conditions revealed no interpretable patterns whatsoever, and so this was not explored any further.

**Table 5.** Mean ratings for determiner phrases (DPs) with an English determiner.

	Unmixed DP			Mixed DP		
	Translation = F	Translation = M	Difference	Noun = F	Noun = M	Difference
	"the house"	"the sun"		"the maison"	"the soleil"	
SimBil	3.19	3.18	0.01	2.65	2.72	0.07
LateBil	3.21	3.21	0.00	2.33	2.52	0.19

An exploratory two-way ANOVA was conducted just for the late bilinguals, with DP Type (unmixed vs. mixed) and Noun Gender as the independent variables. While there was a marginally significant effect of DP Type ( $F(1,12) = 4.437$ ,  $MSE = 165,983.554$ ,  $p = 0.057$ ), neither the effect of Noun Gender nor the interaction were significant ( $p = 0.674$  and  $p = 0.645$ , respectively). Exploratory planned pair-wise comparisons confirmed that the difference in ratings between mixed DPs with English D did not differ significantly depending on the gender of the French N ( $p = 0.840$ ). As such, there is no evidence here that late bilinguals demonstrate any type of mixed DP asymmetry, despite the faint indication of the expected pattern in Table 5.

### 5.2.2. Between-Group Differences: Processing Costs

It was also hypothesized that the group of late bilinguals (but not the simultaneous bilinguals) would show a greater processing cost for mixed DPs where the switch direction was from the stronger (French) to the weaker (English) language (e.g., *la house*) compared to the opposite direction (e.g., *the maison*). Table 6 presents the relevant data.

**Table 6.** Language switch costs: Reading times of the critical noun by condition.

	French Determiner			English Determiner		
	Unmixed	Mixed	Difference	Unmixed	Mixed	Difference
	"la maison/ "le soleil"	"la house"/ "le sun"		"the house"/ "the sun"	"the maison"/ "the soleil"	
SimBil	624.4	823.1	198.7	576.0	686.5	110.5
LateBil	448.8	611.3	162.5	491.3	616.3	125.0

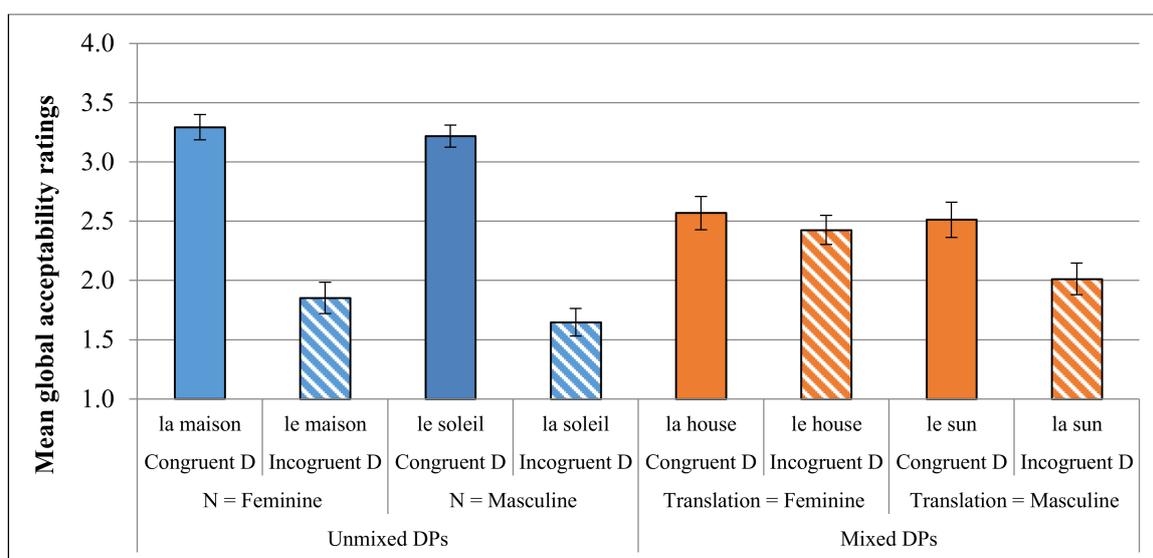
Looking at the patterns alone, it appears that both groups have a greater language switch cost in the French-to-English direction compared to the English-to-French direction. An exploratory two-way ANOVA was conducted for each of the two participant groups, with DP Type (unmixed vs. mixed) and Determiner Language as the independent variables. While there was a significant effect of DP Type for both the late and simultaneous groups ( $F(1,12) = 10.073$ ,  $MSE = 268,640.337$ ,  $p = 0.008$ ; and  $F(1,16) = 19.891$ ,  $MSE = 406,207.887$ ,  $p < 0.001$ , respectively), this factor did not significantly interact with Determiner Language for either group ( $p = 0.439$  and  $p = 0.227$ , respectively). As such, there is no robust evidence of greater processing costs for either group, and contrary to the hypothesis, both groups shared the same pattern. The reason for this similar, but a non-significant pattern will become clearer in the analysis investigating the analogical criterion, presented in the next section.

### 5.3. The Analogical Criterion in Mixed Determiner Phrases with English Nouns

Looking only at experimental conditions with critical DPs containing a French D, a four-way repeated measures ANOVA was conducted for each dependent variable, with DP Type (unmixed vs. mixed), Gender Congruence (congruent vs. incongruent), Noun Gender (feminine vs. masculine), and Group (simultaneous vs. early vs. late) as independent variables.

Results for ratings (Figure 3) showed a significant main effect of Gender Congruence ( $F(1,35) = 95.228$ ,  $MSE = 51.487$ ,  $p < 0.001$ ) as well as of Gender ( $F(1,35) = 14.794$ ,  $MSE = 2.230$ ,  $p < 0.001$ ), but not of DP Type ( $p = 0.150$ ). The interaction between DP Type and Gender Congruence was

significant ( $F(1,35) = 65.262$ ,  $MSE = 23.218$ ,  $p < 0.001$ ), indicating that the difference between DPs with congruent vs. incongruent gender was much larger in unmixed French DPs compared to mixed DPs. Crucially, there was also a very strong trend towards significance for the three-way interaction between DP Type, Gender Congruence and Noun Gender ( $F(1,35) = 4.065$ ,  $MSE = 0.510$ ,  $p = 0.051$ ). Planned pairwise comparisons indicated that for unmixed DPs there was a significant gender congruence effect for nouns with both masculine and feminine gender ( $t(37) = 9.520$ ,  $p < 0.001$ ;  $t(37) = 12.166$ ,  $p < 0.001$ ); however, for mixed DPs the congruence effect was significant only when the gender of the translation equivalent was masculine ( $t(37) = 4.545$ ,  $p < 0.001$ ), and not when it was feminine ( $t(37) = 1.357$ ,  $p = 0.183$ ). There was no significant main effect of Group ( $p = 0.601$ ), nor did it participate in a significant interaction with the other three variables ( $p = 0.138$ ).



**Figure 3.** Ratings results for gender congruence in mixed determiner phrases with French determiner ( $N = 38$ ).

Even though the factor Group did not interact with any of the other independent variables, separate two-way repeated measures ANOVA were conducted for the simultaneous and late bilinguals. This was done because it was hypothesized a priori that these two groups might have different preferences with respect to gender assignment for English nouns. Specifically, it was expected that Late L2-English bilinguals would have a greater tendency to apply analogical gender to English nouns in mixed DPs. As such, these two ANOVAs looked only at mixed DPs with French Ds, and the independent variables were Gender and Congruence. For the late bilinguals, the main effect of Congruence was significant ( $F(1,12) = 12.600$ ,  $MSE = 1.942$ ,  $p = 0.004$ ), but the interaction between Gender and Congruence was not ( $p = 0.496$ ), suggesting that the effect of congruence was the same regardless of the translation’s gender. Indeed, planned pair-wise comparisons indicated that nouns with both feminine and masculine translations were significantly preferred with analogical gender ( $t(12) = 2.758$ ,  $p = 0.017$  and  $t(12) = 2.485$ ,  $p = 0.029$ , respectively). For the simultaneous bilinguals, the main effect of Congruence was significant ( $F(1,17) = 5.316$ ,  $MSE = 2.067$ ,  $p = 0.034$ ) and the interaction between Gender and Congruence was marginally significant ( $F(1,27) = 4.048$ ,  $MSE = 0.467$ ,  $p = 0.060$ ). Planned pair-wise comparisons indicated that the effect of congruence was only significant when the noun’s translation equivalent had masculine gender ( $t(17) = 2.990$ ,  $p = 0.008$ ), but not when it had feminine gender ( $t(17) = 1.061$ ,  $p = 0.303$ ). This supports the hypothesis that the late bilinguals have a stronger preference for assigning analogical gender to English nouns in mixed DPs than do simultaneous bilinguals.

In terms of the four-way ANOVA using RT as the dependent variable (Figure 4), there was a significant main effect of DP Type ( $F(1,34) = 5.783$ ,  $MSE = 247,550.451$ ,  $p = 0.022$ ), but no significant effect of Gender ( $p = 0.386$ ) or of Gender Congruence ( $p = 0.181$ ). There was, however, a significant interaction between DP Type and Gender Congruence ( $F(1,34) = 13.904$ ,  $MSE = 386,115.394$ ,  $p = 0.001$ ). Planned pairwise comparisons suggested that this was again the result of a different pattern for unmixed compared to mixed DPs. For unmixed DPs, gender incongruence between a French N and a French D resulted in significantly slower RTs ( $t(36) = -4.729$ ,  $p < 0.001$ ), as expected; however, for mixed DPs, the RTs were significantly slower when the French determiner was congruent in gender with the English noun’s translation equivalent ( $t(36) = 2.690$ ,  $p = 0.011$ ). This is the opposite of the expected result. While the three-way interaction between DP Type, Gender Congruence and Noun Gender was not significant ( $p = 0.536$ ), planned pair-wise comparisons were conducted in order to further investigate the surprising pattern. The results of these comparisons suggest that it was only in cases where the English N’s translation equivalent was feminine that the RTs were significantly slower in the congruent (e.g., *la house*) compared to the incongruent (e.g., *le house*) condition ( $t(36) = 2.661$ ,  $p = 0.012$ ). When the English noun’s translation equivalent was masculine (e.g., *sun*), there was no significant gender congruency effect ( $t(36) = 0.147$ ,  $p = 0.884$ ). Further, this effect for English Ns with a feminine translation was statistically significant for the late bilinguals ( $t(12) = 2.478$ ,  $p = 0.029$ ), but only marginally significant for the simultaneous bilinguals ( $t(16) = 1.944$ ,  $p = 0.070$ ), despite the greater statistical power of the latter group due to the larger group size. These results will be interpreted and further discussed in the next section.

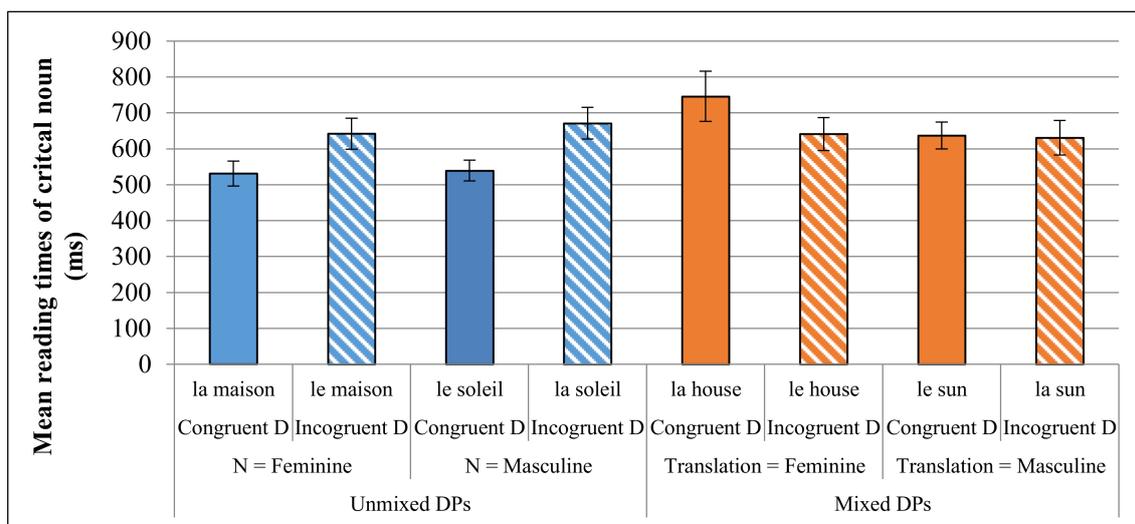


Figure 4. Reading time (RT) results for gender congruence in mixed determiner phrases with French determiner ( $N = 37$ ).

## 6. Discussion

The two primary research questions that the experimental study sought to address were first, whether there was evidence of a French–English mixed DP asymmetry, and second, how the analogical criterion is applied in mixed DPs with an English lexical root. Secondary goals were to investigate the impact of language background factors such as AoA, and to take into account factors that affect processing, such as L2 proficiency and dominance.

With respect to the mixed DP asymmetry, the results for both dependent measures indicated that, overall, there was no preference for switching between a French D and an English N (e.g., *la house*) compared to the reverse (e.g., *the maison*). These results thus support the theoretically motivated hypothesis that there should be no mixed DP asymmetry for French–English.

It must, of course, be conceded that even in Spanish–English language mixing, where the mixed DP asymmetry is expected based on corpora data, AJTs have not provided corroborative evidence of its existence (Licerias et al. 2008, 2016). The experimental task used in the present study, however, differs from those of previous studies in several ways. First, several things were done in order to decrease the saliency of the mixed DPs being investigated—critical DPs were embedded in sentences that had other language-switch points; no unilingual sentences were included, in either critical or filler conditions; and several other types of within- and between-language morphosyntactic anomalies were included in filler items. Second, combining the AJT with SPR might also have had the benefit of reducing the participants’ tendency to be prescriptive or to apply metalinguistic knowledge when making these judgments for two reasons—having comprehension questions after one third of the sentences encouraged participants to focus a significant part of their attention on meaning as opposed to just grammatical structure; and, because the sentences were presented only one word at a time and never in their entirety, participants could not go back and reread sentences. As such, they might not have the same opportunity to deconstruct and consciously analyze their grammatical structure, thus relying more on their intuition as opposed to their metalinguistic knowledge. It is perhaps the application of metalinguistic knowledge that allowed Spanish–English bilinguals to accept DPs such as *the casa* in previous studies; the DM-based analysis in Section 3.1 predicts that such DPs are not ungrammatical per se, but are blocked by the availability of those with a more highly specified determiner (i.e., *la casa*). Intuitively, bilinguals might not like such a DP, but if given time to consider it, they might not know why. Mixed DPs such as *the casa* might even be in their input if there are L1-English L2-Spanish speakers in their environment, for whom the mixed DP asymmetry is not expected. Of course, the only way to determine whether the more implicit methodology used here is capable of detecting a mixed DP asymmetry would be to use it to test a language pair where it is expected, i.e., Spanish–English. One indication that it might be able to do so is that there was robust evidence from the filler items that other uncontroversial language switch violations were indeed rated significantly lower and read significantly more slowly.

It was also hypothesized that late L1-French L2-English bilinguals may show a partial mixed DP asymmetry, specifically a dispreference for mixed DPs consisting of an English D and a feminine French N (e.g., *the maison<sub>F</sub>*). This was hypothesized because such bilinguals may not have a plain *n* in their repertoire of abstract morphemes due to their late AoA, and because the English D *the* is expected to have the same feature specification as the French D *le*. No evidence, however, was found of any asymmetry. This suggests that late bilinguals may have a plain *n*, which could be because plain *n* is a subset of an *n* with a gender feature. Consequently, no “new” morphosyntactic feature actually needs to be acquired by the late learners. It is also possible that such a partial asymmetry could still be found using this methodology with participants having an even later age of L2 immersion, as others assume a later end to the sensitive period for morphosyntax, such as around puberty (e.g., Johnson and Newport 1989; DeKeyser 2000).

The group of late bilinguals was also hypothesized to show increased processing costs for mixed DPs consisting of a French D and an English N because this language switch goes from the dominant to the weak language, following the results of Litcofsky and Hell (2017). Unexpectedly, both the late and simultaneous groups demonstrated this pattern despite the latter group being very balanced. It was later determined that this was because mixed DPs such as *la house* were read significantly more slowly than any other mixed DP with a French D, even those which violated both the analogical criterion and default licensing under the masculine *n* (e.g., *la sun*). This result was also unexpected. It suggests, however, that participants might not have too much difficulty reading/assessing English nouns that appear following the masculine determiner (e.g., *le house*, *le sun*) because any English noun can appear with the default masculine determiner. Likewise, they seem to have little difficulty reading/assessing (and later rejecting) English nouns with masculine translations following a feminine determiner (e.g., *la sun*), suggesting that the issue is not in accessing the licensing conditions of the corresponding French translation equivalent. Rather, it seems that this may be a result of variability

in the input for English nouns with feminine translations (*le/la house*), which is shown to be a robust pattern in Spanish–English corpora (Poplack et al. 1982; Jake et al. 2002; Clegg and Waltermire 2009), and likely also for French (Poplack et al. 1982). Because participants do not have a strong intuition regarding such mixed DPs, this may result in a slower reading time when they encounter one due to the uncertainty. This appears to be particularly so with the late bilinguals, for whom the effect was the most statistically robust. Interestingly, this was also the group who tended to more strongly prefer the analogical criterion, as indicated by their results for acceptability ratings. While this preference was expected for this group, the RT results suggest that its cause is not related to the predictions of the RHM, whereby translations were expected to be more easily accessed by unbalanced sequential bilinguals due to their stronger direct lexical connections. If this were the case, they would have also shown extra processing effort when attempting to access the translation equivalent of masculine nouns following feminine determiners (e.g., *la sun*). Rather, the late bilinguals' preference for analogical gender might reflect a difference in their attitude compared to simultaneous bilinguals, such that the former group attributes greater importance to preserving “French” properties when mixing their languages, thus spending more time deciding what their preference is, and then applying analogical gender preferentially.

These results also suggest that the self-paced reading times of mixed DPs do reflect conscious, strategic preferences to a certain extent. This contrasts with the RTs for unmixed DPs, which really did seem to reflect modulations in processing time related to the grammaticality of particular constructions. Having participants rate the sentences with respect to their acceptability might reduce the benefits of the more implicit RT measure. As such, while it was argued that adding the SPR element to an AJT may decrease the strategic nature of the latter task, this might also increase the strategic nature of the former methodology. It would be interesting to see if the same patterns would persist when the acceptability judgment aspect is removed from the self-paced reading task, rendering it a simple reading task with comprehension questions focusing on meaning.

In a similar vein, the inclusion of within-language morphosyntactic violations may also have affected participants' acceptability ratings of sentences containing mixed DPs. First of all, gender agreement violations in unmixed French DPs are very salient errors to native speakers. As such, drawing participants' conscious attention to such errors may have increased their likelihood of applying the same “rules” to mixed DPs. This could have resulted in an increased tendency to assign analogical gender to English nouns in mixed DPs, particularly for the late bilinguals, who might be more prescriptive. Second, asking participants to use the same acceptability scale to assess within-language violations and potential language switch violations might have also reduced the effectiveness of the task. Bilinguals likely have significant metalinguistic knowledge of within-language morphosyntactic violations as a result of the explicit teaching and correction of grammatical errors, particularly in formal, written (unmixed) language. Language mixing, however, is typically used in informal oral contexts, and bilinguals rarely pay conscious attention to issues of grammaticality. Indeed, the results for the control conditions indicated that participants did have significantly stronger judgments for within-language violations than for language switch violations. This suggests that the scale was more sensitive to the former type of violation, and thus perhaps less able to detect subtle differences in the latter.

## 7. Conclusions

The mixed DP asymmetry of Spanish–English has been investigated from various perspectives over the years, yet there are still many conflicting views with respect to what it means for the underlying linguistic system acquired by bilinguals. The DM-based theoretical account presented in Section 3 assumes a somewhat different underlying system than do previous accounts.

Generally speaking, there are two important theoretical advantages to a DM approach to language mixing. First, because morphology is taken out of the lexicon, this means that the same syntactic processes apply both above and below the word-level (Grimstad et al. 2014; Alexiadou et al. 2015);

“words” are therefore derived from their morphological sub-components in the same way that sentences are derived from their lexical sub-components. As such, word-internal and word-external language mixing, which are often treated separately (e.g., Poplack and Dion 2012; see also Grimstad et al. 2014), can be accounted for, at least formally, using the same mechanisms. Secondly, because VIs from both languages compete for insertion into fully specified syntactic terminal nodes, this allows for optionality when VIs from the two languages have the same feature specifications. This leads to theoretically motivated predictions with respect to where optionality can be expected in mixed language utterances, and where there should be a categorical language preference. This optionality is one of the reasons why it is challenging to model language mixing, yet it is an aspect that DM is particularly well suited to handle.

Demonstrating that the DM framework can account for mixed language data also provides empirical support for the underlying assumptions of this model. In particular, the assumption that gender is not a root property, but a feature on *n*, is supported by the argument that bilinguals have options as to which *n* a translation can become licenced under. Further, the theoretical analysis of Spanish gender and declension class argued for by Kramer (2015) generated specific predictions regarding how this phenomenon was expected to manifest itself in French–English, and also predictions regarding how different paths of dual-language acquisition may affect the patterns. The study presented addressed some of the resulting hypotheses and provided positive evidence for this account, thus lending empirical support to Kramer’s analysis. The two main findings were that there was no evidence for a mixed DP asymmetry for French–English bilinguals, as expected, and that late sequential bilinguals are more likely than simultaneous bilinguals to apply analogical gender to English roots in mixed DPs. The strength of conclusions based on the first result rests on the assumption that such an asymmetry could be found for Spanish–English using this particular experimental methodology. As such, testing this is an important direction for future work. While the second result was predicted based on a model of the bilingual mental lexicon, the reading time data suggest that this pattern might not actually reflect differing types of lexical access across languages for the two groups, but different attitudes and preferences. Conducting experimental studies using even more covert measures, such as event-related potentials and eye-tracking, could help to investigate these questions further.

**Supplementary Materials:** The following are available online at <http://www.mdpi.com/2226-471X/3/2/10/s1>, Table S1: Experimental Stimuli, Table S2: Participant data for acceptability ratings, Table S3: Participant data for reading times, Figure S1: Results for the mixed DP asymmetry by group.

**Acknowledgments:** The author has been funded by a Canada Graduate Scholarship from the Social Sciences and Humanities Research Council, as well as by an Ontario Graduate Scholarship. Many thanks to Éric Mathieu for his valuable guidance regarding theoretical aspects of this paper, as well as to Laura Sabourin for help with the experimental side. Thanks also to Dan Siddiqi for helpful comments on an earlier version of this paper. This paper has also benefited considerably from the comments of two anonymous reviewers and the academic editors. Finally, thanks to members of the Ottawa–Carleton Linguistics Reading Group and the ERP Linguistics Lab at the University of Ottawa for their support and feedback, and of course to all the enthusiastic language mixers who participated in the study.

**Conflicts of Interest:** The author declares no conflict of interest.

## References

- Alexiadou, Artemis, Terje Lohndal, Tor A. Áfarli, and Maren Berg Grimstad. 2015. Language Mixing: A Distributed Morphology approach. In *NELS 45: Proceedings of the Forty-Fifth Annual Meeting of the North East Linguistic Society: Volume 1*. Edited by Thuy Bui and Deniz Özyıldız. North Charleston: CreateSpace Independent Publishing Platform.
- Belazi, Hedi M., Edward J. Rubin, and Almeida Jacqueline Toribio. 1994. Code switching and X-Bar Theory: The Functional Head Constraint. *Linguistic Inquiry* 25: 221–37.
- Brysbaert, Marc, and Boris New. 2009. Moving beyond Kucera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods* 41: 977–90. [CrossRef] [PubMed]

- Burkholder, Michele, Éric Mathieu, and Laura Sabourin. 2017. The role of gender in mixed-language nominal phrases: Insights from Distributed Morphology. Paper presented at Generative Approaches to Second Language Acquisition (GASLA) 14. Southampton, UK, April 7–9.
- Byers-Heinlein, Krista. 2013. Parental language mixing: Its measurement and the relation of mixed input to young bilingual children's vocabulary size. *Bilingualism: Language and Cognition* 16: 32–48. [CrossRef]
- Cantone, Katja Francesca, and Natascha Müller. 2008. Un nase or una nase? What gender marking within switched DPs reveals about the architecture of the bilingual language faculty. *Lingua* 118: 810–26. [CrossRef]
- Chomsky, Noam. 1993. A minimalist program for linguistic theory. In *The view from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*. Edited by Kenneth Hale and Samuel Jay Keyser. Cambridge: MIT Press, pp. 1–52.
- Chomsky, Noam. 1995. Bare phrase structure. In *Government and Binding Theory and the Minimalist Program*. Edited by Gert Webelhuth. Oxford: Blackwell, pp. 71–132.
- Chomsky, Noam. 2005. Three Factors in Language Design. *Linguistic Inquiry* 36: 1–22. [CrossRef]
- Clegg, Jens H., and Mark Waltermire. 2009. Gender assignment to English-origin nouns in the Spanish of the Southwestern United States. *Southwest Journal of Linguistics* 28: 1–18.
- DeKeyser, Robert M. 2000. The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition* 22: 499–533.
- De Vincenzi, Marica, Remo Job, Rosalia Di Matteo, Alessandro Angrilli, Barbara Penolazzi, Laura Ciccarelli, and Francesco Vespignani. 2003. Differences in the perception and time course of syntactic and semantic violations. *Brain and Language* 85: 280–96. [CrossRef]
- Deuchar, Margaret. 2006. Welsh-English code-switching and the Matrix Language Frame model. *Lingua* 116: 1986–2011. [CrossRef]
- Di Sciullo, Anne-Marie, Pieter Muysken, and Rajendra Singh. 1986. Government and code-switching. *Journal of Linguistics* 22: 1–24. [CrossRef]
- Ditman, Tali, Phillip J. Holcomb, and Gina R. Kuperberg. 2007. An investigation of concurrent ERP and self-paced reading methodologies. *Psychophysiology* 44: 927–35. [CrossRef] [PubMed]
- Embick, David, and Rolf Noyer. 2007. Distributed Morphology and the Syntax/Morphology interface. In *The Oxford Handbook of Linguistic Interface*. Edited by Gillian Ramchand and Charles Reiss. Oxford: Oxford University Press, pp. 289–324.
- Fairchild, Sarah, and Janet G. Van Hell. 2017. Determiner-noun code-switching in Spanish heritage speakers. *Bilingualism: Language and Cognition* 20: 150–61. [CrossRef]
- Fernández Fuertes, Raquel, and Juana M. Licerias. 2009. FerFuLice Corpus: English-Spanish Bilingual Twins. CHILDES. Available online: <https://childes.talkbank.org/access/Biling/FerFuLice.html> (accessed on 4 April 2018).
- Fernández Fuertes, Raquel, Juana M. Licerias, and Aurora Bel. 2011. Functional-Lexical Code-Switching in Simultaneous and Sequential Child Bilingualism. Paper presented at the International Symposium on Bilingualism (ISB) 8. Oslo, Norway: University of Oslo, June 15–18.
- González-Vilbazo, Kay, and Luis López. 2011. Some properties of light verbs in code switching. *Lingua* 121: 832–50. [CrossRef]
- Grimstad, Maren Berg, Terje Lohndal, and Tor A. Åfarli. 2014. Language mixing and exoskeletal theory: A case study of word-internal mixing in American Norwegian. *Nordlyd* 41: 213–37. [CrossRef]
- Haberlandt, Karl. 1994. Methods in reading research. In *Handbook in Psycholinguistics*. Edited by Morton Ann Gernsbacher. San Diego: Academic Press, Inc., pp. 1–31.
- Halle, Morris, and Alec Marantz. 1993. Distributed Morphology and the Pieces of Inflection. In *The View from Building 20*. Edited by Kenneth Hale and Samuel Jay Keyser. Cambridge: MIT Press, pp. 111–76.
- Halle, Morris. 1997. Distributed Morphology: Impoverishment and Fission. *MIT Working Papers in Linguistics* 30: 425–49.
- Harley, Heidi. 2014. On the identity of roots. *Theoretical Linguistics* 40: 225–76. [CrossRef]
- Hawkins, Roger, and Cecilia Yuet-hung Chan. 1997. The partial availability of Universal Grammar in second language acquisition: The 'failed functional features hypothesis'. *Second Language Research* 13: 187–226. [CrossRef]
- Jake, Janice L., Carol Myers-Scotton, and Steven Gross. 2002. Making a minimalist approach to code switching work: Adding the matrix language. *Bilingualism: Language and Cognition* 5: 69–91. [CrossRef]

- Jake, Janice L., Carol Myers-Scotton, and Steven Gross. 2005. A response to MacSwan (2005): Keeping the matrix language. *Bilingualism: Language and Cognition* 8: 271–76. [CrossRef]
- Johnson, Jacqueline S., and Elissa L. Newport. 1989. Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology* 21: 60–99. [CrossRef]
- Jorschick, Liane, Antje Endesfelder Quick, Dana Glässer, Elena Lieven, and Michael Tomasello. 2011. German–English-speaking children’s mixed NPs with ‘correct’ agreement. *Bilingualism: Language and Cognition* 14: 173–83. [CrossRef]
- Joshi, Aravind K. 1985. Processing of sentences with intrasentential code switching. In *Natural Language Parsing: Psychological, Computational and Theoretical Perspectives*. Edited by David R. Dowty, Lauri Kattunen and Arnold M. Zwicky. Cambridge: Cambridge University Press, pp. 190–205.
- Julien, Marit. 2005. *Nominal Phrases from a Scandinavian Perspective*. Amsterdam: John Benjamins.
- Kihm, Alain. 2005. Noun class, gender and the lexicon-syntax interface: A comparative study of Niger-Congo and Romance languages. In *The Oxford Handbook of Comparative Syntax*. Edited by Guglielmo Cinque and Richard S. Kayne. Oxford: Blackwell, pp. 459–512.
- Kramer, Ruth. 2015. *The Morphosyntax of Gender*. Oxford: Oxford University Press.
- Kroff, Jorge R. Valdés, Paola E. Dussias, Chip Gerfen, Lauren Perrotti, and M. Teresa Bajo. 2016. Experience with code-switching modulates the use of grammatical gender during sentence processing. *Linguistic Approaches to Bilingualism* 7: 163–98. [CrossRef] [PubMed]
- Kroll, Judith F., and Erika Stewart. 1994. Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language* 33: 149–74. [CrossRef]
- Kučerová, Ivona. 2017.  $\phi$ -Features at the Syntax-Semantics Interface: Evidence from Nominal Inflection. *Linguistic Inquiry*. Available online: <http://ling.auf.net/lingbuzz/003483> (accessed on 4 April 2018).
- Liceras, Juana M. 2014. La adquisición de las lenguas segundas aquí y ahora o... cómo abordar la hipótesis de la interlengua en el siglo XXI. In *El español global*. Edited by J. Santiago Guervós and Y. González Plasencia. Valladolid: Fundación Siglo para el Turismo y las Artes de Castilla y León, pp. 124–48.
- Liceras, Juana M., Raquel Fernández-Fuertes, Susana Perales, Rocío Pérez-Tattam, and Cristina Martínez. 2006. L2A and the pidgin/creole continuum: The pidginization/nativization hypothesis revisited. Paper presented at EUROSLA. Antalya, Turkey: Boğaziçi University, September 13–16.
- Liceras, Juana M., Raquel Fernández Fuertes, Susana Perales, Rocío Pérez-Tattam, and Kenton Todd Spradlin. 2008. Gender and gender agreement in bilingual native and non-native grammars: A view from child and adult functional–lexical mixings. *Lingua* 118: 827–51. [CrossRef]
- Liceras, Juana M., Raquel Fernández Fuertes, and Rachel Klassen. 2016. Language dominance and language nativeness: The view from English-Spanish code-switching. In *Spanish-English Codeswitching in the Caribbean and the U.S.* Edited by R. E. Guzzardo Tamargo, C. M. Mazak and M. C. Parafita Couto. Amsterdam and Philadelphia: John Benjamins, pp. 107–38.
- Lieber, Rochelle. 1992. *Deconstructing Morphology*. Chicago: University of Chicago Press.
- Lipski, John. 1978. Code-switching and the problem of bilingual competence. In *Aspects of Bilingualism*. Edited by Michel Paradis. Columbia: Hornbeam Press, pp. 250–64.
- Litcofsky, Kaitlyn A., and Janet G. Van Hell. 2017. Switching direction affects switching costs: Behavioral, ERP and time-frequency analyses of intra-sentential codeswitching. *Neuropsychologia* 97: 112–39. [CrossRef] [PubMed]
- Lohndal, Terje, and Marit Westergaard. 2016. Grammatical gender in American Norwegian heritage language: Stability or attrition? *Frontiers in Psychology* 7: 344. [CrossRef] [PubMed]
- MacSwan, Jeff. 1999. *A Minimalist Approach to Intrasentential Code Switching*. New York: Garland.
- MacSwan, Jeff. 2000. The architecture of the bilingual faculty: Evidence from intrasentential code switching. *Bilingualism: Language and Cognition* 3: 37–54. [CrossRef]
- MacSwan, Jeff. 2005. Codeswitching and generative grammar: A critique of the MLF model and some remarks on “modified minimalism”. *Bilingualism: Language and Cognition* 8: 1–22. [CrossRef]
- MacWhinney, Brian. 2000. *The CHILDES Project: Tools for Analyzing Talk*. Mahwah: Lawrence Erlbaum Associates, Available online: <https://childes.talkbank.org> (accessed on 4 April 2018).
- Mahootian, Shahrzad. 1993. A Null Theory of Code-Switching. Ph.D. dissertation, Northwestern University, Chicago, IL, USA.

- Marantz, Alec. 1997. No escape from syntax: Don't try morphological analysis in the privacy of your own Lexicon. In *Proceedings of the 21st Annual Penn Linguistics Colloquium: Penn Working Papers in Linguistics 4*. Edited by Alexis Dimitriadis. Philadelphia: Department of Linguistics, University of Pennsylvania, pp. 201–25.
- Moro, Monica. 2001. *The Semantic Interpretation and Syntactic Distribution of Determiner Phrases in Spanish/English Codeswitching*, Paper presented at the 3rd International Symposium on Bilingualism; Bristol, UK, April 17–24.
- Moro Quintanilla, Monica. 2014. The semantic interpretation and syntactic distribution of determiner phrases in Spanish/English codeswitching. In *Grammatical theory and Bilingual Codeswitching*. Edited by Jeff MacSwan. Cambridge: MIT Press, pp. 213–26.
- Moyer, Melissa Greer. 1993. Analysis of Code-Switching in Gibraltar. Ph.D. dissertation, Universidad Autónoma de Barcelona, Barcelona, Spain.
- Müller, Stefan, and Stephen Wechsler. 2014. Lexical approaches to argument structure. *Theoretical Linguistics* 40: 1–76. [CrossRef]
- New, Boris, Christophe Pallier, Marc Brysbaert, and Ludovic Ferrand. 2004. Lexique 2: A new French lexical database. *Behavior Research Methods, Instruments, & Computers: A Journal of the Psychonomic Society* 36: 516–24.
- Pesetsky, David, and Esther Torrego. 2001. T-to-C Movement: Causes and Consequences. In *Ken Hale: A life in Language*. Edited by Michael Kenstowicz. Cambridge: The MIT Press, pp. 355–426.
- Pfaff, Carol W. 1979. Constraints on language mixing: Intrasentential code-switching and borrowing in Spanish/English. *Language* 55: 291–318. [CrossRef]
- Poplack, Shana, Alicia Pousada, and David Sankoff. 1982. Competing influences on gender assignment: Variable process, stable outcome. *Lingua* 56: 139–66. [CrossRef]
- Poplack, Shana, and Nathalie Dion. 2012. Myths and facts about loanword development. *Language Variation and Change* 24: 279–315. [CrossRef]
- Poplack, Shana. 1980. “Sometimes I’ll start a sentence in Spanish y termino en Espanol”: Toward a typology of code-switching. *Linguistics* 18: 581–618. [CrossRef]
- Poplack, Shana. 1981. The syntactic structure and social function of code-switching. In *Latino Language and Communicative Behavior*. Edited by Richard Duran. Norwood: Ablex, pp. 169–84.
- Riksem, Brita Ramsevick. 2017. Language Mixing and Diachronic Change: American Norwegian Noun Phrases Then and Now. *Languages* 2: 3. [CrossRef]
- Sabourin, Laura, Christie Brien, and Michèle Burkholder. 2014. The effect of age of L2 acquisition on the organization of the bilingual lexicon: Evidence from masked priming. *Bilingualism: Language and Cognition* 17: 542–55. [CrossRef]
- Sabourin, Laura, Jean-Christophe Leclerc, Myriam Lapierre, Michèle Burkholder, and Christie Brien. 2016. The Language Background Questionnaire in L2 Research: Teasing Apart the Variables. In *Proceedings of the 2016 Annual Conference of the Canadian Linguistic Association*. Edited by Lindsay Hracs. Available online: [http://cla-acl.ca/wp-content/uploads/actes-2016/Sabourin\\_etal\\_CLA2016\\_proceedings.pdf](http://cla-acl.ca/wp-content/uploads/actes-2016/Sabourin_etal_CLA2016_proceedings.pdf) (accessed on 4 April 2018).
- Sankoff, David, and Shana Poplack. 1981. A formal grammar for code-switching. *Research on Language and Social Interaction* 14: 3–45.
- Siddiqi, Daniel. 2009. *Syntax within the Word: Economy, Allomorphy, and Argument Selection in Distributed Morphology*. Amsterdam: John Benjamins.
- Van Dulm, Ondene. 2009. English-Afrikaans intrasentential code-switching: Testing a feature checking account. *Bilingualism: Language and Cognition* 12: 193–212. [CrossRef]
- Van Gelderen, Elly, and Jeff MacSwan. 2008. Interface conditions and code-switching: Pronouns, lexical DPs, and checking theory. *Lingua* 118: 765–76. [CrossRef]
- Woolford, Ellen. 1983. Bilingual Code-Switching and Syntactic Theory. *Linguistic Inquiry* 14: 520–36.

