

*English compound and non-compound processing in bilingual and multilingual speakers: effects of dominance and sequential multilingualism*

Article

Accepted Version

Gonzalez Alonso, J., Villegas, J. and García Mayo, M. d. P. (2016) English compound and non-compound processing in bilingual and multilingual speakers: effects of dominance and sequential multilingualism. *Second Language Research*, 32 (4). pp. 503-535. ISSN 1477-0326 doi: 10.1177/0267658316642819 Available at <https://centaur.reading.ac.uk/59325/>

It is advisable to refer to the publisher's version if you intend to cite from the work. See [Guidance on citing](#).

To link to this article DOI: <http://dx.doi.org/10.1177/0267658316642819>

Publisher: Sage

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the [End User Agreement](#).

[www.reading.ac.uk/centaur](http://www.reading.ac.uk/centaur)

## **CentAUR**

Central Archive at the University of Reading

Reading's research outputs online

English compound and non-compound processing in bilingual and multilingual speakers: Effects of dominance and sequential multilingualism.

**Jorge González Alonso<sup>+</sup>**

Department of Clinical Language Sciences, University of Reading, UK

**Julián Villegas**

Computer Arts Laboratory, University of Aizu, Japan

**María del Pilar García Mayo<sup>+</sup>**

Departamento de Filología Inglesa, Universidad del País Vasco (UPV/EHU), Spain

<sup>+</sup> It should be noted that ‘González Alonso’ and ‘García Mayo’ are double family names, following Spanish naming customs.

**Corresponding author:**

Jorge González Alonso, School of Psychology and Clinical Language Sciences, Harry Pitt Building, Earley Gate, Reading, UK – RG6 7BE.

Phone no.: (+44) 7459855490

e-mail: [jorge.gonzalezalonso@reading.ac.uk](mailto:jorge.gonzalezalonso@reading.ac.uk)

**Acknowledgements**

Previous versions of this paper were presented at the *Hispanic Linguistics Symposium* in 2012 and at *GASLA 12* in 2013. This work was supported by a doctoral grant [AP2010-2677] awarded by the Spanish Ministry of Education to the first author, and funding from the Basque Government to the *Language and Speech* research group [IT-311-10], coordinated by the third author. We would like to express our gratitude to Tiffany Judy, Drew Long and Anne Lingwall for help with data collection in the USA, and to Jason Rothman, Edith Kaan, Oliver Müller and three anonymous reviewers for helpful comments on earlier versions of this article. Any and all errors, misconceptions and omissions remain our own.

**Declaration of conflicting interest**

The authors declare that there is no conflict of interest.

## Abstract

This article reports on a study investigating the relative influence of the first and dominant language on L2 and L3 morpho-lexical processing. A lexical decision task compared the responses to English NV-er compounds (e.g., *taxi driver*) and non-compounds provided by a group of native speakers and three groups of learners at various levels of English proficiency: L1 Spanish-L2 English sequential bilinguals and two groups of early Spanish-Basque bilinguals with English as their L3. Crucially, the two trilingual groups differed in their first and dominant language (i.e., L1 Spanish-L2 Basque vs. L1 Basque-L2 Spanish). Our materials exploit an (a)symmetry between these languages: while Basque and English pattern together in the basic structure of (productive) NV-er compounds, Spanish presents a construction that differs in directionality as well as inflection of the verbal element ( $V_{[3SG]} + N$ ). Results show between and within group differences in accuracy and response times that may be ascribable to two factors besides proficiency: the number of languages spoken by a given participant and their dominant language. An examination of response bias reveals an influence of the participants' first and dominant language on the processing of NV-er compounds. Our data suggest that morphological information in the nonnative lexicon may extend beyond morphemic structure and that, similarly to bilingualism, there are costs to sequential multilingualism in lexical retrieval.

## **Introduction**

A substantial body of evidence suggests that entries in the lexicon of native speakers include information on the word's morphemic structure, that is, the morphemes that combine to form the word (see, e.g., Fiorentino and Poeppel, 2007; Fiorentino et al., 2014; Stockall and Marantz, 2006; Taft and Forster, 1976, among many others; cf. Bybee, 1995; Kuperman, 2013). This morphemic breakdown is part of the information used during processing, which, in most cases, involves morphological decomposition for complex words (see Marslen-Wilson, 2007, for an extensive review).

Similar studies have been conducted with nonnative speakers, examining the processing of inflection and derivation (e.g., Diependaele et al., 2011; Feldman et al., 2010; Neubauer and Clahsen, 2009; Silva and Clahsen, 2008). While results are mixed, two main positions can be identified. The first assumes that native and nonnative processing of morphology may differ quantitatively (in processing speed and capacity) but not qualitatively (e.g., Diependaele et al., 2011; Duñabeitia et al., 2013). The second view maintains that native and nonnative speakers do employ substantially different mechanisms (e.g., Clahsen et al., 2013). Results of recent studies on English compounds (De Cat et al., 2015; González Alonso, Baquero Castellanos & Müller, submitted) suggest that, at a certain level of proficiency, the nonnative lexicon already contains information on morphemic structure for complex words.

It remains unclear whether the morphological information present in the lexical entries of nonnative speakers is restricted to morphemic structure. Following evidence from native speakers, it could also contain some specification of the semantic and grammatical relationships established between those morphemes, i.e., the word's relational structure. Within the past two decades, research on native compound processing has often highlighted the importance of relational structures for meaning integration in the compositional stage of compound processing (e.g., Gagné and Shoben, 1997; Gagné and Spalding, 2009; Marelli and Luzzatti, 2012; see, e.g., Downing, 1977; Lieber, 1983, for approaches within linguistic theory). These relational structures convey different types of relationships between compound constituents, and, with some variations, are common to large classes of them. English synthetic compounds, in which the head constituent is deverbal in origin, are an illustrative case of how these relational structures productively generalise across large numbers of compounds. Consider the sentences below:

- (1) James may occasionally have to drive or fly home, but never if he has the luxury of time: he is a real *train-admirer*.
- (2) Pixar's new film is a champion of its generation, an example of everything that is right about 21<sup>st</sup> century animation: inspiring, funny, adult friendly, *child-mesmerising*.
- (3) Oh, don't worry! I'll be fine in the heat. Remember I'm a *desert-trained* soldier.

The reason why we can coin and readily interpret novel or highly infrequent words like *train-admirer*, *child-mesmerising* or *desert-trained* is that the assignation of thematic roles involved in their constituent structure is well known to us: compounds such as *taxi driver*, *life-changing* and *homeschooled* encode the same kind of relationships between their constituents as the coinages in (1), (2) and (3), respectively. A number of questions follow from this observation: are these patterns, which extend beyond individual lexical entries, part of a speaker's morphological knowledge of a language? If so, do they have to be learned by nonnative speakers? Can these patterns generalise across various languages in a multilingual system, provided that characteristics of relative constituent position and theta-role assignment are similar? If so, does frequency/recency of use enhance their processing? Would this facilitation override potential hampering effects of sequential multilingualism<sup>1</sup>?

Multilingualism offers an important testing ground for long-standing arguments in psycholinguistics regarding computation versus storage capacity (e.g. Hay and Baayen, 2005). As of yet, it is unclear whether the incorporation of new languages into the multilingual brain increases the cognitive load (De Bot and Jaensch, 2015), thus slightly reducing general processing speed, and how this potential load might be attenuated or even removed at higher levels of proficiency. If it were the case that linguistic processing, in whole or in part, is significantly affected by sequential multilingualism—a frequent scenario in present day societies—this would imply that there are costs to

maintaining more than one language. Two different efforts have been proposed as deriving these costs: the need to keep separate lexicons apart in order to speak fluently in only one language at a time, and the need to select words amongst competitors in a densely populated common lexicon.

The study presented here aimed at examining the potential influence of the first and dominant language in second and third language processing, focusing on the NV-er type of synthetic compounds—e.g. example (1) above. Due to the different characteristics of our participant groups, our design also warrants some speculation on the effects of sequential multilingualism on lexical retrieval.

## **Background**

### *Relational structures in native compound processing*

Within the visual word processing literature, lexical structure has often been probed through various versions of the priming paradigm. This methodological paradigm compares the influence of a previously presented letter string (a word, a morpheme or a nonword; the *prime*) on the processing of another (the *target*).<sup>2</sup> If an effect arises, it is assumed that prime and target are related to a certain extent at one or more levels of representation. Studies on compounding have frequently reported repetition priming from constituents to compounds (e.g. *bell* – *DOORBELL*; Libben et al., 2003). Partial repetition priming from compounds to constituents (e.g. *doorbell* – *BELL*; Fiorentino



and Fund-Reznicek, 2009) and compounds to compounds (e.g., *doorbell* – *DOORSTOP*) have also been reported, even in the absence of full semantic transparency and across constituent positions (e.g. *doorbell* – *BELLHOP*; Duñabeitia et al., 2009). Likewise, some studies have reported semantic priming effects for transparent and partially transparent compounds (e.g. *infant* – *BABYSITTER*; Sandra, 1990; Zwitserlood, 1994). These effects suggest that there is constituent activation during compound processing, and that lexical entries for compounds are morphemically structured. The question remains as to whether the morphological information contained in these entries is more detailed. One way in which this information could be richer is by encoding not only the constituent morphemes of these words, but also the structural relations that hold between them.

Gagné and Shoben (1997) investigated the role that relational structures play in compound processing by manipulating the frequency with which the modifier noun in an N-N compound (e.g. *snowball*) entered into a given relationship with the head noun. Those compounds in which the relational structure was unusual for the modifier (e.g., *mountain magazine*, where *mountain* more generally appears indicating location, as in *mountain cabin*) took longer to process than those in which the modifier-head relationship was frequent for the modifier noun (in the present example, *mountain cloud*). These and other results (see, e.g., Gagné, 2001) have given rise to a theory of conceptual combination in which the weight of compound relational properties crucially

depends on the modifier. Investigating beyond novel compounds, Gagné and Spalding (2009) found that lexicalised compounds such as *snowman* (i.e., a man MADE OF snow) yielded shorter response latencies when preceded by others with the same relational structure (e.g., *snowball*) than they did when preceded by compounds with the same modifier but a different relational structure (e.g., *snow shovel*, a shovel FOR snow). These findings are seen by the authors as evidence that meaning integration takes place even for familiar compounds, and that the lexical entries for these compounds are specified for constituent position and relational structure.

The above discussed studies highlight that, for native speakers, relational structures are part of the morphological information specified in the lexical entries of root compounds (e.g., *snowball*). We turn now to a class of English compounds in which the presence of a high-frequency derivational suffix arguably enhances the saliency of the structural relationship between constituents: synthetic NV-er compounds.

#### *Synthetic NV-er compounds*

NV-er compounds<sup>3</sup> are a productive nominal subtype of English synthetic compounds, a class which also includes the adjectival NV-ed (e.g., *homeschooled*) and the nominal NV-ing (e.g., *housekeeping*). Synthetic compounds contrast with primary or root compounds (e.g., *footnote*) in that their head constituent is derived from a verb through affixation. Some authors (Booij, 2005) consider that the head is verb-derived first and

then takes the non-head noun as its object (i.e., [N[N *heart*] [N[V *break*] *-er*]]). This is because the alternative analysis (i.e., [N[V[N *heart*] [V *break*]] *-er*]) would involve verbal compounding, which is not productive in English (see also Piera, 1985). On the other hand, from the perspective of Level-ordered Morphology (Clahsen et al., 1992; Kiparsky, 1982) {-er} derivation and compounding take place at the same (second) level, which would mean that synthetic NV-er compounds are generated simultaneously with their heads.

Irrespective of the order in which these morphological processes take place, root and synthetic compounds crucially differ in that the head of the latter retains the verb's argument structure. Compared to root compounds, where the non-head constituent can be interpreted as a modifier of different types, synthetic compound's deverbal heads engage in more interpretable, less ambiguous thematic relations with their arguments (Lieber, 2004)—one of which, the object, is almost invariably satisfied internally. This raises the question of whether a straightforward relational structure of this kind can be acquired by nonnative speakers and generalised across compounds of this class. In order to present some of the work that has been conducted on the acquisition and use of these structures by nonnative speakers of English, let us first introduce the equivalent forms in the languages of interest: Spanish and Basque.

*Verb-object compounds in Spanish.* Spanish deverbal compounds differ from their English NV-er equivalents most notably in the presence of a finite verb in these constructions. Compare the examples in (4) and (5) (head constituents in English are underlined):

(4) a. a storyteller

b. \*a storiestteller

c. a mice eater

(5) a. un cuentacuentos [lit. tell-3SG story-PL, 'storyteller']

b. \*un cuentascuentos [tell-2SG story-PL]

c. \*un cuentacuento [tell-3SG story-SG]

d. un crecepelo [lit. grow-3SG hair-SG, 'hair tonic']

e. un crecepelos [lit. grow-3SG hair-PL, 'hair tonic']

f. un ganapán [lit. win-3SG bread-SG, 'breadwinner']

g. \*un ganapanes [win-3SG bread-PL]

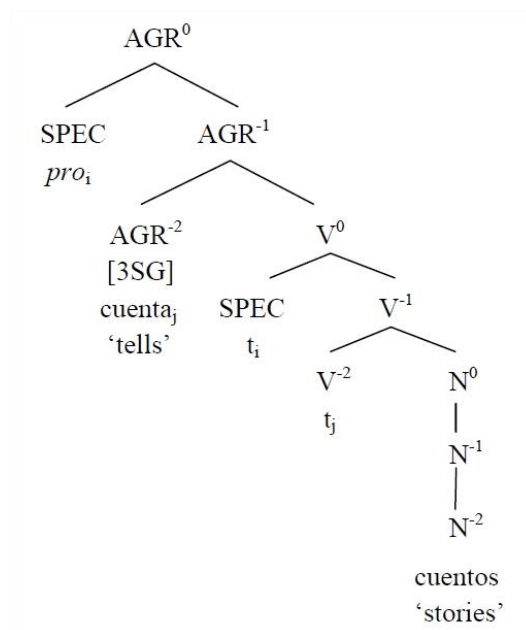
In Spanish verb-object compounds (also called V+N) the verbal element displays a syntax-like leftmost position (Spanish, like English, has head-initial VPs). This constituent is invariably inflected in 3<sup>rd</sup> person singular form (compare 5a and 5b), and takes the second constituent as an object. Since the root verbs of these compounds are usually transitive, an agentive role is assigned to a null *pro* external subject (Di Sciullo,

1991; Lardiere, 1995), while the second constituent is generally assigned a theme role. Under some analyses (e.g. Yoon 2009), this is sometimes a locative, as in *saltamontes* (lit. jumps [on the] mountains, ‘grasshopper’). Note that, in English NV-er compounds, both arguments are satisfied internally: the theme role is assigned to the non-head constituent, and the agent or experiencer to the agentive suffix, {-er} (Di Sciullo, 1992; Lardiere, 1995; Lardiere and Schwartz, 1997).

Unlike English, which restricts the use of plurals within compounds by allowing only irregular forms (cf. 4b and 4c), Spanish allows both singular and plural object constituents only if the non-head is a noun which can be both countable and uncountable (5d-e). Countable nouns (e.g. *cuento*, ‘story’) are always pluralised (5a vs. 5c), while non-countable nouns (e.g. *pan*, ‘bread’) appear always in singular form (5f-g). This default to the plural has been proposed by Lardiere (1995) to reflect the non-referentiality required at the ‘surface’ AGR<sup>0</sup> level, since it is assumed that these compounds are derived by syntactic rules extending to the sub-syntactic (i.e. lexical) level (see also Di Sciullo, 1991, 1992; Roeper, 1988).

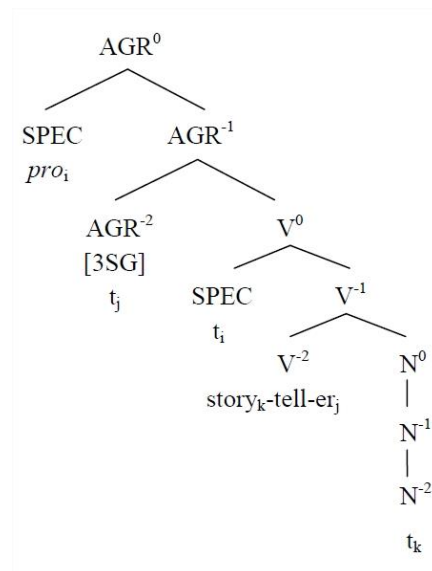
It is not clear whether Spanish deverbal compounds are in fact endocentric, that is, whether one of their constituents can be considered the head of the construction. In a traditional interpretation of headedness, they should be considered exocentric, since none of their constituents bears the main semantic and/or categorisation weight of the compound: *cuentacuentos* is not a type of ‘cuenta’ nor a type of ‘cuento(s)’ (Piera,

1985), and the whole compound is not categorised as a verb, as it should be if the first constituent were the head (Moyna, 2011). In contrast, in English NV-er compounds the head constituent alone already describes the subject referent (e.g., a *taxi driver* is a kind of driver). However, in a ‘syntactic’ analysis of these compounds (e.g., Lardiere and Schwartz, 1997) they would be headed by the verbal element. This element is located at  $\text{AGR}^{-2}$ , itself head of  $\text{AGR}^0$  (i.e., the surface representation of the compound; see Figure 1 for an example analysis). The aim of such syntactic approaches to morphology is to provide a uniform account of derivational computations ‘above and below the  $\text{X}^0$  level’ (Lardiere, 1998: 285).



**FIGURE 1.** Analysis of Spanish deverbal compounds in a syntactic (as opposed to lexicalist) account. Adapted from Lardiere and Schwartz (1997). Note: here and in Figure 2, negative superscripts are used to indicate that these are sub-syntactic levels.

Under such analyses, Spanish and English deverbal compounds are not as different. In English, the head position of AGR, equally specified as [3SG], would initially be occupied by the agentive suffix {-er}, a morpheme that would subsequently move down to  $V^{-2}$  in order to attach to the verb (which does not rise in English). Similarly, the object noun would be incorporated into the verb, losing its plural inflection. (Lardiere and Schwartz, 1997, point to Anderson, 1985, as an example of analyses considering the deletion of nominal inflection as a characteristic trait of incorporation processes). Figure 2 exemplifies this analysis.



**FIGURE 2.** Analysis of English NV-er compounds under a syntactic account. Adapted from Lardiere and Schwartz (1997).

Contreras (1985; see Kornfeld, 2009, for a similar approach) proposes a different analysis, in which the compound would be the (phonologically realised) VP complement of an NP headed by an empty nominal, e.g., ‘un cuentacuentos’  $NP[NP[Det\ Un + N' e] + VP[v\ cuenta + N\ cuentos]]$ , where  $e$  is the empty head. What both syntactic analyses have in common is that they account for the exocentricity of these compounds without resorting to an external element. Varela (2012) is critical of these approaches, and claims that these compounds are endocentric N+N constructions headed by the deverbal left-most element, nominalised through an internal affix. In her view, characterising these compounds as reduced VPs or NPs with an elliptical head does not account for certain constraints on the verbal element. The most important of these are the predominance of transitive-agentive verbs (cf. *\*un duerme*, ‘a sleeps’) and the restriction on light verbs (cf. *\*un dapaseos*, ‘a takes walks’).

No parallel analysis in the line of Contreras (1985) is offered for English semantic equivalents (i.e., NV-er compounds). However, some authors (e.g., Tuggy, 2003) have noted that a rare class of English deverbal compounds is considerably more comparable in structural terms. Nouns like *scarecrow*, *killjoy* or *breakwater* combine a verbal and a nominal element with the same directionality and the same relational structure as Spanish V+N compounds. Still, two notable differences remain. The first is that English



V+N compounds are no longer productive. The second, which still holds for NV-er compounds, is that none of their constituents are inflected. Recall that, in Spanish, the verbal element is conjugated in the third person singular form and the complement noun is generally pluralised. Under a Cognitive Grammar approach (e.g., Tuggy, 2003; Yoon, 2009), English NV-er and V+N compounds are different patterns mapping to the same higher-order abstraction or schema (i.e., the one that encodes the thematic relationship expressed in both). In this sense, English NV-er and Spanish V+N compounds cannot be considered structurally equivalent, even if they eventually abstract to a similar schema. It is unclear what implied predictions this would have for processing. However, Tuggy (2003) emphasises that low-level patterns outrank higher-order constructions in sanctioning novel formations. If this is also true, at least partially, for existing constructions, then structural differences should outweigh semantic similarities between Spanish V+N (e.g., *cuentacuentos*) and English NV-er compounds (e.g., *storyteller*). To our understanding, such approaches predict no effects to arise in processing from an interaction of these constructions.

From a formal perspective, most work done on the acquisition of English synthetic compounds by native speakers of Spanish has focused on the interlanguage representations (and their surface expressions) that result from a gradual adjustment to the internal organisation of English compounds. Lardiere (1995) tested 15 native speakers of Spanish and 11 native speakers of Chinese, all adults at an upper-

intermediate level of proficiency in English. Their production of novel NV-er compounds was elicited by a verbal prompt (e.g., *What would you call someone who wears socks?*; expected response: *A sock-wearer*). Spanish speakers produced a considerable number of regular plurals within compounds (e.g. *\*socks-wearer*). Compared to the much lower percentage of such errors in the Chinese group (21% vs. 47%), these data suggest an influence of the L1 that has not yet been overcome by full acquisition of the English underlying structure. Lardiere and Schwartz (1997) used the same task with a group of 34 native Spanish speakers at different levels of proficiency in L2 English, to examine errors of head-directionality (e.g. *\*wearer socks*). While their error rate was inversely correlated with proficiency, a large number of this type of errors appeared in the responses of the low- and medium-proficiency groups. Taken together, these instances are interpreted by Lardiere (1995) and Lardiere and Schwartz (1997) as reflective of a stage of interlanguage in which nonnative speakers have not yet mastered incorporation (and its associated nominal inflection deletion) in their formation of English NV-er compounds.

It is not clear, however, whether the plural morpheme in Spanish compounds is indeed such a strong feature. In a lexical decision task involving, among others, violations of head-directionality (e.g. *\*latasabre*, lit. ‘can-PL open-3SG’) and plural marking (e.g. *\*abrelata*, lit. ‘open-3SG can-SG’), native speakers of Spanish had the most problems rejecting the latter, with reaction times that differed significantly from

any other nonword condition and an error rate that bordered chance performance (Desrochers et al., 2003). The persistence of the plural morpheme (despite a correct head-final directionality) in the interlanguage of nonnative speakers remains unexplained. While it might be due to a delay in the feature deletion associated with incorporation, as Lardiere and Schwartz (1997) suggest, it could also reflect the use of the plural as a way of encoding non-specificity and non-referentiality.

*Synthetic NV-er compounds in Basque.* The comparable compound construction in Basque can be divided into two sub-classes, depending on the theta-role assumed by the head noun: agentive and instrumental. Unlike Spanish and English, where the same morpheme may carry either thematic role—compare *screwdriver* (instrumental) to *taxi driver* (agentive)—Basque encodes these roles through different morphemes. Consider the examples in (6) (head constituents are underlined):

(6) a. *liburu*        *sal-tzaile-a*

book-SG    sell-AG-Det

‘the book seller’

b. \**liburu-ak*        *sal-tzaile-a*

book-Det<sub>[+PL]</sub>    sell-AG-Det

c. *bide*        *erakus-le-a*

way-SG    show-AG-Det

‘the guide’

d. *lata        ireki-gailu-a*

can-SG    open-INST-Det

‘the can-opener’

e. *Liburu sal-tzaile-a-k        ipuin    hau    gomendatu    dit.*

Book    sell-AG-Det-ERG    tale        this    recommend    has-TO 1SG

‘The book seller has recommended this tale to me.’

*Note:* AG = agentive marker; Det = Determiner; ERG = ergative marker.

Basque, like English, has right-headed deverbal compounds (although see Bauer and Renouf, 2001, for a revision of right-headedness in English). The head noun is derived from the verb stem through affixation, and takes the first constituent as an object (generally with the theta-role of theme). Basque compound nouns, like noun phrases, display no internal inflections or case markings, which attach to the end of the NP (cf. 6e; Laka, 1996). This has the consequence that object nouns within compounds (e.g., *liburu* in *liburu saltzaile*) cannot be inflected for number (6b), since all Basque nouns are pluralised through the determiner (e.g., example 3b would read ‘a the-stories teller’ in English, or ‘un cuenta-los-cuentos’ in Spanish). We will not venture here an analysis in Lardiere and Schwartz’s terms, but it seems reasonable to assume that the object noun’s original determiner would be deleted along with its features upon incorporation into the verb. It seems clear that, even within a syntactic analysis, there is greater

structural similarity between Basque and English deverbal compounds than there is between English and Spanish. The extent to which these similarities will be reflected in processing will likely depend on how much of these computations is predicted to play a role in real-time processing. If surface structure is the only morphological information that can exert cross-linguistic influence, then the asymmetrical pattern of similarities between Spanish, Basque and English has the potential to be visible in processing. If more complex computation is available, differences are less apparent and should therefore have less influence.

Few studies have been conducted on the acquisition and use of English compounds by Basque native speakers. Building on work by Murphy (2000), García Mayo (2006) tested 49 Basque-Spanish bilinguals on their production of novel English NV-er compounds, elicited through a verbal prompt (e.g., *What do you call someone who jumps fences?* Expected response: *A fence-jumper*; see Clark et al., 1986). Low-proficiency participants produced a high number of regular plurals within compounds (51.5%), while their advanced peers did so only in 22% of the cases—still a high proportion, but far from the 75% consistency criterion proposed by Clark et al. (1986). García Mayo found no instances in the L2 data of directionality errors (e.g., *\*wearer socks*). This finding, in stark contrast with previous evidence from studies on Spanish-speaking L2 learners of English, was attributed to a combined influence of Basque

(from which participants could have transferred the appropriate directionality) and Spanish (from which they would have transferred the plural marking).

Our experiment focused on a possible influence of Basque on the recognition English NV-er compounds through their relational structure. In particular, we were interested in a potential modulation of this influence by a variable not controlled for in García Mayo's (2006) study: bilingual dominance. The comparison between L2 and L3 learners allowed us to touch on the subject of processing costs derived from the presence of additional languages.

#### *The bilingual disadvantage*

From very early on (e.g. Cattell, 1887), psycholinguists have agreed that L2 processing is, in general terms, slower than L1 processing. This relatively uncontroversial observation raises the question of whether such a delay stems from the representational or cognitive nature of nonnative languages, or is more generally a consequence of multilingualism taking up more resources. While the answer might lie somewhere in between both accounts, they have largely been explored independently of each other. Importantly, scarce work has been conducted specifically on multilingualism. However, although most of the studies presented in this section were focused on bilinguals, we believe that a majority of their findings can uncontroversially be applied to multilingualism. As De Bot and Jaensch (2015) point out, researchers have

often ignored the distinction between bi- and multilingualism, which makes it likely that some of the findings reviewed here are, in fact, from multilinguals.

In a line of research most notably pursued by Ellen Bialystok and colleagues (see Bialystok, 2009, for an extensive review), the last 25 years have seen an increase in studies highlighting the benefits of bilingualism for domain-general cognition and, in some aspects, for linguistic development more generally. Building on much previous and concurrent work on the effects of experience on neurocognitive plasticity, the main argument that has been put forward is that bilingualism entails a constant exercise in cognitive control that boosts related mechanisms well beyond the linguistic domain. The necessity of selecting only the language that is relevant in a given situation—a demand that changes frequently, and often unexpectedly—enhances the ability of multilinguals to direct attention to a certain stimulus or set of stimuli, while suppressing activation from other sensory or cognitive input that is irrelevant to the task at hand (cf. Duñabeitia et al., 2014). This routine entails a larger processing load, which may explain why bilinguals are found at a disadvantage in high-demand tasks such as speech perception in noisy and reverberant conditions (Rogers et al., 2006; see also Tabri et al., 2011, for similar results in trilinguals).

Parallel to proposals of the so-called ‘bilingual advantage’ in executive functioning, a number of linguistic disadvantages have also been reported, elaborating on the observation highlighted at the beginning of this section: bilinguals are generally slower

at retrieving words. Most of the work conducted to specifically address this issue has focused on speech production (see Sandoval et al., 2010, for a review). When compared to monolinguals, bilinguals have been shown to manage generally smaller vocabulary sizes in both comprehension and production. Portocarrero et al. (2007) tested monolingual and bilingual college students in a number of tasks measuring vocabulary size and verbal fluency in English. They found that, on average, bilingual students had smaller vocabularies than their monolingual peers, and that this effect was mildly correlated with age of arrival (which, in their study, practically equals age of acquisition): the younger the bilingual participants had entered the country, the larger and more diverse their vocabulary size was.

Besides sheer vocabulary size, experimental results in three more measures have generally shown bilinguals at a disadvantage: tip-of-the-tongue (TOT) state rates, verbal (semantic and letter) fluency and picture naming latencies. In TOT elicitation studies (e.g., Ecker, 2004; Gollan and Acenas, 2004), bilinguals have displayed higher TOT state rates than monolingual or quasi-monolingual peers. As regards verbal fluency, bilinguals have been shown to produce fewer exemplars than monolinguals, both in younger (Gollan et al., 2002) and (healthy) older adults (Rosselli et al., 2000). Data from picture-naming studies are particularly informative. In a study examining picture naming and picture classification in English, Gollan et al. (2005) observed that bilinguals presented inflated naming latencies relative to monolinguals, even though



English was their dominant language. This result was not replicated in the classification task (in which participants had to classify pictures as natural or human made), where both groups displayed comparable response times. Ivanova and Costa (2008) compared Spanish monolinguals with highly proficient Spanish-Catalan and Catalan-Spanish bilinguals in a picture-naming task in Spanish. They found that monolinguals were faster than Spanish-dominant bilinguals, who in turn were faster than their Catalan-dominant peers. This effect was maintained through several repetitions of the stimuli, but was significantly smaller for high-frequency words. Ivanova and Costa (2008) was one of the first studies to conflate the L1 factor in terms of dominance and order of acquisition, so that Spanish- and Catalan-dominant bilinguals had acquired their dominant language first. Previous studies, such as Gollan et al. (2005), examined bilinguals with switched dominance. Taken together, the results of Ivanova and Costa (2008) and Gollan et al. (2005) suggest that the differences between bilinguals and monolinguals are not located at the level of conceptual representation, but rather emerge within the lexical domain.

The question remains as to what exactly causes the generally slower lexical retrieval observed in bilinguals. At least two possibilities have been proposed: i) that lexical access is delayed by activation of competing words from the other language (e.g., *cow* would interfere with and delay the retrieval of *vaca* when Spanish-English bilinguals name a picture of a cow in Spanish); and ii) that the relative frequency of words in the

bilingual lexicon is lower as compared to the same lexical items in a monolingual lexicon. Proponents of the latter, known as the Weaker Links Hypothesis (WLH; Gollan et al., 2008), claim that it can simultaneously account for the bilingual disadvantage in lexical access across all ranges of proficiency. Since even high-frequency words have weaker representations in the mental lexicon of novice L2 learners, the effect can only be attenuated to a certain extent by increased proficiency, because even highly proficient bilinguals use their languages in a smaller percentage than monolinguals. By contrast, the first account is based on a language-nonselective view of lexical access, which assumes that all languages of a bi- or multilingual are constantly activated, with their words competing for selection. The so-called *cognate effect*, whereby multilinguals recognise words more easily if their semantic equivalents in the non-relevant languages are also formally similar, has provided some support for this line of research (e.g., Lemhöfer et al., 2004). Both the competition account and the WLH highlight the same problem: for most cases of lexical retrieval, bilinguals are at a disadvantage, be it because the word has a weaker representation (or weaker conceptual-lexical links) or because it has to compete cross-linguistically with other potential candidates.

In a seminal study on multilingualism, Mägiste (1979) reported significantly longer response times for bilinguals and trilinguals as compared to monolinguals, when tested in both the dominant and the non-dominant languages. These effects emerged in

comprehension and production, although they were more noticeable in the latter. Mägiste proposed that a combination of both accounts, that is, relatively lower activation and competition among languages, is the most plausible explanation for the longer latencies. In any case, the studies reviewed in this section suggest that a larger number of spoken languages correlates with slower lexical retrieval.

### **Research hypotheses**

Relational structures play an important role in the processing of compounds by native speakers, particularly at the stage of constituent integration, where meaning composition is attempted and checked against potentially available stored meanings for the compound. It is possible that nonnative representations of compound words, which we take to be morphemically structured (e.g, González Alonso, Baquero Castellanos & Müller, submitted), also contain information regarding the thematic relations established between the constituents. If they do, then recency or frequency of use of comparable relational structures may enhance the acquisition of this kind of information (i.e., its incorporation into the compound's lexical entry) and its use during processing. This influence will most likely interact with dominance—which will in turn determine frequency of use—and proficiency, and may be obscured by factors related to a non-facilitative effect of sequential multilingualism on lexical retrieval.

To explore these questions, we designed a visual lexical decision task with English NV-er compounds (e.g., *cheerleader*) and pseudo-affixed non-words (e.g., *\*clidmancer*) as critical stimuli. We compared the accuracy and response latencies of four different groups: (monolingual) English native speakers, Spanish-English sequential bilinguals (L2 learners), and two groups of early Basque-Spanish bilinguals (L3 learners of English), who were either Basque- or Spanish-dominant. Participants from each of the learner groups were fairly evenly distributed along the proficiency continuum, from elementary to very advanced speakers. We expected to find differences both between native and nonnative speakers (with the former outperforming the later) and among the three nonnative groups. In particular, we expected bilinguals to provide faster responses than trilinguals (i.e., speed of lexical retrieval as an inverse function of the number of languages spoken), and, within these, Basque-dominant trilinguals to be faster than Spanish-dominant ones, as a consequence of the shared relational structure between Basque and English NV-er compounds.

Pseudo-affixed nonword stimuli (hence *pseudocompounds*) may present a particular challenge in that the saliency of the {-er} suffix will prompt an unsuccessful analysis of the word on the basis of the NV-er structure. This is expected to affect all participants, to the extent that the target English relational structure has been acquired and is thus activated during processing. However, any groups who are exposed to this structure in their first and dominant language are likely to be more vulnerable to the effect. This

includes English native speakers and, especially, Basque-dominant trilinguals, who have it in two of their three languages.

## **The study**

### *Participants*

A total of 141 participants (110 females, 31 males) took part in the experiment. They were on average 19 years of age (range = 18–40; median = 20). The experimental groups comprised the following four linguistic profiles. All participants from the non-native groups (1 to 3) had studied English as a foreign language at school, since at least age 10.

- (1) Spanish-English bilinguals, or L1 Spanish–L2 English ( $n = 38$ , 24 females; mean age = 20, range = 18–32): Born and raised in a Spanish-speaking household, with little or no notion of Basque.
- (2) Basque-dominant trilinguals, or L1 Basque–L2 Spanish–L3 English ( $n = 44$ , 37 females; mean age = 20, range = 18–34): Born and raised in Basque-speaking households, not formally exposed to Spanish until primary school. Basque remained the predominant language at home and in school, in most cases including university.<sup>4</sup>

- (3) Spanish-dominant trilinguals, or L1 Spanish–L2 Basque–L3 English ( $n = 44$ , 35 females; mean age = 23, range = 18–40): Born and raised in Spanish-speaking households of the Basque Country. All pre-university schooling, since age 2-3, completed in Basque. These participants had also continued to study in Basque at university and/or had obtained an advanced language certificate in Basque (C1 or C2 level in the Common European Framework of Reference for Languages). Spanish remains the only language spoken at home, and the most frequently used overall.
- (4) English native speakers ( $n = 15$ ; 13 females; mean age = 19, range: 18–23): Born and raised in the USA, these participants had little or no command of a second language. They were recruited among the undergraduate and graduate student community at the University of Florida.

No participants reported any known or diagnosed reading or learning disabilities, nor any history of brain lesion or impairment. All had normal (or corrected-to-normal) vision, and were completely naive with regard to the purpose of the experiment.

### *Materials*

There were 84 critical experimental items, consisting of 42 high-frequency<sup>5</sup> English NV-er compounds (e.g., *taxi driver*) and 42 pseudocompounds (non-words generated by substituting a number of characters in the corresponding compound) retaining the suffix

{-er}. Frequency counts for the compounds were performed on the Corpus of Contemporary American English (Davies, 2008; 450 million words). The most frequent orthographic form—solid ( $n = 29$ ), hyphenated ( $n = 1$ ) or fractured ( $n = 12$ )—was used to maximise the chances for recognition. These 42 items had a mean frequency of 5.41 occurrences per million (range: 1.00–26.66, median = 3.32). The list of pseudo-compounds was designed through a process of semi-random character substitution, using *Wuggy Pseudoword Generator* (Keuleers and Brysbaert, 2010). The output of this process were pronounceable and plausible non-words consistent with English phonotactics. The mean formal overlap between compounds and pseudocompounds was 60.95% (SOLAR system; Davis, 1999, 2006).

The list of stimuli was completed with the inclusion of 168 fillers, divided into 84 non-compound English words (further sub-divided into 21 nouns, 21 adjectives, 21 verbs and 21 adverbs) and 84 pseudowords, generated through the same procedure as above. Since compound words are naturally less frequent, performing a frequency-match between compound and non-compound words would have rendered an all-round low frequency stimuli list, unsuitable for lower proficiency participants.<sup>6</sup> Word fillers and their corresponding pseudowords (Nonword fillers) had a mean overlap of 53.15%.

**TABLE 1.** Descriptive statistics for compound and non-compound stimuli.

STIMULI TYPE	LENGTH			FREQUENCY		
	Mean	SD	Range	Mean	SD	Range
Compound	10.69	1.62	8 – 15	5.41	5.75	1 – 26.66
Pseudocompound	10.69	1.62	8 – 15	NA	NA	NA
Word fillers	4.93	1.68	3 – 10	545.63	348.59	125.02 – 1744.64
Nonword fillers	4.93	1.68	5 – 7	NA	NA	NA

*Note:* Length is measured in characters. Frequencies are measured in occurrences per million.

### *Procedure*

The experiment was conducted individually, and in three stages. First, participants were asked to complete the *Gogo Elebiduna taldearen Galdetegia* questionnaire (Gogo Elebiduna Research Group, 2010), designed to collect data about their language knowledge, exposure and use. This test allowed us to determine their dominance—in trilingual groups—and general language history, including any additional languages they might have relevant knowledge of. Proficiency in Basque and Spanish was self-reported, and assumed to correlate with the participants' language history. Upon completion, participants were tested for language proficiency in English, using the standardised *Oxford Quick Placement* pen-and-paper test (OQPT; Oxford University



Press et al., 2001). Participants were given 30 minutes to complete this test. Their score was measured over a total of 60 points. Participants took the test individually, in a well-lit, soundproof booth.

A computer task was next administered using Neurobehavioral Systems' Presentation™, and monitored from a contiguous room. The software first displayed a set of instructions, followed by 10 practice trials. The overall structure of the task was as follows: the total 252 stimuli were divided in six blocks of 42 items. Each of these blocks, randomly sorted for every participant, contained seven compounds, seven pseudocompounds, 14 word fillers and 14 non-word fillers. These items were also randomly sorted per participant. Pauses between blocks were self-paced. Stimuli were presented in 30 point, lowercase white Cambria font on a black background. For every trial, an asterisk first appeared for 400 ms at the centre of the screen as a fixation point. Next, the test item was displayed centred at the same point and remained on screen until an answer was provided through the response buttons. The initial instructions purposefully avoided the term 'word', to prevent negative judgements for compounds presented in fractured form. Participants were instructed to '...decide whether what you see is real English or not' and respond accordingly, as fast and as accurately as possible, by pressing the button YES or NO on a computer keyboard.<sup>7</sup> The next trial began upon response. No feedback was provided during or after the experiment.

## Results

### *Data analysis*

The total collected data comprised the following variables as predictors: Subject (141 participants), Age (mean = 21:2,  $SD = 4:11$  years), Gender (31 males), Group (Basque-dominant trilinguals | BA; Spanish-dominant trilinguals | ES; Spanish-English bilinguals | ES-bi; English monolinguals | EN), Condition (compounds, pseudo-compounds, word fillers, nonword fillers), Proficiency (English proficiency score: 0–60), Trial (252 repetitions) and Item (252 words). Dependent variables were RT (response time in ms) and Accuracy (evaluation of the participant's response: success, failure). In 43 out of the total 35,532 records (0.12%), RT and Accuracy data were missing due to a glitch in the presentation of the last trial for some participants. No data imputation was conducted, and when necessary non-available (NA) data were removed (e.g., when computing means).

Response times outside  $\pm 2.5$  of a participant's median absolute deviation (MAD) were considered outliers. This exclusion criterion was found to be more appropriate than the absolute time thresholds commonly found in the literature. By imposing a 200ms–3s range, for example, we could have been misrepresenting genuinely fast and slow participants. Note also that using the standard deviation as a criterion for finding outliers (e.g., discarding observations outside  $\pm 2 SD$ ) is not suitable in this case as

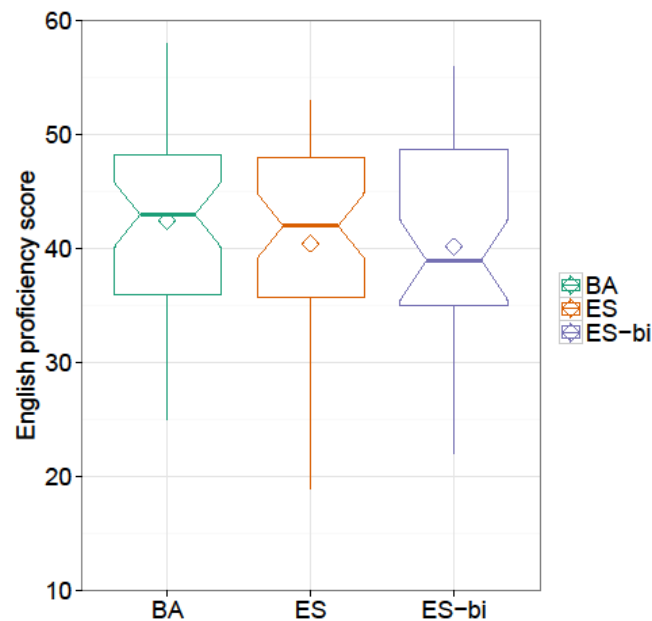
response times are not normally distributed (Leys et al., 2013). Admittedly, the selection of the number of MADs is subjective, but we followed the recommendations of Leys et al. (2013) as to have a ‘moderately conservative’ exclusion criterion. No data points were excluded for being too fast; 717 responses were discarded for being too slow. The total number of outliers excluded with this procedure amounted to 2.14% of the full dataset.

Our response time analyses included the latencies of correct as well as erroneous responses. As put by Salthouse and Hedden (2002), “[m]erely because the overt errors are eliminated from the analyses does not mean that the remaining RTs reflect the true duration of the relevant processing, because the guesses that were correct by chance are still represented in the data.” Their claim resonates with the findings of Diependaele et al. (2012), who found that lexical decision tasks have a certain degree of noise, that is, presented with the same item, subjects are not always consistent with their previous answers. They also noted that word frequencies below 10 occurrences per million (as in some stimuli in the current experiment) may yield consistency rates near chance. Taken together, these findings call into question the somewhat default assumption that all correct guesses reflect recognition of the stimuli.

Response times were transformed by taking their reciprocal and scaling it by  $-1000$  as a way of minimising the (positive) skewness in the RT distribution (see, e.g., Kliegl et al., 2010). The transformed times were accordingly labelled ‘Speed’ and used as the

dependent variable in the RT analyses. Since long answers are usually associated to incorrect responses, we included Accuracy as a factor in our response time analyses.

Among nonnative speakers, the average English proficiency score of Basque-dominant trilinguals was numerically higher than the other groups (see Figure 3): Spanish-English bilinguals, mean = 40, range = 22-56; Basque-dominant trilinguals, mean = 42, range = 25-58; Spanish-dominant trilinguals, mean = 40, range = 19-53. This difference, however, was not statistically significant,  $F(2, 123) = 0.907$ ,  $p = .406$ . For subsequent analyses, Proficiency scores and the Trial variable were centred around zero.



**FIGURE 3.** English proficiency scores for nonnative groups. Bold horizontal lines and diamonds show the median and mean score for each group, respectively. The whiskers extend to  $\pm 1.5$  the inter-quartile range. (BA = Basque-dominant trilinguals; ES = Spanish-dominant trilinguals; Es-bi = Spanish-English bilinguals).

### *Response times*

The data summarized in Table 2 show that native English speakers were the fastest group on average for all conditions, followed by Spanish-English bilinguals. The results of the two trilingual groups are almost identical, with a small numerical lead of Spanish-dominant trilinguals in all conditions except Compounds, where the Basque-dominants' mean response latency was 10 ms shorter. With the notable exception of native speakers, similar trends were found across all groups for the different conditions: responses to pseudocompounds were the slowest, followed by compounds, nonword fillers and, fastest, word fillers. This suggests that, for our nonnative speakers, there is a prominence of word length over lexicality, with shorter items being responded to faster, irrespective of whether they were known to them or not.<sup>8</sup> Subsequent analyses focus on the nonnative data only.

**TABLE 2.** Mean response times (in milliseconds) and Accuracy percentages for each participant group in the Compound, Pseudocompound and filler conditions.

GROUP	CONDITION	RESPONSE TIME (ms)	ACCURACY (%)
EN	Compounds	669	98.4
	Pseudocompounds	927	94.1
	Word fillers	550	99.4
	Nonword fillers	801	81.6
ES-bi	Compounds	995	80.8
	Pseudocompounds	1140	87.6
	Word fillers	610	99.3
	Nonword fillers	895	76.8
BA	Compounds	1090	87.2
	Pseudocompounds	1360	84.1
	Word fillers	643	99.6
	Nonword fillers	1020	74.1
ES	Compounds	1100	83.0
	Pseudocompounds	1350	87.8
	Word fillers	639	99.5
	Nonword fillers	979	76.1

A correlation between response times of previous and current trials was observed for some subjects, especially in immediately successive trials (see, e.g., Kuperman et al., 2009). For this reason, we included the response time of the previously presented trial as a predictor (labelled ‘prevRT’ in (6) below). The data were analysed through a series of linear mixed models, beginning with the model presented in (5) and moving towards more complex structures by means of a stepwise forward selection procedure.

Computation of these models was done using the *lmer* routine in the *lme4* package (Bates et al., 2013), within the *R* software (R Core Team, 2013).

$$(5) \quad \text{Speed} \sim \text{Condition} + (1|\text{Subject}) + (1|\text{Item}) + (1|\text{Age}) + (1|\text{Gender}) + (1|\text{Accuracy}) + (1|\text{prevRT})$$

The equation in (5) specifies that Speed is explained by the fixed factor Condition, and by Subject, Item, Age, Gender and Accuracy as random factors. More complex models were created by adding a single fixed factor at a time and performing a likelihood ratio test between subsequent models. If no significant difference between the two models was found ( $\alpha = .95$ ), the most parsimonious model was kept. When all theoretically justified factors were included, we proceeded in the same manner with interactions between fixed factors. When it was not possible to add more fixed factors (or their interactions) to the model, we simplified it when possible by subtracting apparently irrelevant random effects (starting from those with lowest variance) via model comparisons, as previously, until we arrived at the maximal random structure justified by the design (Barr et al., 2013). The resulting model is presented in (6) below (colons indicate interactions in *R* notation).

$$(6) \quad \text{Speed} \sim \text{Group} + \text{Condition} + \text{Group}:\text{Condition} + \text{Trial} + (1|\text{Subject}) + (1|\text{Item}) + (1|\text{Accuracy}) + (1|\text{prevRT})$$

In this and following analyses, visual inspection of residuals was used to verify the model's assumptions; unless otherwise stated, no obvious deviations from normality or homogeneity of variance were found. While English proficiency was not a determining factor,  $\chi^2(1) = 3.4$ ,  $p = .065$ , both Group,  $\chi^2(2) = 6.8$ ,  $p < .05$ , and its interaction with Condition,  $\chi^2(6) = 49.9$ ,  $p < .001$ , had a significant effect on Speed. A further post-hoc analysis based on multiple comparisons of means with Tukey contrasts revealed that these significant differences did not lie between the trilingual groups ( $p = .948$ ) or between the Spanish-dominant groups ( $p = .090$ ), but rather between Spanish-English bilinguals and Basque-dominant trilinguals ( $p < .05$ ). Specifically, bilinguals were significantly faster answering in the Nonword filler ( $p < .05$ ) and Pseudocompound ( $p < .05$ ) conditions.

#### *Accuracy*

The Accuracy analysis was divided into two parts. First, we examined the differences in overall accuracy (i.e., the percentage of successful classifications of items as words or non-words) between the groups. A close inspection of the patterns that emerged in this analysis suggested that there were between-group differences at the response level that went beyond pure accuracy. For this reason, we computed discriminability and bias measures for each participant, which allowed us to repeat our analyses with new information on the participants' tendency to a particular response—either in general



terms or as a function of the different factors under investigation. We will begin by looking at the general Accuracy analysis.

*Accuracy.* The Accuracy column in Table 2 indicates that native speakers were better than the rest of groups at classifying items in all conditions, with the exception of word fillers, where a ceiling effect is likely to obscure proficiency-based differences. Among the nonnative speaker groups, results displayed complex patterns: Basque-dominant trilinguals were better than the other two groups at classifying compounds, but worse at pseudocompounds. The opposite holds true for Spanish-dominant trilinguals and Spanish-English bilinguals, with a small numerical lead of the trilinguals. The same trend was observed in filler items, with the L1-Spanish groups performing better at non-words. These data were analyzed in a similar fashion to response times. However, since accuracy results are binary, we fit a logit linear mixed effects model with a Bernoulli error distribution. The original model fit was as follows:

$$(7) \quad \text{Accuracy} \sim \text{Condition} + (1|\text{Subject}) + (1|\text{Item})$$

After several iterations of model building, we obtained the following fit:

$$(8) \quad \text{Accuracy} \sim \text{Condition} + \text{Proficiency} + (1|\text{Subject}) + (1|\text{Item})$$

The Group factor did not display a significant main effect,  $\chi^2(2) = 0.85, p = .650$ . The fact that proficiency significantly affected the accuracy but not the speed of responses

(where it had only a very marginal effect, see above) is relatively surprising, and a likely indicator that some sort of bias might be affecting the performance of some of the groups: lower-proficiency participants made more errors, and incorrect responses tend to take longer. If the significantly higher number of erroneous responses did not produce, as a result, a significant difference in RTs, then something might be making some of these participants miss faster than they would under normal circumstances. Note that this might stem both from a liberal bias (because it would make them over-accept) or a conservative one (because they would over-reject). Indeed, although the comparable performance in accuracy among learner groups matched our predictions, we noted that two of the groups were more accurate at classifying pseudocompounds (a nonword category) than compounds (a word category). This was interpreted as an indication that these participants may have a tendency towards negative responses—which, in an extreme case, would entail a perfect score on nonword items and a zero score on word items. These two groups shared their first and dominant language, one of the central variables in our study.

*Sensitivity.* In order to compute the appropriate rates, we grouped the Compound and Pseudocompound conditions into ‘Compound-like’ and the filler conditions into ‘Noncompound-like’ levels of a factor named ‘Class’. This grouping allowed us to determine, for each participant, hit and miss rates, and false alarm and correct rejection rates. As suggested by Snodgrass and Corwin (1988), to compute the hit and false alarm

rates we added 0.5 to the hit and false alarm frequencies, and divided them by one plus the total number words, non-words, respectively. This procedure is necessary to treat hit rates equal to unity and false alarms equal to zero, for which their  $Z$ -scores are infinite.

Non-parametric discriminability ( $A'$ ) and response bias ( $B''_D$ ) were computed in  $R$  following the procedure suggested in Pallier (2002). Note that  $A'$  and  $B''_D$  are estimates of sensitivity in the same fashion as  $d'$  and  $\beta$ . In this implementation,  $A'$  values near unity indicate good discriminability, while values near 0.5 indicate chance performance. Negative, equal to zero, and positive  $B''_D$  values indicate a liberal bias (i.e., a tendency to respond ‘Yes, [this item is correct in English]’), no bias, and a conservative bias (i.e., a tendency to respond ‘No, [this item is not correct in English]’), respectively. Discriminability and response bias results are summarised in Table 3.

**TABLE 3.** Nonparametric discriminability power ( $A'$ ), response bias ( $B''_D$ ), and hit and false alarm (FA) rates in the nonnative groups.

CLASS	GROUP	$A'$	$B''_D$	Hit rates	FA rates
Compound-like	ES-bi	0.908	0.226	0.806	0.128
	BA	0.915	-0.000	0.867	0.163
	ES	0.913	0.216	0.824	0.125
Noncompound-like	ES-bi	0.937	-0.813	0.988	0.235
	BA	0.931	-0.924	0.990	0.259
	ES	0.936	-0.912	0.989	0.240

At least two facts in this table help interpret our overall Accuracy results. The first is that all groups display a strong liberal bias for non-compound items (i.e., fillers), which simultaneously explains their native-like performance in the word filler condition while missing around 25% of responses to nonword fillers. The second is that, while Basque-dominant participants show almost zero bias in the Compound-like class, their Spanish-dominant peers are moderately biased towards conservation. These differences in  $A'$  and  $B''_D$  between groups, along with potential effects of other measures (e.g., Proficiency), were tested through linear mixed models as in previous analyses. For discriminability, the resulting model was:

$$(9) \quad A' \sim \text{Class} + \text{Proficiency} + (1|\text{Subject})$$

For response bias, a crucial variable changed in the fixed effects structure of the final model:

$$(10) \quad B''_D \sim \text{Class} + \text{Group} + (1|\text{Subject})$$

The absence of Group as a predictor in (9) indicates that this factor does not influence their discriminability power,  $\chi^2(2) = 0.22$ ,  $p = .900$ , but as expected, English proficiency does,  $\chi^2(1) = 21.6$ ,  $p < .001$ . On the other hand, a post-hoc analysis on the bias model indicates that, when confronted with noncompound-like stimuli, Spanish-dominant groups had a liberal bias. This contrasts with the conservative bias observed

in their responses to compound-like items ( $p < .001$ ). Basque-dominant trilinguals, while being the most liberal of the three groups in the noncompound-like class, displayed practically no bias in compounds and pseudocompounds. The bias differences between bilinguals and Basque-dominant trilinguals were significant ( $p < .05$ ), but none emerged between the trilingual groups ( $p = .278$ ) or between the Spanish dominant groups ( $p = .510$ ). Although enlightening, these patterns add to the complexity of the RT and Accuracy results.

## **General discussion**

The experiment presented in this article consisted of a lexical decision task in which the critical items were English NV-er synthetic compounds, a class which has an almost identical (structural) correlate in Basque but is missing as such in Spanish, where similar verb-object relations are expressed through a different compound pattern. We hypothesised that, if sequential multilingualism and the language dominance factor play a significant role in morpho-lexical processing, we would be able to see differences between native and nonnative speakers, and among our three nonnative groups: Spanish-English bilinguals and two trilingual Basque-Spanish-English groups, who were either Spanish- or Basque-dominant. In particular, we predicted speed of lexical retrieval to be inversely proportional to the number of languages spoken, and an advantage of Basque-dominant trilinguals over the Spanish-dominant ones by virtue of

the shared compound structure between English and Basque. Let us now use these predictions as a guide for discussion.

### *Native speaker performance*

As expected, native speakers outperformed all learner groups in Accuracy and RT measures. This difference, however, was not as marked in certain conditions (e.g., word fillers). Two factors may be separately or jointly responsible for this. The first is that the overall high frequency of word filler stimuli made them easily recognisable, creating a ceiling effect. The second is that this might be due to an extremely liberal bias in non-compound stimuli on the part of nonnative speakers. This bias would have taken them to incorrectly accept many non-words, but also to correctly classify word fillers as words. If this second account is true, an indeterminate number of these hits will have been correct by chance.<sup>9</sup>

### *The multilingual disadvantage*

Our second prediction concerned differences between bilingual and trilingual groups. These are not as clear-cut as those between native and nonnative speakers. However, the performance of these groups does seem to stand along a continuum: Basque-dominant trilinguals and bilinguals at the extremes, and Spanish-dominant trilinguals somewhere in between (closer to the Basque-dominant group, especially in response times). This becomes apparent from the average response latencies of the groups, both overall (ES-

bi: 910 ms; ES: 1017 ms; BA: 1028 ms) and across conditions. The small disadvantage of the Basque-dominant group is hardly attributable to proficiency, since they were on average more proficient than the other two groups (recall Figure 3). Group means seem to suggest some kind of ‘multilingual disadvantage’ in our data, but the non-significance of some of these differences preclude strong conclusions. Table 4 shows *p* values for multiple comparison post-hoc tests performed on the response time data of the three learner groups.

**TABLE 4.** P-values for Tukey HSD post-hoc tests on the RT data, performed by level of Condition among the nonnative groups. Significant values at  $\alpha = .95$  and at  $\alpha = .90$  are highlighted in grey and light grey, respectively.

		ES	BA
ES-bi	Compounds	.450	.459
	Pseudocompounds	.066	.035
	Word fillers	.734	.489
	Nonword fillers	.450	.033
BA	Compounds	1.00	
	Pseudocompounds	1.00	
	Word fillers	1.00	
	Nonword fillers	.935	

Spanish-English bilinguals were significantly faster than Basque-dominant trilinguals on the two nonword conditions, but there were no differences in the Compound and

Word filler conditions. With respect to the Spanish-dominant trilinguals, the difference between them and bilinguals was only marginally significant in the Pseudocompound condition. Two possibilities may explain this pattern. One is that the learner groups are not numerous enough to show a significant difference. While this might be true for the difference between the two L1-Spanish groups in the Pseudocompound condition, other comparisons display  $p$  values far from a conventional significance threshold ( $p \leq .05$ ).

A second possibility is that the conservative bias of L1-Spanish groups results in an alignment of their response times, pulling the trilingual group away from their Basque-dominant peers, who do show a difference with bilinguals. This does not seem fully consistent with our results. The best candidate factor to explain the conservative response bias of L1-Spanish groups seems to be word length: Basque's rich agglutinative morphology entails a higher number of long words than those found in Spanish or English. Being a salient perceptual feature, length should have an early effect on response times, affecting both conditions of the concerned class (compound-like) equally. This was not the case: differences between bilinguals and both groups of trilinguals are virtually the same in the Compound condition, whereas responses to pseudocompounds, where the advantage of bilinguals over Basque-dominant trilinguals is most apparent, yielded an only marginally significant difference between the L1-Spanish groups. This conditional asymmetry in the distance between L1-Spanish groups



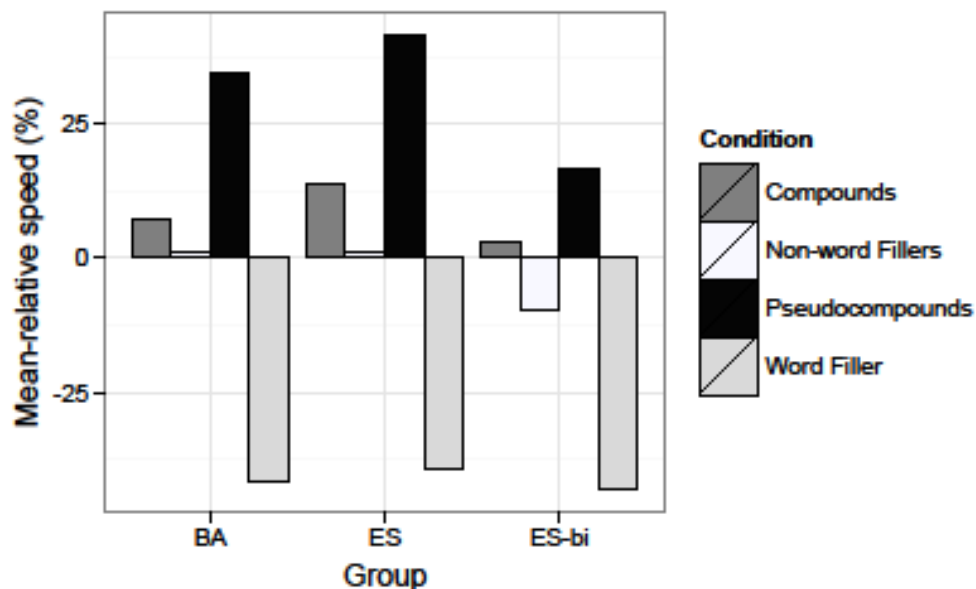
suggests that, while their shared conservative bias may be grounded on L1-specific factors, it is not the sole responsible for the graded pattern of RTs in our nonnative data.

### *The dominance factor*

Our third and most complex prediction was that Basque-dominant trilinguals would display an advantage over their Spanish-dominant peers in the Compound condition, which contained a familiar relational structure. Such an advantage did not emerge in standard RT and Accuracy analyses. However, we observed that L1-Spanish participants were better at rejecting pseudo-compounds than they were at accepting compounds. This is uncommon because non-words are typically associated with higher error rates. An analysis of response bias showed that L1-Spanish participants had a moderate tendency towards rejecting compound-like words, unlike the Basque-dominant trilinguals who displayed virtually no bias.

Any attempts at a successful account of this finding should take into consideration the common features of the Compound-like class that might have triggered the bias once participants had noticed them. Furthermore, whatever aspect we consider as a candidate should also be different between Spanish, on the one hand, and Basque and English on the other, since the differences in response bias seem to be linked to this grouping variable in our data. Given these two conditions, two language-specific features provide reasonable explanations of the effect: average word length and the NV-

er relational structure. The first would have (very) early effects on the L1-Spanish groups' responses to the compound-like conditions, while the second would most likely affect the composition stage (e.g., Gagné and Spalding, 2009; Ji et al., 2011). Figure 4 provides some useful information to evaluate the first account.



**FIGURE 4.** RTs (expressed as percentages) for each Group and Condition relative to each participant's overall mean response times.  $y = 0$  represents the mean for all conditions; negative values indicate faster-than-average responses; positive values indicate slower-than-average responses.

If word length, perceptually a very salient feature, is responsible for the bias in the L1-Spanish groups, we should expect relatively shorter latencies (i.e., a smaller deviation from the mean) in the compound-like conditions. This should be true even

more for pseudocompounds, where an initial morpho-orthographic analysis (e.g., Rastle and Davis, 2008) will not, in most cases, yield any recognisable lexemes. However, Figure 4 does not show an alignment of L1-Spanish groups in these conditions. To examine this, we fit a linear mixed model with the mean-relative RTs as the dependent variable, Group, Condition and their interaction as fixed factors, and Subjects and Items as crossed random factors. Post-hoc analyses on the interaction (Tukey contrasts) revealed no significant differences between the two trilingual groups (ES – BA, compounds:  $z = 2.23$ ,  $p = .38$ ; pseudocompounds:  $z = 1.97$ ,  $p = .551$ ). In fact, while bilinguals showed a difference with respect to both trilingual groups in the Pseudocompound condition (ES-bi – ES:  $z = -7.24$ ,  $p < .01$ ; ES-bi – BA:  $z = -5.34$ ,  $p < .01$ ), it was with the Spanish-dominant group that a significant difference emerged in the Compound condition (ES-bi – ES:  $z = -4.02$ ,  $p < .01$ ; ES-bi – BA:  $z = -1.87$ ,  $p = .618$ ). These results suggest that, whatever the nature of the conservative bias in the L1-Spanish groups, we are not seeing early effects on response times. It is difficult to determine to what extent these data undermine the word length account, but a second possible factor should be explored: the NV-er compound structure.

Our trilingual participants were early bilinguals, but they were neither simultaneous nor balanced: Basque and Spanish, respectively, were both their first and dominant languages. This entails that, despite being native or native-like in both, the frequency of use of each language is the most notable between-group difference. We selected

participants with clear-cut profiles to increase the detectability of any potential effects, and avoid confounding factors like switched dominance (see Ivanova and Costa, 2008, for discussion). Frequency of use could result in cross-linguistic influence of morphemic and relational structures which are common in the dominant language. The absence of a response bias in Basque-dominant trilinguals might be a reflex of this effect. The saliency of the {-er} suffix, also present in the Pseudocompound condition, would trigger a composition process in terms of the NV-er relational structure. Basque-dominant participants, for whom the structure has a purportedly higher resting activation level, would proceed to an evaluation of the compound (or pseudocompound) by attempting to fit this structure with the morphemic breakdown output by the morphological decomposition process. In contrast, Spanish-dominant speakers would already acquire a bias against these items, which they would have more problems interpreting correctly, because the intended thematic structure is most often mapped to different surface configurations in the largest part of their everyday language experience.

## **Conclusions**

The experiment presented in this article investigated the use of relational structures in nonnative compound processing, as well as the possible interaction between the hampering effect of sequential multilingualism—as a plausible extension of the ‘bilingual disadvantage’—and the potential cross-linguistic priming of these structures.

A group of native speakers of English and three groups of learners (Spanish-English bilinguals, Basque-Spanish-English and Spanish-Basque-English trilinguals), all at different levels of proficiency, performed a lexical decision task which measured their response to compound and non-compound words, with a particular focus on English synthetic NV-er compounds.

While native and nonnative speakers significantly differed in accuracy and average response times, differences among the three nonnative groups display more complex patterns. Most importantly, the two trilingual groups were slower and generally more similar between them than they were to the bilingual group, despite the fact that bilinguals and Spanish-dominant trilinguals shared their first and dominant language. A notable difference did emerge as a function of this factor: both L1-Spanish groups showed a moderate conservative bias in the compound-like conditions of the experiment. Two possible accounts were put forward to explain this effect. The first is that Basque-dominant speakers, more used to processing long words, are less prone to rejecting them on sight. The second account relies on the absence of comparable surface structures in Spanish. We have speculated that this may predispose Spanish-dominant learners to reject these items if the relational structure is not initially recognised, even if the compound's constituents can be individually parsed. Measures with higher temporal resolution might provide the necessary information to adjudicate between both accounts.

Our data suggest that there are costs in language processing, at least with regard to lexical retrieval, associated with a larger number of languages spoken. Relational structures within compound lexical representations are likely to develop later than basic morphemic structure as part of nonnative morpho-lexical knowledge. For that reason, they may not be as readily available as the former during processing. We have observed, however, that morphological properties of the most frequently active language may have an influence on the processing of other languages. Further research should be conducted to determine what level of specification is present in the morphological structure of nonnative lexical representations, and to what extent this knowledge may impact on the parsing of words in a different language.

## References

- Anderson SR (1985) Typological distinctions in word formation. In: Shopen T (ed.), *Language Typology and Syntactic Description (vol. 3)*, Cambridge: Cambridge University Press, pp. 3–56.
- Aronin L and Singleton D (2012) *Multilingualism. Impact* (Amsterdam. 1997), Amsterdam: John Benjamins.
- Barr DJ, Levy R, Scheepers C, et al. (2013) Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of memory and language*, Elsevier Inc., 68(3), 255–278.
- Bates D, Maechler M, Bolker B, et al. (2013) lme4: Linear mixed-effects models using Eigen and S4. Vienna, Austria: R Foundation for Statistical Computing.
- Bauer L and Renouf A (2001) A Corpus-Based Study of Compounding in English. *Journal of English Linguistics*, 29(2), 101–123.
- Bialystok E (2009) Bilingualism: The good, the bad, and the indifferent. *Bilingualism: Language and Cognition*, 12(1), 3–11.
- Booij G (2005) *The Grammar of Words: An Introduction to Morphology*. Oxford: Oxford University Press.
- Bybee J (1995) Regular morphology and the lexicon. *Language and Cognitive Processes*, 10, 425–455.
- Cattell JM (1887) Experiments on the association of ideas. *Mind*, 12, 68–74.
- Clahsen H, Rothweiler M, Woest A, et al. (1992) Regular and irregular inflection in the acquisition of German noun plurals. *Cognition*, 45, 225–255.
- Clahsen H, Balkhair L, Schutter J-S, et al. (2013) The time course of morphological processing in a second language. *Second Language Research*, 29(1), 7–31.
- Clark E V., Hecht BF and Mulford RC (1986) Coining complex compounds in English: affixes and word order in acquisition. *Linguistics*, 24(1), 7–29.

- Contreras H (1985) Spanish exocentric compounds. In: Unessel F (ed.), *Current Issues in Hispanic Phonology and Morphology*, Bloomington, IN: Indiana University Linguistics Club, pp. 14–27.
- Davies M (2008) The Corpus of Contemporary American English: 450 million words, 1990-present. Available from: <http://corpus.byu.edu/coca/>.
- Davis CJ (1999) The self-organising lexical acquisition and recognition (SOLAR) model of visual word recognition. University of New South Wales.
- Davis CJ (2006) Orthographic input coding: A review of behavioural data and current models. In: Andrews S (ed.), *From inkmarks to ideas: Current issues in lexical processing*, Psychology Press, pp. 180–206.
- De Bot K and Jaensch C (2015) What is special about L3 processing? *Bilingualism: Language and Cognition*, 18(2), 130–144.
- De Cat C, Klepousniotou E and Baayen RH (2015) Representational deficit or processing effect? An electrophysiological study of noun-noun compound processing by very advanced L2 speakers of English. *Frontiers in Psychology*, 6(77), 1–17.
- Desrochers A, Licerias JM, Spradlin T, et al. (2003) Can an on-line response latency task shed light on native and non-native competence in the deverbal compounds of Spanish? In: Licerias JM, Zobl H, and Goodluck H (eds), *6th Generative Approaches to Second Language Acquisition Conference (GASLA 2002)*, Somerville, MA, pp. 64–70.
- Di Sciullo A-M (1991) On the structure of deverbal compounds. *Working Papers in Linguistics*, 1, 1–26.
- Di Sciullo A-M (1992) Deverbal compounds and the external argument. In: Roca IM (ed.), *Thematic structure: Its role in grammar*, Dordrecht: Foris, pp. 65–78.
- Diependaele K, Duñabeitia JA, Morris J, et al. (2011) Fast morphological effects in first and second language word recognition. *Journal of Memory and Language*, 64(4), 344–358.
- Diependaele K, Brysbaert M and Neri P (2012) How noisy is lexical decision? *Frontiers in Psychology*, 3(348), 1–9.



- Downing P (1977) On the creation and use of English compound nouns. *Language*, 53, 810–842.
- Duñabeitia JA, Laka I, Perea M, et al. (2009) Is Milkman a superhero like Batman? Constituent morphological priming in compound words. *European Journal of Cognitive Psychology*, 21(4), 615–640.
- Duñabeitia JA, Dimitropoulou M, Morris J, et al. (2013) The role of form in morphological priming: Evidence from bilinguals. *Language and Cognitive Processes*, 28(7), 967–987.
- Duñabeitia JA, Hernández JA, Antón E, et al. (2014) The inhibitory advantage in bilingual children revisited: Myth or reality? *Experimental Psychology*, 61(3), 234–251.
- Ecke P (2004) Words on the tip of the tongue: A study of lexical retrieval failures in Spanish-English bilinguals. *Southwest Journal of Linguistics*, 23, 33–63.
- Feldman LB, Kostić A, Basnight-Brown DM, et al. (2010) Morphological facilitation for regular and irregular verb formations in native and non-native speakers: Little evidence for two distinct mechanisms. *Bilingualism: Language and Cognition*, 13(2), 119–135.
- Fiorentino R and Fund-Reznicek E (2009) Masked morphological priming of compound constituents. *The Mental Lexicon*, 4(2), 159–193.
- Fiorentino R and Poeppel D (2007) Compound words and structure in the lexicon. *Language and Cognitive Processes*, 22(7), 953–1000.
- Fiorentino R, Naito-billen Y, Bost J, et al. (2014) Electrophysiological evidence for the morpheme-based combinatoric processing of English compounds. *Cognitive Neuropsychology*, 31(1-2), 123–146.
- Gagné CL (2001) Relation and lexical priming during the interpretation of noun–noun combinations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27, 236–254.
- Gagné CL and Shoben EJ (1997) The influence of thematic relations on the comprehension of modifier-noun combinations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23, 71–87.

- Gagné CL and Spalding TL (2009) Constituent integration during the processing of compound words: Does it involve the use of relational structures? *Journal of Memory and Language*, Elsevier Inc., 60(1), 20–35.
- García Mayo MP (2006) Synthetic Compounding in the English Interlanguage of Basque–Spanish Bilinguals. *International Journal of Multilingualism*, 3(4), 231–257.
- Gogo Elebiduna Research Group (2010) *Gogo Elebiduna taldearen Galdetegia (GEG 2.0.)*.
- Gollan TH and Acenas L-AR (2004) What is a TOT? Cognate and translation effects on tip-of-the-tongue states in Spanish-English and tagalog-English bilinguals. *Journal of experimental psychology: Learning, memory, and cognition*, 30(1), 246–69.
- Gollan TH, Montoya RI and Werner GA (2002) Semantic and letter fluency in Spanish–English bilinguals. *Neuropsychology*, 16, 562–576.
- Gollan TH, Montoya RI, Fennema-Notestine C, et al. (2005) Bilingualism affects picture naming but not picture classification. *Memory & cognition*, 33(7), 1220–34.
- Gollan TH, Montoya RI, Cera C, et al. (2008) More use almost always a means a smaller frequency effect: Aging, bilingualism, and the weaker links hypothesis. *Journal of memory and language*, 58(3), 787–814.
- González Alonso J, Baquero Castellanos S and Müller O (submitted) Masked constituent priming of English compounds in native and nonnative speakers. *Language, Cognition and Neuroscience*.
- Hay JB and Baayen RH (2005) Shifting paradigms: gradient structure in morphology. *Trends in cognitive sciences*, 9(7), 342–348.
- Ivanova I and Costa A (2008) Does bilingualism hamper lexical access in speech production? *Acta psychologica*, 127(2), 277–88.
- Ji H, Gagné CL and Spalding TL (2011) Benefits and costs of lexical decomposition and semantic integration during the processing of transparent and opaque English compounds. *Journal of Memory and Language*, 65(4), 406–430.

- Keuleers E and Brysbaert M (2010) Wuggy: a multilingual pseudoword generator. *Behavior research methods*, 42(3), 627–33.
- Kiparsky P (1982) From cyclic phonology to lexical phonology. In: van der Hulst H and Smith H (eds), *The structure of phonological representations. Part 1*, Dordrecht: Foris.
- Kliegl R, Masson MEJ and Richter EM (2010) A linear mixed model analysis of masked repetition priming. *Visual Cognition*, 18(5), 655–681.
- Kornfeld LM (2009) IE, Romance: Spanish. In: Lieber R and Štekauer P (eds), *The Oxford Handbook of Compounding*, Oxford: Oxford University Press, pp. 436–452.
- Kuperman V (2013) Accentuate the positive: semantic access in English compounds. *Frontiers in Psychology*, 4(203), 1–10.
- Kuperman V, Schreuder R, Bertram R, et al. (2009) Reading Polymorphemic Dutch Compounds: Towards a Multiple Route Model of Lexical Processing. *Journal of Experimental Psychology: Human Perception and Performance*, 35(3), 876–895.
- Laka I (1996) *A Brief Grammar of Euskara, the Basque Language*. Leioa/Donostia: Euskal Herriko Unibertsitatea.
- Lardiere D (1995) L2 acquisition of English synthetic compounding is not constrained by level-ordering (and neither, probably, is L1). *Second Language Research*, 11(1), 20–56.
- Lardiere D (1998) Parameter Resetting in Morphology: Evidence from Compounding. In: Beck M-L (ed.), *Morphology and Its Interfaces in Second Language Knowledge*, Amsterdam/Philadelphia: John Benjamins, pp. 283–306.
- Lardiere D and Schwartz BD (1997) Feature-marking in the L2 development of deverbal compounds. *Journal of Linguistics*, 33, 327–35.
- Lemhöfer K, Dijkstra T and Michel M (2004) Three languages, one ECHO: Cognate effects in trilingual word recognition. *Language and Cognitive Processes*, 19(5), 585–611.

- Leys C, Ley C, Klein O, et al. (2013) Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology*, 49(4), 764–766.
- Libben G, Gibson M, Yoon YB, et al. (2003) Compound fracture: the role of semantic transparency and morphological headedness. *Brain and language*, 84(1), 50–64.
- Lieber R (1983) Argument Linking and Compounds in English. *Linguistic Inquiry*, 14(2), 251–285.
- Lieber R (2004) *Morphology and Lexical Semantics*. Cambridge: Cambridge University Press.
- Mägiste E (1979) The competing language systems of the multilingual: A developmental study of decoding and encoding processes. *Journal of Verbal Learning and Verbal Behavior*, 18(1), 79–89.
- Marelli M and Luzzatti C (2012) Frequency effects in the processing of Italian nominal compounds: Modulation of headedness and semantic transparency. *Journal of Memory and Language*, 66(4), 644–664.
- Marslen-Wilson WD (2007) Morphological processes in language comprehension. In: Gaskell G (ed.), *The Oxford Handbook of Psycholinguistics*, Oxford: Oxford University Press, pp. 175–193.
- Moyna MI (2011) *Compound Words in Spanish*. Amsterdam/Philadelphia: John Benjamins.
- Murphy VA (2000) Compounding and the representation of L2 inflectional morphology. *Language Learning*, 50(1), 153–197.
- Neubauer K and Clahsen H (2009) Decomposition of Inflected Words in a Second Language. *Studies in Second Language Acquisition*, 31(3), 403–435.
- Oxford University Press ., Cambridge University . and Association of Language Testers in Europe . (2001) *Quick placement test: Paper and pen test*. Oxford: Oxford University Press.
- Pallier C (2002) *Computing discriminability and bias with the R software*. Available from:

<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.218.7769&rep=rep1&type=pdf>.

- Piera C (1985) On Compounding in English and Spanish. In: Campos ER and Kempchinsky P (eds), *Evolution and Revolution in Linguistic Theory*, Washington, DC: Georgetown University Press, pp. 302–315.
- Portocarrero JS, Burright RG and Donovan PJ (2007) Vocabulary and verbal fluency of bilingual and monolingual college students. *Archives of Clinical Neuropsychology*, 22(3), 415–22.
- R Core Team T (2013) R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing, Available from: <http://www.r-project.org/>.
- Rastle K and Davis MH (2008) Morphological decomposition based on the analysis of orthography. *Language and Cognitive Processes*, 23(7-8), 942–971.
- Roeper T (1988) Compound syntax and head movement. In: *Yearbook of Morphology 1*, pp. 187–228.
- Rogers CL, Lister JJ, Febo DM, et al. (2006) Effects of bilingualism, noise, and reverberation on speech perception by listeners with normal hearing. *Applied Psycholinguistics*, 27, 465–485.
- Rosselli M, Ardila a, Araujo K, et al. (2000) Verbal fluency and repetition skills in healthy older Spanish-English bilinguals. *Applied neuropsychology*, 7(1), 17–24.
- Salthouse TA and Hedden T (2002) Interpreting Reaction Time Measures in Between-Group Comparisons. *Journal of Clinical and Experimental Neuropsychology*, 24(7), 858–872.
- Sandoval TC, Gollan TH, Ferreira VS, et al. (2010) What causes the bilingual disadvantage in verbal fluency? The dual-task analogy. *Bilingualism: Language and Cognition*, 13(02), 231–252.
- Sandra D (1990) On the representation and processing of compound words: Automatic access to constituent morphemes does not occur. *The Quarterly Journal of Experimental Psychology Section A*, 42(3), 529–567.

- Silva R and Clahsen H (2008) Morphologically complex words in L1 and L2 processing: Evidence from masked priming experiments in English. *Bilingualism: Language and Cognition*, 11(2), 245–260.
- Snodgrass JG and Corwin J (1988) Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General*, 117(1), 34–50.
- Stockall L and Marantz A (2006) A single route, full decomposition model of morphological complexity: MEG evidence. *The Mental Lexicon*, 1, 85–123.
- Tabri D, Chacra KM and Pring T (2011) Speech perception in noise by monolingual, bilingual and trilingual listeners. *International Journal of Language & Communication Disorders*, 46, 411–422.
- Taft M and Forster KI (1976) Lexical Storage and Retrieval of Polymorphemic and Polysyllabic Words. *Journal of Verbal Learning and Verbal Behavior*, 15, 607–620.
- Tuggy D (2003) Abrelatas and Scarecrow Nouns: Exocentric Verb-Noun Compounds as Illustrations of Basic Principles of Cognitive Grammar. *International Journal of English Studies*, 3(2), 25–61.
- Varela S (2012) Derivation and Compounding. In: Hualde JI, Olarrea A, and O'Rourke E (eds), *The Handbook of Hispanic Linguistics*, Noida, India: Wiley-Blackwell, pp. 209–226.
- Yoon J (2009) Constructional meanings of verb–noun compounds in Spanish: Limpiabotas vs. tientaparedes. *Language Sciences*, 31(4), 507–530.
- Zwitserslood P (1994) The role of semantic transparency in the processing and representation of Dutch compound words. *Language and Cognitive Processes*, 9, 341–368.

## APPENDIX A: Full list of stimuli

**TABLE A.1.** Target compounds and corresponding pseudocompounds, with descriptive statistics (length and lemma frequency) and overlap percentages.

COMPOUND	PSEUDOCOMPOUND	LENGTH	FREQUENCY	OVERLAP (%)
eyeliner	eathoner	8	1	55.56
bodybuilder	lodybearer	11	1.04	54.54
hairstylist	lansplessor	11	1.3	54.54
factory worker	fennory hester	13	1.51	60
pathfinder	pamelonder	10	1.51	60
nutcracker	nupscander	10	1.62	50
bookseller	boodsether	10	1.69	70
troublemaker	treaglemaner	12	1.7	66.67
shoemaker	groumayer	9	1.71	55.56
prize winner	prist marner	11	1.79	54.54
grasshopper	grastdopper	11	1.83	77.78
lawn mower	lan trower	9	2.05	77.78
screwdriver	scribslamer	11	2.23	44.44
taxi driver	teth prawer	10	2.27	44.44
onlooker	onstucker	8	2.3	77.78
screenwriter	scrowngriter	12	2.35	66.67

shopkeeper	fridseeper	10	2.4	50
dishwasher	destbucter	10	2.99	44.44
tape recorder	torse recober	12	3.09	75
minority leader	mendally geader	14	3.12	57.14
housekeeper	houthdroper	11	3.23	45.45
songwriter	sortshiter	10	3.42	77.78
security adviser	decutiny antiser	15	3.48	60
skyscraper	physcriper	10	3.49	70
typewriter	teansliter	10	3.8	66.67
babysitter	basysilmer	10	4.07	75
cheerleader	chettfeader	11	4.21	63.64
staff writer	stane gliter	11	4.47	77.78
storyteller	stomytowder	11	4.7	75
blockbuster	thockfasher	11	4.75	54.55
drug dealer	snib fealer	11	4.94	55.56
football player	neatbade pocker	14	5	50
bartender	fintimper	9	7.15	44.44
fund-raiser	fure-paiser	10	8.31	77.78
school teacher	squeet tearner	13	9.34	46.15
filmmaker	hoistayer	9	10.4	44.44
caregiver	carnsamer	9	11.52	55.56



firefighter	filtpigher	11	12.47	77.78
shareholder	shaundolder	11	14.21	77.78
homeowner	faleowler	9	14.31	55.56
police officer	porice osapher	13	23.88	77.78
taxpayer	tebmicer	8	26.66	44.44

---

*Note:* Frequencies, measured in occurrences per million, apply only to compounds. Length applies to both conditions.

**TABLE A.2.** Word fillers and corresponding nonword fillers, with descriptive statistics (length and lemma frequency) and overlap percentages.

WORD FILLER	NONWORD FILLER	LENGTH	FREQUENCY	OVERLAP (%)
people	pundle	6	1747.50	50
time	sile	4	1601	33.33
year	aste	4	790.37	33.33
way	faw	3	1031.74	33.33
thing	thab	5	466.23	66.67
day	vay	3	768.84	66.67
world	wonde	5	752.11	66.67
life	lomb	4	711.2	33.33
man	tam	3	681.20	33.33
president	predigent	9	612.42	77.78
school	scrool	6	608.25	83.33
history	farmory	7	282.08	55.56
state	stass	5	564.09	66.67
children	chelthen	8	558.53	66.67
house	hount	5	557.73	66.67
woman	gemen	5	532.73	40
percent	perfine	7	501.81	50

family	bedily	6	489.49	50
student	stumell	7	487.06	66.67
morning	mowling	7	284.62	71.14
work	welk	4	477.91	33.33
other	oster	5	1386.59	60
new	naw	3	1094.68	66.67
good	boud	4	907.98	50
american	aderimen	8	544.25	75
great	creet	5	464.44	33.33
old	elt	3	461.60	33.33
big	bew	3	458.42	33.33
high	humb	4	449.23	25
national	daboutal	8	422.09	37.5
different	debberent	9	414.30	66.67
small	glell	5	394.26	33.33
little	timple	6	380.34	50
black	prack	5	378.36	66.67
important	impennant	9	366.53	66.67
political	motominal	9	365.29	58.33
social	roviaI	6	344.03	66.67
long	lomp	4	338.40	66.67

young	doong	5	327.34	40
right	raunt	5	322.06	40
best	bers	4	320.08	50
white	whiss	5	319.045	66.67
just	jesh	4	1744.64	50
also	adro	4	1192.04	50
well	jeld	4	1054.63	50
how	zow	3	1015.80	66.67
only	onty	4	960.67	75
even	edan	4	922.28	50
still	stigs	5	758.11	60
when	whab	4	686.40	50
really	deenly	6	680.79	50
never	nyler	5	668.78	66.67
why	whe	3	603.29	66.67
already	alfeily	7	294.18	66.67
where	whost	5	498.73	33.33
always	adlage	6	459.12	50
all	ild	3	445.23	33.33
more	mure	4	405.04	75
often	upten	5	356.97	66.67

however	fabeyer	7	345.44	55.56
once	enge	4	321.73	33.33
much	mude	4	321.24	50
least	leams	5	301.12	66.67
know	kneb	4	996.93	50
think	snink	5	971.83	60
want	wals	4	517.56	50
say	bly	3	438.55	33.33
see	mee	3	418.98	66.67
get	keb	3	389.30	33.33
understand	untercrope	10	213.08	40
need	nood	4	299.06	50
mean	meap	4	283.56	75
look	moul	4	260.25	25
come	coss	4	249.19	50
make	mamp	4	235.14	50
let	lut	3	232.88	66.67
take	rade	4	231.35	50
thank	thash	5	203.57	60
believe	beseave	7	174.47	66.67
feel	foak	4	168.16	25

like	lind	4	157.63	50
tell	teck	4	135.61	50
use	ure	3	128.61	66.67
call	cank	4	125.02	50

---

*Note:* Frequencies, measured in occurrences per million, apply only to word fillers. Length applies to both conditions.

## Notes

<sup>1</sup> Throughout this paper, we use the term ‘sequential multilingualism’ to refer to the linguistic situation of individuals that have acquired their third or further language non-simultaneously with their first and/or second languages. We have avoided the term ‘additive’ (and its counterpart, ‘subtractive’) because the literature has typically used them to distinguish between situations of sequential multilingualism in which different sociolinguistic factors affect the relationship between the languages at play (see Aronin and Singleton, 2012: 28-29, for discussion). We thank an anonymous reviewer for pointing this out.

<sup>2</sup> Priming is a very broad concept in psycholinguistics, and may be explored between many other types of stimuli and across many other dimensions (e.g., semantic). For relevance, we only refer here to morphological and orthographic priming in the visual domain.

<sup>3</sup> ‘Noun-verb-er’. We will use this name throughout the text, although these compounds have been referred to in many different ways in the literature. A similar and frequent example is OV-er (object-verb-er), which emphasises the lexical syntax of the class.

<sup>4</sup> As Spanish citizens, these participants have a native-like command of Spanish, which is official in the whole territory and still the majority language in the Basque Country.

---

The strong presence of Spanish on TV, the Administration and often their university environment ensures a high proficiency in both languages.

<sup>5</sup> Meaning here *within-class* frequency. These were among the highest-frequency NV-er synthetic compounds in the corpus, but it should be noted that English compound words are generally low-frequency, and within these, NV-er compounds are on average less frequent than root compounds (e.g., *football*).

<sup>6</sup> T-tests conducted between the critical and filler conditions showed that there were significant differences in length,  $t(124) = 18.34$ ,  $p < .001$ , and frequency,  $t(83.09) = 14.2$ ,  $p < .001$ . between compound and noncompound stimuli. Admittedly, this may have enhanced the salience of compounds and pseudocompounds, counteracting the distracting effect of the filler/critical stimuli ratio (2:1). An alternative list with longer and less frequent filler words, however, would have made it impossible to test lower proficiency participants.

<sup>7</sup> The YES and NO buttons were manipulated by the right and left hands, respectively, regardless of hand preference. We repeated all analyses with handedness as a factor, which did not improve the fit of the statistical models. It is possible that the disparate proportions of left- and right-handed participants (only 3 participants out of the total 141 were left-handed) prevent seeing any potential effects, since this factor is practically controlled for. We thank an anonymous reviewer for bringing up this issue.



---

<sup>8</sup> Length is partially confounded with frequency in our stimuli: word fillers are, on average, both shorter (5 vs. 11 characters) and much more frequent (546 vs. 5.5 occurrences per million) than compounds. The importance of length alone, however, can be observed in the much faster responses to nonword fillers than to pseudocompounds, a comparison in which frequency does not play a role. Furthermore, these were also faster than responses to compounds, which have the advantage of lexicality.

<sup>9</sup> In our view, this further supports the idea that considering all responses might give us a better estimate of the processing times under scrutiny. However, as one anonymous reviewer points out, it might also be indicative that there is noise in the data, in the form of processing activity beyond or besides what is specifically targeted by the task (NV-er compound processing).