International Journal of Bilingualism
http://ijb.sagepub.com/

Lone English-origin nouns in Spanish: The precedence of community norms

Jessi Elana Aaron

International Journal of Bilingualism published online 14 January 2014
DOI: 10.1177/1367006913516021

The online version of this article can be found at:
http://ijb.sagepub.com/content/early/2014/01/06/1367006913516021

Published by:
SAGE
http://www.sagepublications.com

Article

The International Journal of Bilingualism

International Journal of Bilingualism
201X, Vol. XX(X) 1–22

DOI: 10.1177/1367006913516021
ljb.sagepub.com
SAGE

Lone English-origin nouns in Spanish: The precedence of community norms

Jessi Elana Aaron
University of Florida, USA

Abstract

This paper offers an examination of morphosyntactic factors that are generally understood to measure grammatical integration—and therefore used to help determine the status of other-language-origin nouns as borrowings or code-switches—through the lens of discourse, semantics, and lexical patterns. A total of 820 lone English-origin nouns surrounded by otherwise Spanish discourse are compared to Spanish and English nouns from the recorded speech of the same bilingual speakers in New Mexico. The semantic domains most open to English-origin nouns include both those traditionally expected, such as technology, and those generally thought to be unborrowable, such as kinship terms. In the case of determiner patterning, lone English-origin nouns' propensity to occur with indefinite articles or as bare is linked to its use in the predicating function. Regarding gender, the preference for masculine assignment for lone English-origin nouns is tied to both nonreferentiality and the general patterns found in Spanish. The impact is felt here not from English, but from the conventions of the local community. Among their many functions, these nouns are best suited in this community for naming kin, classifying individuals as belonging to a certain occupation, and creating verbal compounds. It is argued that the morphosyntactic patterns found reflect the community norms, in which English-origin nouns tend to perform certain discourse functions. Systematic quantitative analysis thus reveals the powerful role of discourse referentiality of nominal forms, in tandem with local practices.

Keywords

Code-switching, borrowing, community practice, discourse, referentiality, Spanish

1. Introduction

The literature on contact, including that focused on Spanish, is replete with commentary on the use of lexical items from one language in the discourse of another, as in (1), in which English-origin *cards* is surrounded by Spanish-language discourse.

Corresponding author:
Jessi Elana Aaron, Spanish and Portuguese Studies, University of Florida, P.O. Box 117405, Gainesville, FL 32611-7405, USA.
Email: jeaaron@ufl.edu

(1)

Francisco	*mucha gente juega muchas **cards** áhi en las casas no?*	'lots of people play **cards** there in the houses right?'

[18 Las minas, 0:44:36-0:44:40][1]

The use of single other-language-origin content words is the most frequent code-mixing phenomenon in contact situations (Jake, Myers-Scotton, & Gross, 2002, p. 72; Poplack & Meechan, 1998, p. 127). Scholars have generally focused on two main questions: a) *why* speakers use certain lexical items in certain contexts; and b) *how* and *to what extent* speakers incorporate these items into the grammar of the language surrounding them. Regarding the motivations behind such usage, it is often suggested that some new items enter a language to refer to new cultural concepts, constituting what Smead (2000) terms "unique loans" (p. 292), as in (2), in which *quarter* refers to a culture-specific item.

(2)

Anita	*.. papá le pagaba a mijo un **quarter** pa' que le hablara.*	'.. dad used to pay my son a **quarter** so that he would speak to him.'

[14 Calcetines, medias y mallas, 0:20:21-0:20:24]

Clearly, however, not all such items represent a lexical gap. In (3), for example, Francisco considers both *running* and *correr* as options in talking about winning a basketball game, and he uses them both within two contiguous Intonation Units (represented on separate lines in this and following examples).

(3)

Francisco	*.. (H) ... en todo el año de=l,*	'.. (H) … all year during the,
	.. State Champion[ship],	.. State Champion[ship],
Gabriel	*[aquí] [2en Peñas2]co?*	[here] [2in Peñas2]co?
Francisco	*[2I mean en2] –*	[2I mean in2] --
	(H) porque=,	(H) because,
	*.. era puro **running**,*	.. it was just **running**,
	*gana asi[na puro **correr**].*	you win [like that just **running**].
Gabriel	*[oh yeah].*	[oh yeah].
	[2yeah2].	[2yeah2].
Francisco	*[2(H)2] puro **correr**.*	[2(H)2] just **running**.'

[18 Las minas, 0:23:57-0:24:06]

Regardless of the general speaker motivations behind code-mixing, the grammatical status of single items is superficially ambiguous. Do they represent one-word code-switches, in which two grammars are juxtaposed, or are these items grammatically incorporated into the language that surrounds them, and thus best understood as borrowings (Poplack & Meechan, 1998)? In an earlier corpus of New Mexican and Southern Colorado Spanish (NMCOSS) (Bills & Vigil, 2008), Torres Cacoullos and Aaron (2003, p. 292) found that when incorporating English-origin nouns into their Spanish-language discourse, the speakers drew on Spanish, not English, grammar. This was evidenced in the patterns of bare (i.e. determiner-less) English-origin nouns in Spanish discourse, which followed those of bare Spanish-origin nouns in Spanish discourse and were unlike those of

bare English-origin nouns in the English discourse of the same speakers. With this, they conclude that, on the whole, these items were treated as instances of borrowing, not code-switching, the latter of which is understood as "the juxtaposition of sentences or sentence fragments, each of which is internally consistent with the morphological and syntactic (and, optionally phonological) rules of its lexifier language" (Poplack, 1993, p. 255).

A reliable view of how English and Spanish are combined in discourse can only be obtained through systematic quantitative examination of actual usage in a particular bilingual community. This paper investigates the status of single English-origin nouns in bilingual discourse in New Mexico, as nouns were not only the most numerous single-word other-language items, but are also an open word class that is particularly susceptible to borrowing. However, unlike the earlier NMCOSS corpus, the data examined here also include copious unambiguous multi-word code-switching. Through quantitative exploration of morphosyntactic, semantic, and discourse factors, I address the question of whether these nouns can be best understood as grammatically integrated or non-integrated into the recipient language (that is, as borrowings or single-word code-switches). At the same time, these factors reveal community-specific practices that intersect with—and perhaps supersede—the issue of morphosyntactic integration. The results will illustrate the power local practices have in defining linguistic norms.

2. Data and methods

2.1 Participants and corpus

The data for this study are drawn from the New Mexico Spanish-English Bilingual corpus (NMSEB) (Torres Cacoullos & Travis, in preparation), constituted by approximately 320,000 words and 30 hours of spontaneous speech. The participants are speakers of a Spanish that has been spoken in New Mexico for centuries, and not descendants of recent immigrants. The Northern New Mexican communities represented in this corpus have been in contact with English for 150 years, and the speakers all acquired both Spanish and English during childhood (Bills & Vigil, 2008). This diachronic depth affords us a wealth of information regarding the long-term effects of code-mixing within a community. NMSEB offers a particularly unique perspective on US Spanish phenomena, as it comprises spontaneous use of English and Spanish by a community of bilingual speakers who regularly use both languages, including during the same conversation (for more details on the corpus, see Torres Cacoullos & Travis, introduction; Travis & Torres Cacoullos, 2013). For this study, a sub-corpus of 18 interviews was examined, or approximately 180,000 words encompassing 17.5 hours of recorded speech from 20 speakers.

Single English-origin nouns occurred in 15 of these 18 interviews, and in the speech of 16 of the 20 speakers. (Interviews 4, 10 and 12 contained none, though these interviews did contain multi-word code-switches to English; see Wilson & Dumont, 2014, submitted. Interviewer data were not included.)[2] Among the speakers who did use single English-origin items, rates of use were highly variable, with a normalized frequency ranging from 6 to 215 tokens per 10,000 words (see Appendix).

2.2 Data: Lone English-origin nouns

The data include all English-origin nouns and compound nouns surrounded by Spanish discourse by the same speaker (N=1112 nouns, from 2449 English-origin words, including adjectives, verbs, and discourse markers), irrespective of Intonation Unit breaks. From this initial extraction, proper nouns, including names and places (N=629), as in (4), and nouns that acted as a title, and which

were thereby connected to a proper noun (*N*=32), as in (5), were excluded from the study because proper nouns may not be subject to the same processes of integration as common nouns (Poplack, Sankoff, & Miller, 1988, p. 99, n. 8).

(4)

Anita ... *no tenemos que tenerle miedo a* ***Russia***. '… we don't have to be afraid of **Russia**.'

[14 Calcetines, medias y mallas, 0:58:27-0:58:30]

(5)

Sandra *y mi* ***grandpa*** *Pablo?* 'and my **grandpa** Pablo?'

[03 Dos comadres, 0:32:28-0:32:29]

Also excluded were examples in which the noun was used in a metalinguistic context in which the noun itself was the topic of conversation (*N*=18), as in (6).

(6)

Manuel ... *allá les dicen* ***concession***. '… there they call them **concessions**.'

[16 Trip to Africa pt.1, 0:2:35-0:2:37]

Numbers, when treated as nominals as in (7), were included in this study; however, adjectival uses, as in (8), were excluded. Finally, items in which the noun of interest was truncated (*N*=12) were excluded.

(7)

Rafael *(H) pero compró como* ***six***. '(H) but he bought like **six**.'

[06 El túnico, 0:35:12-0:35:14]

(8)

Monica *me retiré cuando tenía* ***sixty*** *años=*. 'I retired when I was **sixty** ((*lit*. had sixty years)).'

[11 El trabajo, 0:0:15-0:5:18]

Beyond these exclusions, the English-origin lone noun dataset was further refined to ensure, first, that the noun in question would not generally be considered part of a monolingual Spanish speaker's lexicon, and second, that the compounds were conventionalized compounds among monolingual American English speakers (and therefore could be treated as single words). To satisfy the first criterion, the English-origin nouns that were listed in the most recent edition of the *Diccionario de la Real Academia* (DRAE, http://www.rae.es/rae.html), were excluded from these data (99 lexical types, *N*=102; see Appendix). Retained were cases where the definition provided in the DRAE was entirely different from the semantics found in the data (*N*=29). These uses were not judged to be equivalent to the DRAE listings, for nouns such as *brecas* 'brakes,' *carton* (as a container), *chanza* 'chance,' *craques* 'cracks,' *cricket* (the insect), *home* (rest home), *light* (as in *rays of light*), *machine, post, robe, semi* (truck), *trust*, and *yarda* 'yard' (as in a piece of land next to a house).

To satisfy the second criterion, that the compounds were conventionalized compounds among monolingual American English speakers (and therefore could be treated as single items), I only included potential compounds that were found in *Merriam-Webster's English Dictionary* (http://www.merriam-webster.com/), like *alarm clock, chain saw, day care, duct tape, ghost town, mental illness*, and *monkey bars* (see Appendix for excluded compounds). This left 820 lone English-origin nouns.

Though adjectives in general were excluded from this study, a handful (N=11) of lone adjectives were found with articles, and therefore functioning as nouns, as in (9).[3] This structure is generally acquired between the second and third year among monolingual Spanish speakers, which may be taken to indicate a certain level of syntactic complexity (Snyder, Senghas, & Inman, 2001, p. 162). Its presence in these data—both with Spanish-origin and English-origin nouns—speaks to the extent to which English-origin words have been incorporated into everyday language.

(9)

[06 El túnico, 0:7:37-0:7:39]

Ivette	*como que les ponía más atención a los* ***fourth*** *que a los* ***fifth***.	'It's like she paid more attention to the **fourth** ((graders)) than the **fifth** ((graders)).'

2.3 Baseline comparisons

There are several measures to determine whether an other-language-origin item is a borrowing or a code-switch, the former of which involves recourse to only one grammar, and the latter of which involves the juxtaposition of two grammars. Items that are frequent and that occur among many speakers (i.e. are diffuse) are expected to function as borrowings, but the primary criterion by which borrowing status is determined is on the basis of morphosyntactic integration: an established borrowing like *troca* 'truck' should pattern no differently from a Spanish-origin noun like *camioneta* 'truck.' The morphosyntactic criterion, however, applies regardless of whether the item is established; that is, any item that shows clear signs of morphosyntactic integration can be seen as a borrowing.

However, with languages as typologically similar as English and Spanish, morphosyntactic integration is often not easy to identify. This is particularly true because in many instances Spanish and English grammars show variability, such that their contrast is probabilistic, not absolute. To determine the grammatical status of lone other-language-origin nouns in code-mixed discourse on a larger scale, it is necessary to identify a conflict site, i.e. "a form or class of forms which differs functionally, structurally, and/or quantitatively across comparison varieties" (Poplack & Meechan, 1998, p. 132; Poplack & Tagliamonte, 2001, p. 101). In other words, the distributional patterns of contentious items (i.e. those that are neither frequent nor diffuse) must be compared with those of other items whose grammatical status is clear. Nouns in monolingual stretches of speech (e.g. Spanish-origin nouns surrounded by Spanish) are unambiguous. Moreover, established borrowings, as determined by diffusion within the speech community, also demonstrate morphosyntactic integration.

To identify conflict sites and examine distributional patterns, comparison datasets must be prepared. In this case, the NMSEB offers us the unique opportunity to explore the patterns found within the two languages of the speech community itself, instead of resorting to an idealized monolingual norm or some other regional variety. The use of Spanish and English from the same speakers allows us to avoid the "comparative fallacy" (Bley-Vroman, 1983), since we are assuming that the most appropriate measure of language norms, and therefore of integration or lack thereof, is within the bilingual community itself (see Poplack & Meechan, 1998; Pires & Rothman, 2009; Torres Cacoullos & Aaron, 2003).

To create comparison datasets for each of the languages in contact, a roughly equivalent number of Spanish-origin nouns surrounded by Spanish discourse (N=856) and up to the same number of

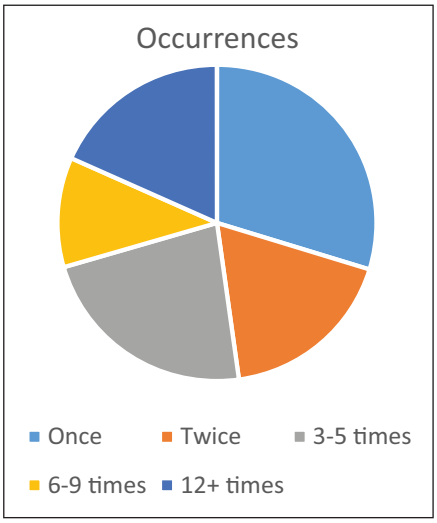

Figure 1. Breakdown of lone English-origin items according to number of occurrences (*N*=820). For example, 30% of all tokens are of items that occurred only once.

English-origin nouns surrounded by English discourse (*N*=608) was extracted from each speaker, taking into account the number of occurrences provided by each in the lone English-origin dataset.[4] This ensured that the monolingual data were comparable with the English-origin data in terms of any individual biases or tendencies. For these datasets, extraction began at the 15-minute mark for each interview.

3. Results

3.1 Frequency distribution of lone English-origin nouns

The 820 lone English-origin nouns represent a total of 405 types. Of the 820 occurrences, 30% (*N*=244) appeared only once in these data (these will be referred to as "Singleton" lone items), and another 18% (*N*=148, with 74 types) appeared only twice. Another 23% (*N*=186, with 52 types) appeared three to five times, and 11% (*N*=92, 13 types) appeared six to nine times. The remaining 20% of the data (*N*=150) includes six types, which occurred between 12 and 51 times. These six types were: *uranium* (*N*=12), *troca/truck* (*N*=15), *grandpa* (*N*=19), *daddy* (*N*=25), *grandma* (*N*=28), and *dad* (*N*=51). Figure 1 shows this distribution.

However, given that 70 of the repeated items, such as *uranium* (*N*=12), were simply repeated multiple times by the same speaker, a better gauge for how widespread the English-origin words were is how many speakers used them. There were 314 types (*N*=433) used by only one speaker; of these, 244 occurred only once. There were 52 types (*N*=154) produced by two speakers. Finally, 12 types (*N*=51) were produced by three speakers, five types (*N*=53) by four speakers, and six types (*N*=129; *grandpa, mom, weekend, troca* 'truck', *grandma*, and *dad*) by six or more speakers. Those produced by three or more speakers will be considered "Diffuse" in this paper. Figure 2 shows this distribution both in terms of number of occurrences and number of types.

To recap, I will call those produced only once "Singletons," and those produced by three or more speakers "Diffuse." This allows us to clearly distinguish between those items that are commonly used within the community and are therefore likely loanwords (and may be transmitted by

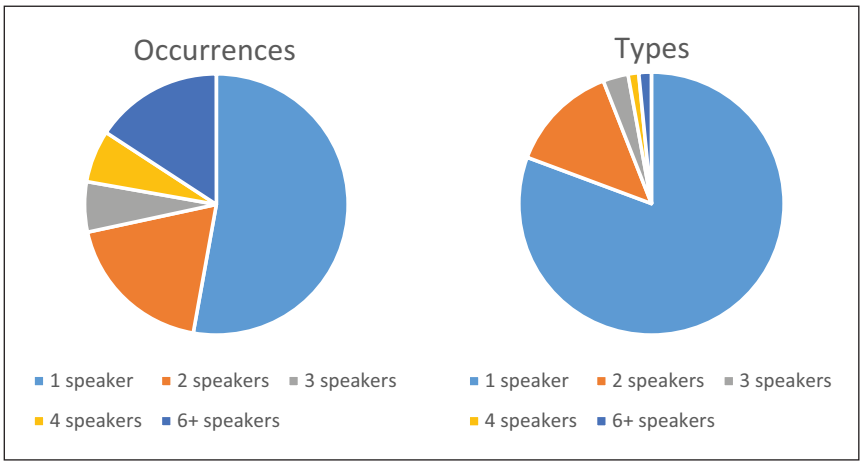

Figure 2. Breakdown of lone English-origin nouns according to number of speakers (for "Occurrences," *N*=820; for "Types," *N*=389). For example, 129 of all 820 occurrences (i.e. 16% of the data) were produced by six or more speakers; however, these 129 occurrences are made up of only six types (i.e. distinct lexemes), making up under 2% of the 433 lexical items.

other speakers), from those whose possible loanword status is not established (and which therefore may entail active access to English). In the case of Diffuse items, we would expect these to pattern like Spanish-origin nouns (which, for the most part, they do). In contrast, the Singleton items may pattern like English nouns, revealing them to be, in the aggregate, single-word code-switches. Alternatively, they may pattern like the Diffuse English-origin nouns and like the Spanish nouns, revealing their aggregate status as nonce loans that are instantly integrated morphosyntactically (Poplack, 2012; Poplack & Meechan, 1998).

3.2 Semantic domains

It has often been argued that lexical borrowings tend to be drawn from certain semantic domains, particularly those that have to do with new cultural items that are associated with the contact culture. Smead (2000) identifies two types of loanwords: "unique" loanwords, defined earlier, and "synonymic" loanwords, i.e. "those that compete for the same semantic space with a native language term" (p. 292). Teschner's (1974) annotated bibliography reveals several semantic domains that appear more likely to have English-origin words, including education, food, sports, and technology. It is generally assumed that highly frequent or core vocabulary items (Swadesh, 1971) are not likely to be borrowed (e.g. Smead, 2000, p. 282). However, Myers-Scotton and Okeju (1973) argue that (based on evidence from the East African language Ateso) "borrowings within the core vocabulary itself are also very common, given sufficiently extensive contact with another culture" (pp. 872–873).

The preliminary categories coded here were gleaned from the previous literature. These were then supplemented as the analysis was guided by the data. The 17 semantic categories coded included: kinship terms (e.g. *dad*); everyday items (e.g. *bag*); events and places (e.g. *birthday*); person (not kin) (e.g. *firefighter*); year or number; work- or money-related (e.g. *spending money*); related to the land or earth (e.g. *mountain*); technology (e.g. *treadmill*); academia (or school) (e.g. *homework*); food, drink, and smoke (e.g. *popcorn*); vehicle or transport-related (e.g. *troca*); institution (e.g. *social services*); animal (e.g. *bird*); abstract time/space concept (e.g. *weekend*); health and body (e.g. *eyedrops*); domestic life (e.g. *sewing*); and linguistics or language (e.g. *Spanish*).

Table 1. Semantic domains per lone English-origin nouns vs. Spanish and English nouns in corresponding samples.

	Singleton		Diffuse		Spanish		English		Total	% English-origin
	N	%	*N*	%	*N*	%	*N*	%		
Everyday items	34	14	4	2	84	10	65	11	187	20
Events and places	28	12	8	3	34	4	54	9	124	29
Other	28	12	18	8	86	10	78	13	210	22
Person	24	10	17	7	83	10	60	10	184	22
Year or number	19	8	4	2	4	<1	14	2	41	56
Technology	19	8	4	2	21	2	14	2	58	40
Earth	14	6	6	3	58	7	22	4	100	20
Food, drink, and smoke	14	6	0	–	41	5	25	4	80	18
Health	11	4	0	–	37	4	38	6	86	13
Academic	9	4	11	5	33	4	14	2	67	30
Nominalized adjective	9	4	0	–	19	2	4	<1	32	28
Animals	8	3	0	–	23	3	13	2	44	18
Domestic items	7	3	0	–	24	3	56	9	87	8
Institutional	7	3	0	–	10	1	17	3	34	21
Work or money	7	3	0	–	61	7	27	4	95	7
Vehicles	4	2	19	8	14	2	7	1	44	52
Abstract time/space	3	1	8	3	101	12	52	9	164	7
Kinship terms	2	<1	134	58	74	9	28	5	238	57
Linguistic	2	<1	0	–	40	5	20	3	62	3
Total	244	100	233	100	856	100	608	100	1941	25

Nominalized adjectives and other items that did not fall into one of these domains were coded separately.

The tendency for both culture-specific and more universal (or "core") terms to be borrowed in this community (cf. Torres Cacoullos & Aaron, 2003) is seen in Table 1. The last column shows the proportion of each domain that is produced as lone English items (based on a comparison with the Spanish and English sample described above).

This analysis shows that English-origin nouns do indeed occur relatively more often in particular semantic domains. These domains, however, are not all on the list of usual suspects, which, as mentioned, include technology, food items, and academia. The most favorable contexts for English-origin lone nouns in these data are kinship terms (which were produced as lone English-origin nouns 57% of the time), years and numbers (56%), technology (40%), and vehicles (52%). The items referring to events, technology, or vehicles, as in (10), (11), and (12), may indeed represent cultural novelties (though in the case of vehicles, 79% (15/19) of the Diffuse items are *troca*). However, the category of kinship belongs to the core vocabulary (Smead, 2000, p. 282) and cannot in this case be said to be culturally specific or represent lexical gaps.

(10)

Anita <*VOX .. anda llévanos pa'l* ***drive-in*** *VOX*>. '<VOX .. come on take us to the **drive-in** VOX>.'

[14 Calcetines, medias y mallas, 0:44:28-0:44:30]

(11)

Francisco *y tienen un* ***power plant*** *grande ahí en Prewitt*, 'and they have a big **power plant** there in Prewitt,'

[18 Las minas, 0:33:55-0:33:59]

(12)

Manuel *.. (H) .. y esto no lo puedo cargar en el* ***airplane***, '.. (H) .. and I can't carry this onto the **airplane**,'

[16 Trip to Africa pt.2, 0:12:01-0:12:03]

The English-origin kinship terms found in these data exhibit certain lexical patterns that provide evidence that the conventionalization of English-origin nouns occurs at both the lexical and semantic level (cf. Poplack et al., 1988), since we find both *dad* and *grandma* (but not other kinship terms) among the most frequent. For example, *dad* and *daddy* combined occurred 76 times as lone nouns, compared with four occurrences of *papá* among the Spanish nouns. In contrast, *mom/momma* occurred only eight times in the lone nouns, and *mamá* 15 times in the Spanish nouns. *Grandma* had 30 occurrences; *abuela* only one. *Grandpa* occurred 20 times, *abuelo* three. There were no occurrences of *brother* or *sister*, while mentions of *hermano, hermana*, and *hermanos* were made 11 times. *Son* and *daughter* did not appear (though *son-in-law* occurred once); *hija* and *hijo* occurred 18 times combined.

These patterns are mirrored in the corpus as a whole: English terms are used over 1.5 times as often as Spanish terms to refer to dads (305 vs. 182 tokens), while for moms, the breakdown is reversed (178 tokens of English terms vs. 219 in Spanish).[5] Given the overwhelming preponderance of kinship terms, they will be excluded from the analyses of morphosyntactic and discourse patterns in the Diffuse dataset, as they tend to display peculiar tendencies that would skew the results (for example, they show a greater tendency than other nouns to be referential (as in ex. 18), to occur in subject position, and with possessive marking).

In the datasets in this study, while the core items of kinship show evidence of high rates of English-origin nouns, the core items in the domain of abstract time or space are much less likely to appear in English within Spanish discourse (2%, N=15), when compared with these types of items

in Spanish (12%, *N*=101). Over half of these English-origin items are *weekend(s)* (*N*=8); items like *day* and *year* did not appear among the lone nouns, though there were 14 and 23 occurrences of these in the Spanish nouns, respectively. Nonetheless, the relatively high rate of specific years and numbers in the English-origin data shows that use of English in this related context is a local community norm. Indeed, the 36 occurrences of years (some as Singletons, some repeated across speakers, some Diffuse), as shown in (13), (14), and (15), are best understood as instances of a construction in which it is established practice to use English in this community.

(13)

Susan	... *nació en* ***eighteen sixty nine***,	'he was born in eighteen sixty-nine,'

[01 El abuelo, 0:29:43-0:29:46]

(14)

Francisco	*(TSK) (H) en el* ***eighty*** *algo.*	'(TSK) (H) in **eighty** something.'

[18 Las minas, 0:22:47-0:22:48]

(15)

Ivette	... *(H) se me hace que en ..* ***seventy one*** *me fui a vivir con él.*	'…(H) I think in .. **seventy-one** I went to live with him.'

[06 El túnico, 0:2:35-0:2:38]

A third area that at first blush seems to be disproportionately favorable to lone nouns is that of vehicles. However, 15 of the 19 occurrences are represented by *troca*, which indicates that there is a lexical effect, rather than a semantic class effect. Thus, while some use of English-origin nouns are domain-specific (e.g. numbers, kinship), there are also lexical effects.

3.3 Gender

Gender has been a primary target for the examination of the status of lone items because it can be understood as one indication of morphosyntactic integration: if the gender patterns are similar to those of Spanish-origin nouns, then it can be argued that they are morphosyntactically integrated, and thus not code-switches (e.g. Poplack, Pousada, & Sankoff, 1982; Zamora, 1975; Zamora Munnt & Btjar, 1987). A preference for masculine gender assignment has been suggested in several previous studies, among L2 learners (Martinez-Gibson, 2011), bilinguals (Chaston, 1996; Clegg, 2010: 16; Garcia, 1998; Montes-Alcalá & Lapidus Shin, 2011; Sánchez, 1995: 134-137), and monolinguals (Banfield, 1994; Pérez-Pereira, 1991; Smith, Nix, Davey, López Ornat, & Messer, 2003).

To determine gender for both Spanish-origin and English-origin nouns, overt cues beyond the noun itself were used. These cues included the determiner gender and adjectival modifiers, as in (16), in which *machine* was coded as feminine due to the feminine marking on *otra*. If no overt cues were present, the noun was coded as unmarked for gender. In cases in which the semantics of the noun dictated one or the other gender, such as *dad*, a separate code was used, to ensure that all nouns coded as having been assigned a gender were, at least in theory, sites at which either gender could have been assigned.[6] Overt cue-based gender coding was necessary across all datasets to ensure comparability of the coding.

Table 2. Distribution of gender marking.

	No info		Masc.		Fem.		Total
	N	%	*N*	%	*N*	%	*N*
Singleton	122	51	101	42	16	7	239
Diffuse	31	35	39	44	19	21	89
Spanish	313	39	239	30	259	32	811

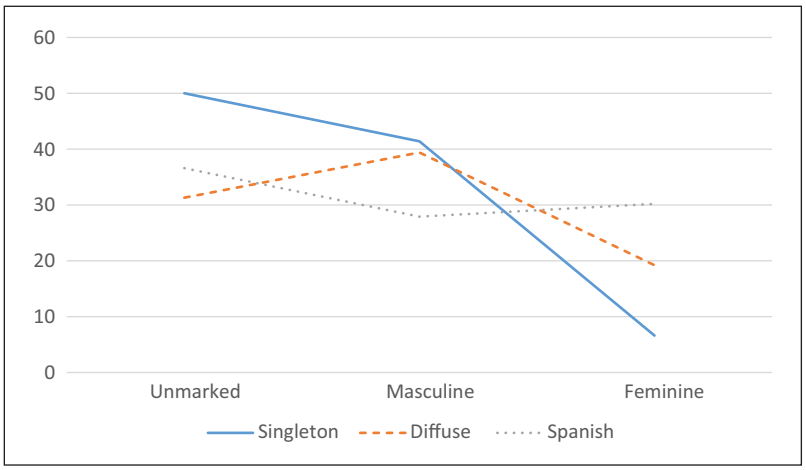

Figure 3. Gender.

(16)

Ivette *tenías que irte pa' otra* ***machine***. 'you had to go to *another* **machine**.'

[06 El túnico, 0:35:32-0:35:33]

Table 2 and Figure 3 show the results regarding gender assignment for these data. Note that if we compare only masculine and feminine, we get a rate of 14% (16/117) feminine for Singletons. This is nearly identical to the results reported by others for other bilingual populations (Mota, cited in Otheguy & Lapidus, 2003, pp. 214–215; Otheguy & Lapidus, 2003; Poplack, 1982). Also notable here is the high rate of unmarked nouns, at 51% for Singletons. A superficial analysis of Singletons alone might lead us to take this as evidence of the lack of grammatical integration. However, a comparison with the rate of unmarked Spanish nouns, also high at 39%, discourages such an interpretation. This slightly elevated rate of the absence of gender marking can be attributed not to a lack of grammatical integration, but rather to the greater proportion of the Singletons used in nonreferential contexts, as will be discussed below.

The preference for masculine gender reported in other studies of English-origin nouns in Spanish is found among these speakers as well, for Singleton, and Diffuse lone items, at 41% and 39%, respectively, compared with 28% for the Spanish nouns dataset. In contrast, only around 6% of the Singleton data was assigned feminine gender, compared with 19% for the Diffuse lone items

and 30% for the Spanish nouns data. Figure 3 illustrates the contrast between Spanish and Singletons; Diffuse lone English-origin items fall in the middle.

The preference for masculine, however, likely has nothing to do with code-mixing tendencies per se, but may rather simply follow from patterns and preferences that are internal to Spanish. In a connectionist analysis of a longitudinal database of parental production, Smith et al. (2003) noted that "while regular feminine nouns were slightly more frequent than regular masculine nouns, irregular masculine nouns outnumbered irregular feminine nouns by roughly 2 to 1" (p. 306). Based on this input, which preserved type and token frequencies, a computer-generated model produced a similar bias toward masculine gender assignment to novel words, suggesting that the frequency distribution has a direct role in gender assignment.[7] When applied to lone English-origin items, such experiential patterns would likewise lead to a preference for masculine gender assignment.

3.4 Discourse function

Lacking from most previous studies of English-origin nouns in Spanish is consideration of patterns of referentiality (Hopper & Thompson, 1984; Thompson, 1997). This turns out to be a particularly profitable context where we may be able to measure the extent to which lone English-origin nouns are grammatically integrated, and can therefore be considered to be functioning as Spanish nouns and not as code-switches to English.

The data were coded for referentiality, which, in Hopper and Thompson's (1984, p. 711) terms, refers to the "manipulability" of the referent; that is, "a noun phrase is *referential* when it is used to speak about an object as an object, with continuous identity over time" (DuBois, 1980, p. 208). Thompson and Hopper (2001) identified three contexts in which nouns perform nonreferential functions: to form intransitive predicates with semantically weak verbs, to classify, and to orient. These contexts are illustrated in (17), (18), and (19), respectively.

(17)

Ivette	*.. él me daba **ride***.	'.. he would give me a **ride**.'
		[06 El túnico, 0:52:27-0:52:28]

(18)

Susan	*...(1.1) era también **farmer***,	'…(1.1) he was a **farmer** too,'
		[01 El abuelo, 0:31:50-0:31:53]

(19)

Manuel	*...(0.8) tenemos un toothpick*, *...(1.0) es u=n*, *.. credit card*, *ése*, *casi lo pudiera acarrear en la **wallet***, *.. pero no te dejan.*	'…(0.8) we have a toothpick, …(1.0) it's a, .. credit card, that one, you could almost carry it in your **wallet**, .. but they don't let you.'
		[16 Trip to Africa pt.2, 0:12:42-0:12:48]

Referentiality is key to the study of single other-language-origin nouns, as Torres Cacoullos and Aaron (2003) found that nonce-loan nouns in New Mexican Spanish often occurred in predicating

Table 3. Referentiality.

	Orienting		Classifying		Predicating		Total Non-ref.		Total
	N	%	*N*	%	*N*	%	*N*	%	*N*
Singleton	59	24	41	17	64	26	164	67	244
Diffuse	34	34	4	4	25	25	63	64	99
Spanish	259	30	69	8	159	19	487	57	856
English	170	30	85	14	95	16	350	58	608

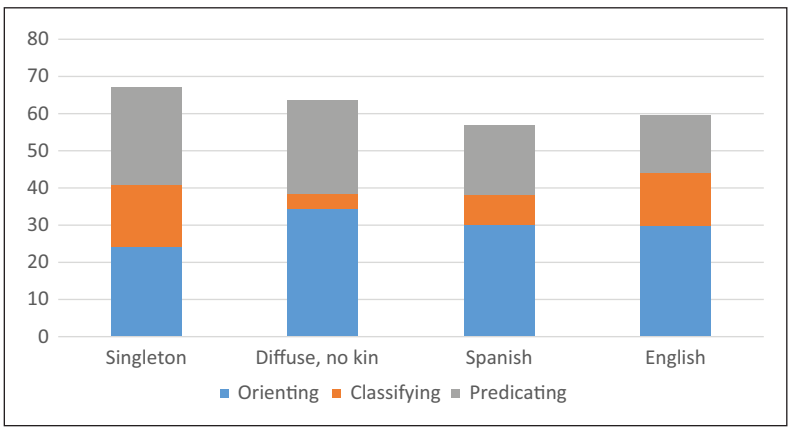

Figure 4. Nonreferential uses.

functions, and were disproportionately frequent in classifying contexts, where they designated occupation or social status as predicate nominals. This led them to conclude that "It is not lack of grammatical integration, but these nonreferential uses in recipient-language predicates, that is manifested in bare nonce-loan nouns" (p. 293).

If we examine Table 3 and Figure 4, we see that both lone nouns datasets (Singleton and Diffuse) showed higher rates than both English nouns and Spanish nouns in the predicating role. This means that the former were more likely to appear as nonreferential objects of semantically weak verbs, thereby forming verbal compounds (Thompson & Hopper, 2001), as in (17). However, given that both Diffuse and Singleton lone nouns demonstrate this tendency, and this is not a conflict site between Spanish and English because nouns from both languages readily occur in this context, this cannot be understood as a measurement of grammatical integration. Instead, we may posit that this reflects the community practice of creating novel compounds, perhaps at times to express culturally specific concepts, as in (20). This could potentially be linked to the similar community practice of creating verbal compounds with *hacer* 'do' + V, which is a novel bilingual device (Wilson, 2013; Wilson & Dumont, this volume).

(20)

Mónica .. *están agarrando* ***credit***, '.. they *are getting* **credit**,'

[11, El trabajo, 0:9:37-0:9:39]

In fact, in terms of discourse referentiality, the only notable difference between English nouns and Spanish nouns is found in the classifying role, as in (18). Here, we find that English nouns and Singletons are more likely to fulfill this function. Notably, 22% (9/41) of the classifying Singletons refer to occupation, compared with 9% (8/85) of the English nouns, 9% (6/69) of the Spanish nouns, and none of the Diffuse nouns. (Notably, however, 22% (19/85) of the English classifying nouns are human.) The nonreferential use of Singletons in occupation-related classifying function was also found in Torres Cacoullos and Aaron (2003), and may explain in part some of the gender-marking prevarication noted by Otheguy and Lapidus (2003) in this context. For example, in (21), the lone item does not actually refer to the person herself, but only a category to which she belongs. Since the category itself—social worker—is unmarked for gender (i.e. both women and men can fulfill this role), its gender assignment would not necessarily be linked to that of the woman, who represents only one example of somebody in that category.

(21)

¿Y tu mamá? Ella es un social worker, una trabajadora social…	'And your mother? She is **a social worker** (M), **a social worker** (F)…'

[228D, cited in Otheguy & Lapidus, 2003, p. 216]

The explanation for the elevated occurrence of English nouns in classifying role, however, is not the same. Instead, it comes from variations of the expression *that's the thing*, which accounted for another 9% (8/85) of the English nouns data; in contrast, *cosa* 'thing' occurs only once in classifying position in the Spanish nouns dataset. Finally, overall, Singletons are less likely to perform orienting functions (24%) than nouns in all other datasets; examples include numbers and words such as *north, winter, month*, and *weekend*.

A final site in which the ramifications of the discourse functions of English-origin nouns in New Mexico can be found is in the determiner patterns. Determiners have been cited as an ideal area in which transfer might be identified (e.g. Montrul & Ionin, 2010). However, the scant evidence that exists on the question of transfer in article interpretation among Spanish-English bilinguals offers mixed results. One difficulty lies in the relative scarcity of obvious conflict sites in English and Spanish.

Two clear loci of inter-lingual difference do exist, however. First, in contexts of inalienable possession, as in (23), Spanish generally has a definite article, while English would use a possessive pronoun. Second, in generic contexts, as in (24), Spanish generally has a plural definite article, while English nouns in these contexts tend to be bare. Unfortunately, these two contexts were rare in the data (N=1 for inalienable possessions; N=20 for generics), and thus could not be analyzed here. In order to gain a fuller understanding of conflict sites between typologically similar languages, then, we must move beyond these clearer cases. Corpus studies are particularly well suited for more fine-grained analysis.

(23)

Mónica	*le duele el* ***espinazo***.	'*his* **spine** hurts.'

[11 El trabajo, 0:20:23-0:20:24]

(24)

Rafael	*... y qué tan ra- -- rápido estabas yendo?*	'… and how qui- – quickly were you going?

	.. qué tan veloz?	.. how fast?
Ivette	*.. pues,*	.. well,
	no se me hace que venía muy recio porque me pude ir pa' atrás pa' adentro.	it doesn't seem to me that I was coming that quickly because I could go back inside.
	... y luego,	… and then,
	después me acuerdo que,	afterwards I remember that,
	*... % .. tenía tanto miedo a las **stop signs**.*	… % .. I was so afraid of ((the)) **stop signs**.'

[06 El túnico, 0:49:07-0:49:17]

Determiner types considered here included bare (25), definite (26), indefinite (27), possessive (28), demonstrative (29), and quantifier (30).

(25)

Adriana	*... no sé la historia.*	'… I don't know the story.
	.. muy bien pero,	.. very well but,
	*(H) ...(1.6) **granma** tiene .. un --*	(H) …(1.6) **grandma** has .. a --
	... like an angel?	*like an angel?*'

[03 Dos comadres, 8:19-8:21]

(26)

Molly	*lo van a dejar vivir en la casa de <u>la</u> **grandma**,*	'they are going to let him live at **grandma**'s house,'

[09 La salvia, 0:53:25-0:53:28]

(27)

Fabiola	*... pero ese tenía <u>una</u> **girlfriend**,*	'… but that one had *a* **girlfriend**,'

[09 La salvia, 0:54:53-0:54:54]

(28)

Sandra	*<X muy X> linda <u>mi</u> **grandma**.*	'<X really X> lovely *my* **grandma**.'

[03 Dos comadres, 0:2:30-0:2:31]

(29)

Bartolomé	*... (H) y nosotros agarramos parte de <u>esos</u> **crews**,*	'… (H) and we would grab part of *those* **crews**,'

[02 La marina, 0:32:23-0:32:28]

(30)

Francisco	*(TAP) (H) mucha gente juega <u>muchas</u> **cards** áhi en las casas no?*	'(TAP) (H) a lot of people play *a lot of* **cards** there in the houses no?'

[18 Las minas, 0:44:36-0:44:40]

Table 4. Distribution of determiner types.

	No Det.		Definite		Indefinite		Possessive		Quant.		Demon.		Total	
	N	%	*N*	%	*N*	%	*N*	%	*N*	%	*N*	%	*N*	%
Singleton	114	**47**	70	**29**	42	**17**	7	3	8	3	3	1	244	100
Diffuse	30	30	43	**43**	11	11	3	3	7	7	5	5	99	100
Spanish	253	30	369	**43**	62	7	73	8	67	8	32	4	856	100
English	176	29	180	**30**	97	**16**	96	16	38	6	21	3	608	100

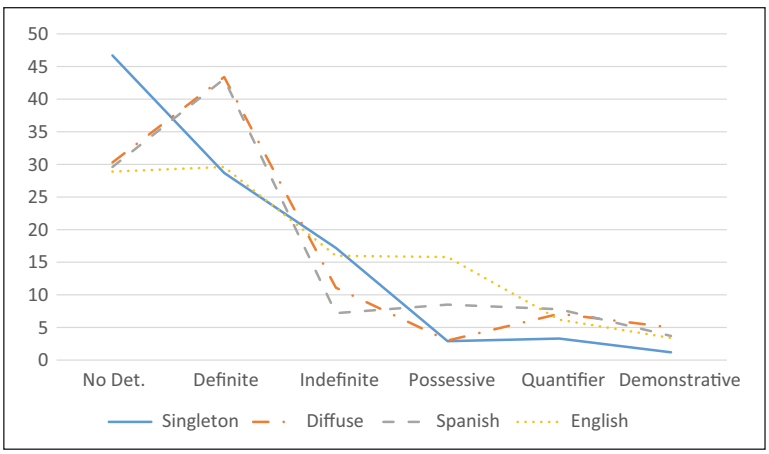

Figure 5. Determiners.

Table 4 and Figure 5 show the distribution of determiner types for all four datasets. Regarding definiteness, it has been proposed that definite articles are more frequent, and indefinite articles less frequent, in Spanish than in English (see, e.g. Vargas-Barón, 1952, p. 410). This characterization is borne out here: the Spanish data show 43% definite articles, compared with 30% in English; conversely, Spanish shows 7% indefinite articles, compared with 16% in English. The Diffuse items line up with Spanish in terms of definite articles, at 43%, and are closer to Spanish than to English in indefinite article rates. Singletons, however, as shown clearly in Figure 5, are in line with English for both, at 29% definite articles and 17% for indefinite articles. This suggests that, at least sometimes, speakers may treat these items like English in terms of (Spanish) article usage. However, if possessives and definite articles are combined into one category, the English looks much like the Spanish, suggesting that the "conflict" may not be as great as it appears at first glance.

One context in which Singletons do indeed seem to align with English is in the predicating context. Only 1% (15/159) of the predicating nouns in Spanish occurred with indefinite articles, as in (31); much more common was zero marking, as in (32), produced by the same speaker only moments earlier. This compares with 19% (12/64) indefinite articles among predicating Singletons and 24% (23/95) in English, as shown in Figure 6. Notably, the indefinite article with Singletons in this context, as in (33), tends to occur with culture-specific nouns, such as *guzzler, recreational vehicle*, and *B-47*. Thus, the predicating noun construction may be a context in which we may see some one-word code-switches into English.

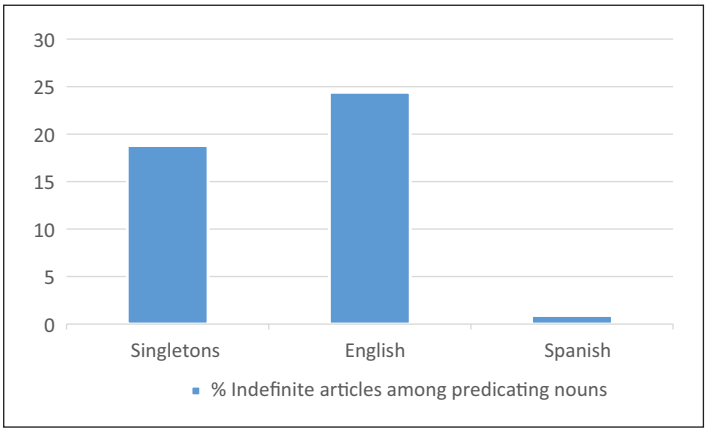

Figure 6. Percentage of indefinite articles among predicating nouns in Singletons, English nouns, and Spanish nouns (*N*=64, 95, and 159, respectively).

(31)

Fabiola *tiene una **novia**?* 'does he have *a* **girlfriend**?'

[09 La salvia, 0:26:41-0:26:42]

(32)

Fabiola *...(1.2) y él no tiene **Ø novia**?* '…(1.2) and he doesn't have ((a)) **girlfriend**?'

[09 La salvia, 0:26:36-0:26:39]

(33)

Pedro *.. antes le ponían un **cast** a uno*, '.. before they would put a **cast** on you,'

[07 Basketball teams, 0:26:33]

Nevertheless, Singletons do not always align with English when it comes to bare marking. Singletons clearly stand out; at 47% bare, these nouns are marked with zero about 16 percentage points more than nouns in the other datasets, all of which hover near 30%. It is worth noting that 13% (15/114) of the bare singletons are numbers or years, compared with 7% (13/176) of the English nouns. So, while the prevalence of numbers in the Singletons may play some part in the elevated rate of bare nouns, the overall trend goes beyond lexical tendencies: it is linked to nonreferentiality. Nonreferentiality is fairly common among bare nouns, at 75% (*N*=86/114) of the bare singletons, and 83% (*N*=184/253) of Spanish nouns. The higher proportion of nonreferential contexts with Singletons, then, like in Torres Cacoullos and Aaron (2003), would explain the higher rate of bare nouns in this dataset.

In sum, an examination of morphosyntactic factors that are generally understood to measure grammatical integration—and are therefore used to help determine the status of other-language-origin nouns as borrowings or code-switches—through the lens of discourse, semantics, and lexical patterns, has revealed the fingerprint of this community's code-mixing norms. The morphosyntactic

patterns found are in fact reflections not of a lack of grammatical integration, but rather of the use of English-origin words to perform certain discourse functions. These norms are established locally; there is nothing about English nouns in particular that makes them more or less suitable in these contexts. As we saw in section 3.2 above, the power of the community norm in linguistic usage is found at the semantic and lexical levels as well, with years and certain kinship terms generally expressed as English-origin nouns.

4. Discussion

While some of the evidence here suggests that Singletons behave grammatically like English, and Diffuse nouns pattern like Spanish, a greater part of the evidence extends beyond the question of grammatical integration: the distributional patterns show that English-origin nouns are mostly different not because they function like English nouns in English discourse, but rather because they serve specific, locally determined discourse functions that have been conventionalized within this community. For instance, the semantic domains found to be most open to English-origin nouns included both those traditionally expected, such as technology, and those generally thought to be unborrowable (Smead, 2000, p. 282), such as kinship terms. In fact, the nouns in latter group were the most prevalent overall. In terms of discourse function, both Singletons and Diffuse English-origin nouns were found to be used more often to create verbal compounds, a trend not shared in either English or Spanish. While Singletons showed similar rates to English for classifying roles, diverging from Spanish and Diffuse nouns, this effect was due to different underlying causes: Singletons are often used to classify individuals according to their occupation, while in English expressions like *that is the thing* are common.

Utilizing the comparative method (Poplack & Meechan, 1998) to examine patterns of English-origin nouns and their Spanish-origin counterparts in a corpus of New Mexican speech, this study has examined various semantic and discourse factors relevant to the occurrence of English-origin nouns in Spanish discourse. In general, two trends were revealed. First, Singleton nouns were found to differ in some instances from Spanish and Diffuse English-origin nouns, aligning instead with English; this was not the case with Diffuse English-origin nouns. Second, community norms were revealed through certain contexts in which English-origin items distinguish themselves distributionally from both English and Spanish, illustrating the tendency to use certain types of English-origin nouns in certain discourse-functional contexts in this community.

In the case of determiner patterning, it was found that the English-origin nouns, when compared with Spanish nouns, were more likely to occur with an indefinite article or as bare, and less likely to occur with a definite article. These findings are different from those found in an earlier, less-bilingual corpus (with less English and less multi-word switching) from the same community, in which Singleton and Diffuse English-origin nouns patterned like Spanish (Torres Cacoullos & Aaron, 2003). In particular, here it was found that Singletons occurred with indefinite articles in the predicating function, a pattern that aligns with English and not with Spanish. This may indicate transfer in this particular construction, or one-word code-switching.

Spontaneously used other-language-origin nouns may not always behave like their recipient-language counterparts. Some of the patterns found here suggest that, at times, in some communities, they may constitute one-word code-switches (cf. Poplack & Meechan, 1995). Nonetheless, the impact is felt here not from English, but from the conventions of the local community. Among their many functions, these nouns are best suited in this community for naming kin, classifying individuals as belonging to a certain occupation, and creating verbal compounds. These results remind us that language is a local phenomenon, molded by speakers in everyday life as they build a grammar that works in their lives.

Acknowledgements

I would like to thank Meagan Day for her able assistance with some of the coding for this project. I would also like to thank Rena Torres Cacoullos and Catherine Travis for their generous input, and Naomi Shin for invaluable consultation regarding gender and code-mixing.

Funding

This work was partially supported by National Science Foundation (NSF) grant #1019112/1019122 awarded to Rena Torres Cacoullos and Catherine E. Travis to support the development of the NMSEB corpus.

Notes

1. All examples given are from New Mexico Spanish-English Bilingual corpus (NMSEB) (Torres Cacoullos & Travis, in preparation) and are reproduced verbatim from the transcripts (see transcription conventions in Appendix 2 of the Introduction to this volume). Within brackets is the recording number, name, and time stamp. In examples where Spanish and English are used, the original appears on the left, and the translation on the right.
2. In all cases, the transcription, and not the sound file, was used as the basis for analysis. This means that at times, particularly with Spanish-English cognates, the transcriber's interpretation of the production (e.g. *situación* vs. *situation*) was used to determine the linguistic origin of the word.
3. For an exploration of nominal ellipsis in Spanish from an information-structure perspective, see Eguren (2010).
4. Not all speakers produced a sufficient amount of stretches of monolingual English discourse to match the number of extracted occurrences exactly.
5. I am grateful to Chris Koops for making these frequency figures available to me.
6. As Smead (2000) notes in his study of English-origin words in Chicano Spanish, "if the underlying referent is human, biological sex always determines gender assignment" (p. 295), so these cases are excluded here.
7. Clegg (2011) suggested that a frequency-based system for teaching gender in Spanish-language pedagogy should replace those based on Bull's (1965) findings regarding terminal phonemes.

References

Banfield, R. L. (1994). *Gender assignment of English loanwords in the Spanish language*. Master's thesis, Brigham Young University.

Bills, G., & Vigil, N. (2008). *The Spanish language of New Mexico and Southern Colorado: A linguistic atlas*. Albuquerque, NM: UNM Press.

Bley-Vroman, R. (1983). The comparative fallacy in interlanguage studies: The case of systematicity. *Language Learning*, *33*, 1–17.

Bull, W. (1965). *Spanish for teachers: Applied linguistics*. New York, NY: Ronald.

Chaston, J. M. (1996). Sociolinguistic analysis of gender agreement in article/noun combinations in Mexican American Spanish in Texas. *Bilingual Review/Revista Bilingüe*, *21*, 195–202.

Clegg, J. H. (2010). Native Spanish speaker intuition in noun gender assignment. *Language Design*, *12*, 5–18.

Clegg, J. H. (2011). A frequency-based analysis of the norms for Spanish noun gender. *Hispania*, *94*, 303–319.

DuBois, J. W. (1980). Beyond definiteness: The trace of identity in discourse. In W. Chafe (Ed.), *The pear stories: Cognitive, cultural, and linguistic aspects of narrative production* (pp. 203–274). Norwood, NJ: Ablex.

Eguren, L. (2010). Contrastive focus and nominal ellipsis in Spanish. *Lingua*, *120*, 435–457.

Garcia, M. (1998). Gender marking in a dialect of Southwest Spanish. *Southwest Journal of Linguistics*, *17*, 49–58.

Hopper, P., & Thompson, S. A. (1984). The discourse basis for lexical categories in universal grammar. *Language*, *60*, 703–752.

Jake, J. L., Myers-Scotton, C., & Gross, S. (2002). Making a minimalist approach to codeswitching work: Adding the matrix language. *Bilingualism: Language and Cognition*, *5*, 69–91.

Martinez-Gibson, E. A. (2011). A comparative study on gender agreement errors in the spoken Spanish of heritage speakers and second language learners. *Porta Linguarium*, *15*, 177–193. Retrieved from http://www.ugr.es/~portalin/articulos/PL_numero15/10.%20ELIZABETH%20A%20MARTINEZ%20GIBSON.pdf

Montes-Alcalá, C., & Lapidus Shin, N. (2011). *Las keys* vs. *el key*: Feminine gender assignment in mixed-language texts. *Spanish in Context*, *8*, 119–143.

Montrul, S., & Ionin, T. (2010). Transfer effects in the interpretation of definite articles by Spanish heritage speakers. *Bilingualism: Language and Cognition*, *13*, 449–473.

Myers-Scotton, C., & Okeju, J. (1973). Neighbors and lexical borrowings. *Language*, *49*, 871–889.

Otheguy, R., & Lapidus, N. (2003). An adaptive approach to noun gender in New York contact Spanish. In R. Cameron, L. López, & R. Núñez-Cedeño (Eds.), *A Romance perspective on language knowledge and use* (pp. 209–232). Amsterdam: John Benjamins.

Pérez-Pereira, M. (1991). The acquisition of gender: What Spanish children tell us. *Journal of Child Language*, *18*, 571–590.

Pires, A., & Rothman, J. (2009). Disentangling contributing variables to incomplete acquisition competence outcomes: What differences across Brazilian and European Portuguese heritage speakers tell us. *International Journal of Bilingualism*, *13*, 211–238.

Poplack, S. (1982). Competing influences on gender assignment: Variable processes, stable outcome. *Lingua*, *57*, 1–28.

Poplack, S. (1993). Variation theory and language contact. In D. Preston (Ed.), *American dialect research* (pp. 251–286). Amsterdam: John Benjamins.

Poplack, S. (2012). What does the Nonce Borrowing Hypothesis hypothesize? *Bilingualism: Language and Cognition*, *15*, 644–648.

Poplack, S., & Meechan, M. (1998). Introduction: How languages fit together in codemixing. *International Journal of Bilingualism*, *2*, 127–138.

Poplack, S, & Meechan, M. (1995). Orphan categories in bilingual discourse: A comparative study of adjectivization strategies in Wolof/French and Fongbe/French. *Language Variation and Change*, *7*, 169–194.

Poplack, S., Pousada, A., & Sankoff, D. (1982). Competing influences on gender assignment: Variable processes, stable outcome. *Lingua*, *57*, 1–28.

Poplack, S., Sankoff, D., & Miller, C. (1988). The social correlates and linguistic processes of lexical borrowing and assimilation. *Linguistics*, *26*, 47–104.

Poplack, S., & Tagliamonte, S. (2001). *African American English in the diaspora*. Oxford: Blackwell.

Sánchez, M. F. (1995). *Clasificación y análisis de préstamos del inglés en la prensa de España y México*. Lewiston, NY: Edwin Mellen.

Smead, R. C. (2000). On the assignment of gender to Chicano Anglicisms: Processes and results. *Bilingual Review/La Revista Bilingüe*, *25*, 277–297.

Smith, P., Nix, A., Davey, N., López Ornat, S., & Messer, D. (2003). A connectionist account of Spanish determiner production. *Journal of Child Language*, *30*, 305–331.

Snyder, W., Senghas, A., & Inman, K. (2001). Agreement morphology and the acquisition of noun-drop in Spanish. *Language Acquisition*, *9*, 157–173. Retrieved from http://www.jstor.org/stable/20011507

Swadesh, M. (1971). *The origin and diversification of language*. Edited post mortem by Joel Sherzer. Chicago, IL: Aldine.

Teschner, R. V. (1974). A critical annotated bibliography of Anglicisms in Spanish. *Hispania*, *57*, 631–678.

Thompson, S. A. 1997. Discourse motivations for the core-oblique distinction as a language universal. In Akio Kamio (Ed.), *Directions in functional linguistics* (pp. 59–82). Amsterdam: John Benjamins.

Thompson, S. A., & Hopper, P. J. (2001). Transitivity, clause structure, and argument structure: Evidence from conversation. In J. Bybee & P. J. Hopper (Eds.), *Frequency and the emergence of linguistic structure* (pp. 27–59). Amsterdam: John Benjamins.

Torres Cacoullos, R., & Aaron, J. E. (2003). Bare English-origin nouns in Spanish: Rates, constraints, and discourse functions. *Language Variation and Change*, *15*, 289–328.

Torres Cacoullos, R., & Travis, C. E. (2014). Gauging convergence on the ground: Code-switching in the community. *International Journal of Bilingualism. Special issue, Gauging convergence on the ground: code-switching in the community*, C. E. Travis & R. Torres Cacoullos (eds.).

Torres Cacoullos, R and Travis, C. E. (In preparation). New Mexico Spanish-English Bilingual (NMSEB) corpus, National Science Foundation 1019112/1019122. http://nmcode-switching.la.psu.edu/.

Torres Cacoullos, R., & Travis, C. E. (2014, this volume). Introduction. Gauging convergence on the ground: Code-switching in the community. *International Journal of Bilingualism*.

Travis, C. E., & Torres Cacoullos, R. (2013). Making voices count: Corpus compilation in bilingual communities. *Australian Journal of Linguistics*, 33(2), 170–194.

Vargas-Barón, A. (1952). The function of the definite article in Spanish. *Hispania*, *35*, 410–414.

Wilson, D. (2013). One construction, two sources: *Hacer* with an English infinitive in bilingual discourse. In A. M. Carvalho & S. Beaudrie (Eds.), *Selected Proceedings of the 6th Workshop on Spanish Sociolinguistics* (pp. 123–134). Somerville, MA: Cascadilla Proceedings Project.

Zamora, J. C. (1975). Morfología bilingüe: La asignación de género a los préstamos. *Bilingual Review/ Revista Bilingüe*, *2*, 239–247.

Zamora Munnt, J. C., & Btjar, E. C. (1987). El género de los préstamos. *Revista Española de Lingüística*, *1*, 131–137.

Author biography

Jessi Elana Aaron is an assistant professor in the Department of Spanish and Portuguese Studies at the University of Florida. Her research interests include language variation and change, language contact, and usage-based models. She earned her PhD in Spanish Linguistics and an MA in Anthropology from the University of New Mexico. She also holds an MA in Latin American Studies from Stanford University.

Appendix

Table A shows the production rate of each speaker in each interview, including both the raw number of tokens of lone English-origin nouns and a standardized number, which shows how many such tokens each speaker produced per 10,000 words.

Table A. Number of Lone English-origin Nouns Produced by Each Speaker, by Interview.

Interview	Speaker	Lone English-origin nouns	Standardized Nl one English-origin nouns per10,000 words
1 El Abuelo	Susan	8	12
2 La Marina	Bartolomé	20	74
3 Dos Comadres	Sandra	53	49
5 Las Tortillas	Rocío	9	11
6 El Túnico	Ivette	136	98
7 Basketball Teams	Pedro	7	18
7 Basketball Teams	Samuel	13	49
8 Graduación Familiar	Inmaculada	4	7
9 La Salvia	Fabiola	44	66
9 La Salvia	Molly	44	114
11 El Trabajo	Mónica	101	86
13 La Acequia	Betty	1	3
14 Calcetines/Medias	Anita	35	63
14 Calcetines/Medias	Inmaculada	18	51
15 Las Cosas Viejas	Aurora	5	6
16 Trip to Africa, pt. 1 and 2	Manuel	44	51
17 La Comadreja	Javier	34	35
18 Las Minas	Francisco	174	215

The excluded nouns found in DRAE included *alumni, area, babysitter, base, baseball, basket, basketball, bingo, boom, bus, carbon, carro, chef, chequeadita, claustrophobia, closet, estaca 'stake,' football (fútbol), honor, kinder, kindergarten, lonche (lunch), magazine, material, miss, mister, mobile (móvil), nylon (nailon), panels, panties, party, pajamas, pizza, plataforme (plataforma), rails, record(s), renta 'rent,' rifle, satin, shorts (short), supervisor, test(s), tour, tractor,* and whiskey.

The excluded compounds (N=102) included: *a lot later, antique mall, antique stores, atomic bomb test, baby grade, baseball team, big star, birthday party, bus depot, business manager, camper trailer, card game, check stubs, child support, egg farm, electric heaters, first/second/third/fourth/fifth/sixth/eighth/ninth grade(s)/graders, first holy communion, fishing reports, fried egg, fun pocket, game wardens, grocery cart, hard metal, heavy equipment, ice fishing, last one, last period, lunch bag, major news, mechanic school, mid high, nylon dresses, power saw, rain gear, rock wall, roe sacks, second time, security guard, senior citizen center, senior year, sesame truck, sewing kit, shower stall, state championship, state fair, third one, third prize, tire tube, trophy fee, twenty minutes, VIN number, and Vista volunteers*.