Sankoff, David, Poplack, Shana, & Vanniarajan, Swathi. (1990). The case of the nonce loan in Tamil. *Language Variation and Change* 2(1):71–101.
Santamaría, Francisco J. (1959). *Diccionario de mejicanismos*. México: Porrúa.
Scheibman, Joanne. (2002). *Point of view and grammar. Structural patterns of subjectivity in American English conversation*. Amsterdam: Benjamins.
Silva-Corvalán, Carmen. (2001). *Sociolingüística y pragmática del español*. Washington, DC: Georgetown University Press.
Smead, Robert N. (2000). Phrasal calques in Chicano Spanish: Linguistic or cultural innovation? In Ana Roca (ed.), *Research on Spanish in the U.S.* Somerville, MA: Cascadilla Press. 162–172.
Suñer, Margarita. (1982). Syntax and semantics of Spanish presentational sentence-types. Washington, DC: Georgetown University Press.
Thompson, Sandra A. (1997). Discourse motivations for the core-oblique distinction as a language universal. In A. Kamio (ed.), *Directions in functional linguistics*. Amsterdam: Benjamins. 59–82.
U.S. Census Bureau. Percent of persons who are Hispanic or Latino (of any race): 2000, New Mexico by county. http://factfinder.census.gov/servlet/BasicFactsServlet (accessed April 4, 2003).
Vendryes, Joseph. (1925). Language. *A linguistic introduction to history*. London: Routledge & Kegan Paul.
Weinreich, Uriel. (1953). *Languages in contact*. The Hague: Mouton.

Language Variation and Change, **15** (2003), 289–328. Printed in the U.S.A.
© 2003 Cambridge University Press 0954-3945/03 $9.50
DOI: 10.1017/S0954394503153021

Bare English-origin nouns in Spanish: Rates, constraints, and discourse functions

RENA TORRES CACOULLOS AND JESSI ELANA AARON
University of New Mexico

ABSTRACT

We test the hypothesis that single other-language-origin words are nonce loans (Sankoff, Poplack, & Vanniarajan, 1990) as opposed to code-switches in a corpus-based study of English-origin nouns occurring spontaneously in New Mexican Spanish discourse. The object of study is determinerless nouns, whose status is superficially ambiguous. The study shows that, even with typologically similar languages, variable rule analysis can reveal details of the grammar that constitute conflict sites, even when relative frequencies for variants are similar. Though the rate of bare nouns is identical, their distribution patterns in Spanish and English differ. Linguistic conditioning parallel with the former, and at odds with the latter, shows that the contentious items are loanwords. In information flow terms (Dubois, 1980; Thompson, 1997), it is not lack of grammatical integration but nonreferential uses of nonce-loan nouns to form recipient-language predicates that is manifested in zero determination.

ISSUES IN THE CLASSIFICATION OF SINGLE OTHER-LANGUAGE-ORIGIN NOUNS

It is uncontroversial that borrowing and code-switching are two different manifestations of language contact (though the terminology remains far from uniform; see Muysken, 2000). Lexical borrowing is the incorporation of individual words originating in a donor, or lexifier, language into the discourse of a recipient, or host, language. Intrasentential code-switching is the mixing of two languages within the confines of a single clause. The processes involved are distinct. Borrowing involves recourse to one grammar only, that of the recipient language, whereas code-switching entails the grammars of both languages.

Heated debate persists, however, about the characterization of single content words of origin in one language appearing in a clause or constituent otherwise in another. Examples of single, or lone, English-origin nouns appearing in a clause otherwise entirely in Spanish appear in (1) through (3).[1]

We are grateful to Neddy A. Vigil for access to the New Mexico–Colorado Spanish Survey tapes. Mayra Cortes-Torres, Matt Alba, Jens Clegg, and Mark Waltemire helped with data transcription and extraction. This work was supported by a University of New Mexico Research Allocations Committee grant to Torres Cacoullos (#02-01). Work was completed during a postdoctoral fellowship for Torres Cacoullos at the University of Ottawa Sociolinguistics Laboratory, for which we thank Shana Poplack. A preliminary version was presented at NWAV-31, Stanford University, October 2002.

(1) *I'd wash the floor* de rodillas y le daba wax (117.23)
'*I'd wash the floor* on my knees and wax it' [literally: give it wax]
(2) yo arreaba gatos, arreaba loaders y trocas y todo eso (214.5)
'I drove tractors, I drove loaders and trucks and all that'
(3) y le puse complaint a ese chota (219.16)
'And I filed a complaint against that cop'

In (1), the English multiword sequence *I'd wash the floor* is followed by a code-switch at the prepositional phrase *de rodillas* 'on my knees', which initiates a Spanish multiword sequence. These multiword fragments are straightforward examples of code-switching. In (2), the English-origin single noun troca 'truck' is an indisputable case of lexical borrowing. In fact, this noun is an established loanword that manifests phonological adaptation to Spanish and satisfies the extralinguistic criteria of diffusion and dictionary attestation (e.g., Cobos, 1983; Galván & Teschner, 1989). Trocas does not constitute a switch from Spanish into English, which plays solely an etymological, not a grammatical, role. But what of wax (1), loaders (2), and complaint (3)?

Recent research has shown that lexical borrowing is productive beyond established loanwords. Nonce borrowing (Weinreich, 1953:11) involves single lexical items, mostly nouns. Unlike established loanwords and like code-switches, nonce loans are neither recurrent nor attested in dictionaries and require some degree of bilingualism. Like established loanwords, however, they are grammatically indistinguishable from native words (Poplack & Meechan, 1998a; Poplack, Sankoff, & Miller, 1988; Sankoff, Poplack, & Vanniarajan, 1990).

Nonce borrowing sometimes involves phonological integration. Nevertheless, we cannot rely on phonetic criteria to settle the status of a word like wax, pronounced with the final consonant cluster, as opposed to complaint, realized with a final [n] in accordance with Spanish phonotactics. The correlation between phonological adaptation and grammatical integration or social acceptance is an empirical question for each speech community. Yet there is substantial consensus that phonology is not a reliable gauge for distinguishing between borrowing and code-switching. In addition to inter- and intraspeaker variability, it may be that in a bilingual community the phonetics of the contact varieties are similar (Sankoff et al., 1990:73) or that speakers have borrowed a word together with its phonemes (Poplack & Meechan, 1998a:134). In any case, there is evidence that, more often than not, single other-language-origin words retain their original phonology (Jake, Myers-Scotton, & Gross, 2002:75–76).

In Spanish and English, as in other contact situations, the most frequent manifestation of language mixing is precisely single other-language-origin content words such as wax, loaders, and complaint (Jake et al., 2002:72; Poplack & Meechan, 1998a:127). Because such words are so prevalent in bilingual corpora, their classification is of utmost consequence for code-switching theories.

The distinction between nonce borrowing and code-switching is implicated in the kind of relationship that code-switching models posit between the languages in contact, whether one of asymmetry and insertion or one of juxtaposition (see

Poplack, 2001). When lone items are considered code-switches, one language is seen as dominant in providing the "matrix" into which other-language-origin items are "embedded" (e.g., Myers-Scotton, 1993). When lone items are delimited from multiword sequences, a number of empirical observations, including the monolingual well-formedness of the alternating fragments in each language, indicate that both languages play a role. Thus, Poplack (1993:255) defined code-switching 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."

Despite the mounting empirical evidence from different language pairs that single other-language-origin words are overwhelmingly nonce borrowings (see Blas Arroyo & Tricker, 2000; Poplack & Meechan, 1998b), lone English-origin nouns in Spanish discourse might be code-switches nonetheless, if only single-word fragments. In many instances, determiners, which like other closed-class (or function) words are seldom borrowed, can indicate the language (grammar) of the noun phrase (Sankoff et al., 1990:77, 93). For Spanish and English, however, the mere presence of a determiner might be insufficient for diagnosis. The equivalence constraint (Poplack, 2000; Sankoff, 1998) states that the boundary between the alternating language fragments that comprise code-switching occurs where the word order is homologous. This means that for Spanish and English, switching is permitted between a determiner in one language and a noun in the other, because these elements are ordered the same way in both languages. For example, in (4), the Spanish definite article los precedes the noun grades, which initiates a multiword English sequence, clearly a code-switch to English. It is not impossible, then, that (5) might represent a code-switch within the NP, between the article las and the noun beauticians.

(4) pero, **los** *grades are so bad* también, oiga (088.10)
'but the *grades are so bad* too, you know'
(5) lo dejó *because* no hacen dinero **las** beauticians (318.37)
'she quit it *because* beauticians don't make money'
[literally: ... don't make money the beauticians]

How to distinguish between borrowing and code-switching for single other-language-origin nouns? The case of beauticians (5) actually affords clear indices of grammatical integration. One is the postverbal position of this English-origin noun in subject role, compatible with Spanish, but not English, word order. Another is the use of a definite article, in consonance with Spanish marking of generic uses. Determinerless nouns, however, may present no superficial clues.

Given the inherently ambiguous appearance of cases such as wax, loaders, and complaint in (1)–(3), we cannot classify tokens on a case-by-case basis. When this is attempted, one researcher's judgments are susceptible to another's contestation, which can in turn be challenged, and so on. The noun in "era como [null] foreign language," for example, was judged in Jake et al. (2002:80–81), to require a determiner in standard Spanish; no scientific advances are gained if we were to disagree (or agree).

Only quantitative studies of distribution patterns can reveal the status of lone English-origin nouns. We use the variationist comparative method (Poplack & Meechan, 1998a). Comparative, because to determine whether English-origin nouns surrounded by Spanish words are behaving grammatically like Spanish as opposed to English nouns requires comparison with the source Spanish and English varieties going into the bilingual mix. Variationist, because these varieties, as all spoken language, are characterized by structural variability. It is the inherent variability of the contentious elements that can be used to disambiguate their status. If they show statistical parallels with recipient-language nouns, they are borrowings. A requisite methodological tool is the conflict site, "a form or class of forms which differs functionally, structurally, and/or quantitatively across comparison varieties" (Poplack & Meechan, 1998a:132; Poplack & Tagliamonte, 2001:101). In other words, shared or parallel structures occurring with the same frequency in both languages are inadequate for evaluating the grammatical integration of the contentious nouns. It is necessary to identify divergent structures in which the marking of comparable functions is distributed differently in each language.

Even though Spanish and English are typologically very similar, noun determination is to some extent a conflict site, as comparison of example (5) with its English gloss suggests. In this study, we examine patterns of distribution of determinerless English-origin nouns in New Mexican Spanish discourse, which we will refer to as bare or zero forms (see Dubois, 1980:212). We assess their status by considering not only rates of occurrence but also constraint hierarchies of conditioning factors, through variable rule analysis (Rand & Sankoff, 1990; Sankoff, 1988). Although the status of each and every bare English-origin noun cannot be determined on a case-by-case basis, quantitative methods provide statistical evidence of systematic tendencies in the bilingual group studied. If the conditioning of bare single English-origin nouns is parallel to that of Spanish bare nouns and also different from English, we can conclude that, because only Spanish grammar is involved, the English-origin words are loans, in the aggregate.

This study is the first to use variable rule analysis to identify conflict sites and to submit single other-language-origin words to this kind of multivariate analysis. This is also the first test of the nonce loan hypothesis (Sankoff et al., 1990) for Spanish–English contact in the United States Southwest. We find that the group of words that includes complaint and beauticians are borrowings clearly distinguishable from code-switches. These single English-origin nouns follow Spanish, not English, patterns of zero marking. In contrast, bare nouns in multiword English fragments follow English, not Spanish, distribution patterns. The former (nonce loans) are grammatically identical to their established counterparts and native Spanish nouns, whereas the latter (code-switched English fragments) are grammatical by English patterns. The disparate behavior of single items and multiword sequences belies code-switching theories that rely on data amalgamating these two classes. A further new finding on Spanish and English code mixing strategies concerns two-word adjective + noun combinations: patterns of determiner marking reveal these to be compound borrowings.

In terms of information flow categories in a discourse-based approach to grammar, zero determination in single English-origin nouns follows from nonreferential uses of nouns as part of a predicate or to serve a classifying function (Dubois, 1980; Thompson, 1997). To anticipate our final discussion, wax (1), loaders (2), and complaint (3) are not arguments but function as predicating nouns that form intransitive predicates with semantically weak support verbs. We also find nonce loans disproportionately serving a classifying function as predicate nominals with the class of nouns designating occupations or social status. It is not lack of grammatical integration, but these nonreferential uses in recipient-language predicates, that is manifested in bare nonce-loan nouns.

The remainder of the article is organized as follows. First we describe the bilingual speech community and corpus and then present the lone English-origin noun data. These comprise frequent lexical types, which turn out to be overwhelmingly established loanwords, and lexical types occurring only once in the corpus, which are the candidates for nonce borrowing. The comparison data, monolingual (or unmixed) Spanish and English, as well as code-switched multiword fragments, are presented. Next we show the distribution of determiners across the comparison groups. Spanish and English differ in the proportion of definite and indefinite marking, but not in the relative frequency of bare nouns. We consider the linguistic conditioning of bare nouns, in terms of discourse, semantic, and structural factors. Though Spanish and English largely coincide, a few small grammatical areas constitute conflict sites. This enables evaluation of the behavior, whether Spanish or English, of the single English-origin nouns. Finally we discuss the discourse functions, in terms of information flow properties, in which nonce loans may be preferentially used.

THE COMMUNITY AND CORPUS

New Mexico is the oldest Spanish-speaking area in what is presently the United States. After explorations, expeditions, and colonization endeavors in the 1500s, Santa Fe was established as (Spanish) provincial headquarters in 1610. Permanent settlement by Spanish speakers dates to the 1693 reconquest of Santa Fe. From the remote colony of New Spain, to the outskirts of independent Mexico, to part of the United States since 1848, northern New Mexico has been one of the most isolated areas of the Spanish-speaking world, Santa Fe lying 1500 miles (2400 km) from Mexico City. New Mexican Spanish is said to retain characteristics of 16th and 17th century Spanish (see Lipski, 1994:281), though many settlers of "La Nueva México" had been born in the New World and brought with them a language already evolving from peninsular Spanish varieties (Bills & Vigil, 1999:43). With the introduction of the railroad and public school system in the 1880–1890s, the influx of English speakers increased rapidly and the shift to English has been documented in a number of studies (e.g., Bills, Hernández Chávez, & Hudson, 1995). A more recent contribution to the language mix in New Mexico is from accelerated immigration from Mexico, especially in the last decade.

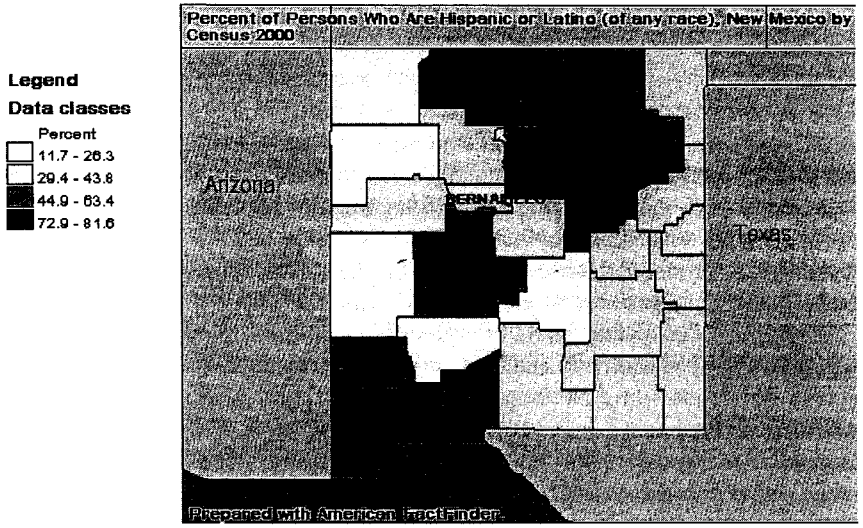

FIGURE 1. Percent of persons who are Hispanic or Latino (of any race): 2000 New Mexico by county (U.S. Census Bureau).

Data from the 2000 census indicate a continued strong presence of Spanish in the area. In Bernalillo county, which includes Albuquerque, 25% of the population speaks Spanish at home, with 8% of this group speaking English less than "very well." Corresponding figures are higher in counties to the north: Mora, 68% and 24%; Río Arriba, 60% and 16%; Taos, 46% and 12%. The Hispanic or Latino population in these counties is shown in the map in Figure 1 (U.S. Census Bureau).

New Mexican Spanish is one of the most studied varieties, beginning with the pioneering work of Aurelio M. Espinosa in the early 1900s (Espinosa, 1914–1915, 1975). Nevertheless, previous studies of Spanish–English language mixing in the southwestern United States have relied on subjective lists of words or on dictionary listings. Though such studies offer some clues, they can only provide indirect evidence about bilingual processes. At the moment when speakers use a word from one language in the discourse of another, is the linguistic process one of integration into the recipient language (borrowing) or one of maintaining the structure of the donor language (code-switching)?

In this study we draw on a corpus of recorded sociolinguistic interviews, during which English-origin words appeared in the spontaneous discourse of the speakers. The New Mexico–Colorado Spanish Survey (NMCOSS) comprises interviews with 355 Spanish-speaking people who are native to New Mexico or Southern Colorado, carried out in the early 1990s with the goal of producing a linguistic atlas (Bills & Vigil, 1999). Demographic information for the sample of interviews in the present study appears in Table 1. The speakers are primarily

TABLE 1. *Speaker characteristics**

Interview Number	Sex	Date of Birth	Age at Interview	Formal Instruction	Former or Current Occupation	City or County
004	F	1911	80	3 years	rancher	Rio Arriba Co.
010	F	1904	88	two winters	housekeeper	Española
020	F	1910	82	book 5	cook, sheep shearing	Taos Co.
047	F	1939	54	8 years	housewife	Albuquerque
076	F	1918	74	book 8	bus driver	Mora Co.
088	M	1947	45	12 years	truck driver	Albuquerque
102	M	1944	48	10+ years	wood artist	Taos Co.
117	F	1930	62	8 years	cook	Bernalillo Co.
142	F	1912	80	16 years	teacher	Taos Co.
144	M	1909	83	16 years	pastor	Rio Arriba Co.
147	F	1926	66	17 years	teacher	Taos Co.
156	M	1920	73	10 years	?	Rio Arriba Co.
190	M	1923	71	12 years	mechanic/service manager	Tucumcari
214	M	1917	76	book 4	truck driver	Mora Co.
219	F	1897	96	3 years	weaver	Rio Arriba Co.
270	M	1907	87	12 years	?	Catron Co.
311	M	1903	90	book 7	highway dept.	Rio Arriba Co.
318	F	1921	72	book 10	?	SW Colorado
M01	F	1953	48	12 years	library	Mora Co.
M02	M	1951	50	16 years	teacher/principal	Mora Co.
A01	F	1911	90	nursing school	health worker	Albuquerque

*Speakers are from NMCOSS (Bills & Vigil 1999).

residents of Río Arriba, Taos, Mora, and Bernalillo counties, between 45 and 96 years old, with 2 to 17 years of school, 12 women and 9 men.

These interviews were chosen to include speakers who regularly and/or spontaneously use Spanish in everyday interactions and who can provide a base of comparison for future studies of younger speakers; additionally, the sample was restricted to interviews with large amounts of relatively casual natural speech. This is reflected in the occurrence of code-switching, which is common in everyday speech in New Mexico. Of the 21 speakers in the present study, seven code-switch 1 to 4 times and eight do so 5 or more times in the transcribed portions of the interview. Isolated English-origin nouns, the object of the present study, appear throughout, though their frequency is not uniform across speakers. Six speakers had 12–20 tokens each, eight had 21–50, two had 51–100, and five speakers had more than 100 (the amount of material per speaker varies as well, however). The data analyzed were extracted from a little over 200,000 transcribed words, corresponding to approximately 20 hours of recorded material.

SINGLE ENGLISH-ORIGIN NOUNS

We extracted all of the single (lone) nouns of English origin, that is, those preceded and followed by Spanish words.[2] For example, we included the occurrence of grades in *buenos grades sacaba* (6a), but not the one that initiates the sequence of English words *grades are so bad* (6b), which we consider part of a multiword alternation, or intrasentential code-switch; similarly, we included the occurrence of town surrounded by Spanish words in (7a), but not the one that occurs in the multiword fragment in (7b).

(6) a. buenos grades sacaba (088.8)
'he got good grades'
b. pero, los *grades are so bad* también, oiga (088.10)
'but the *grades are so bad* too, you know'

(7) a. mi grandpa vivía en un town pobre (041.4)
'my grandpa lived in a poor town'
b. los ojos están acá *about eighteen miles from the spring, from the town* y el agua llega allá (270.19)
'the eyes (springs) are here *about eighteen miles from the spring, from the town* and the water reaches there'

Included as single nouns are two-word units that have a separate dictionary entry, for example, ice cream, high school, and business man. Other two-word adjective + noun combinations not attested in English dictionaries, for example, ringer-type machine (8) were kept apart and will be treated in a following section. Also set aside were the few cases of English-origin nouns with English determiners otherwise surrounded by Spanish, as in (9). Such examples were rare, with a total of 14 tokens, concentrated in one speaker (318, 8/14), who also presented a fair amount of intrasentential multiword code-switching.[3]

(8) yo tenía una ringer-type machine (117.6)
'I had a ringer-type machine'

(9) ahí están the graves que nos enseñaron (190.6)
'over there are the graves that they showed us'

Excluded from further study here were loan translations (Otheguy, 1993; Weinreich, 1953), or phrasal calques (Smead, 2000), such as escuela alta 'high school', literally 'school high', since by definition these are integrated at all levels, phonological, morphological, and syntactic (Silva-Corvalán, 2001:311). They also happen to be relatively infrequent in this speech sample. For example, escuela alta was used by three speakers a total of four times, while high school (with varying degrees of phonological adaptation) appeared in three times as many speakers and with five times as many tokens, 9 and 21, respectively. Also excluded were proper nouns, because they may be treated differently than common nouns in processes of integration (Poplack et al., 1988:99, note 8) and in determiner distribution (Alarcos Llorach, 1972:175). These comprise a

TABLE 2. *Single English-origin nouns* (N = *1071*)

I.	Words of English etymology (first used in English), with (a) or without (b) a non-English-origin equivalent (*N* = 875)
	a. ahora camina la gente en los puros boses (010.2) 'now people always get around by bus'
	b. ahora les tienen televisión (MO2.6) 'now they have television for them'
II.	Cognates used with their primary English meaning (*N* = 91) yo me estuve en la en la escuela [...] al grado tres no más llegué (219.2) 'I was in school [...] I just got up to the third grade'
III.	Cognates used with English phonology (*N* = 105) estaba todavía muy disturbed, con su separation (017.4) 'she was still very disturbed, with her separation'

substantial number (*N* = 336) of place names (Truchas Peak), stores (Wal-Mart), hospitals (Lovelace), newspapers (Taos News), and institutions or organizations (English Plus).

Table 2 shows the three kinds of single English-origin nouns in the corpus, within the parameters delimited above. Of a total of 1072, the vast majority (82%) are words of English etymology. Some of these have different equivalents in New Mexican or other varieties of Spanish, for example, troca 'pick-up truck' is also camioneta; sute 'suit' is also traje or vestido (Bills & Vigil, 1999:49); and bos 'bus' is also camión (e.g., Mexico), guagua (e.g., Puerto Rico), or autobús (e.g., Spain). Others have no alternative that is not of English origin, for example, televisión, fútbol, rifle. English origin was checked in the *Diccionario de la Real Academia Española* (DRAE) and in Corominas's *Diccionario crítico etimológico castellano e hispánico*.

Verifying English origin is not always straightforward. In some cases the word is attributed to Greek and/or Latin, as in the case of televisión, from tele- and vision. In cases like these we consulted the *Oxford English Dictionary* to determine if the word was first used in English with the meaning that appears in the corpus. Benveniste (1974:164–165), for example, made the case that microbe is French, not Greek, and Vendryes wrote (1925:194) that flert < fleurette was borrowed by French from English. Strict adherence to this criterion led us to include televisión and aeroplano, for example, but exclude helicóptero and tractor, which were first used in French. For varieties of Spanish spoken in the United States, determining English origin is even thornier. It is likely that some words entered New Mexican Spanish directly from English, rather than from French or through transmission from other varieties of Spanish. In the absence of studies documenting first attestations in this variety, we kept a handful of words associated with U.S. institutions even though they are not ascribed English origin and are listed in the DRAE with a meaning encompassing the one in the corpus, for example, receso (school recess).[4]

(10) a. no nos dejaban salir al receso (147.3)
'they wouldn't let us go out to recess'
b. te daban unos cinco o seis minutos para que fueras al recess (M02.6)
'they gave five or six minutes to go out to recess'

Far less numerous than the nouns of English etymology, but also counted as English-origin nouns, are cognates from a common ancestor (e.g., Latin) if they were used according to their main English meaning. An example is grado 'grade, year in school'. This kind of word has traditionally been characterized as a semantic extension (Weinreich, 1953), in which the process involved is one of meaning change in an existing recipient-language form. Equally valid is the alternative hypothesis that considers words such as grado to be borrowings, taken with their meaning and form directly from English. In fact, the same lexical item appeared without phonological adaptation, as in (11), which also shows the Spanish equivalent in this variety for 'grade', *libro* 'book'.

(11) no me acuerdo, cuando estaba en esos grades, cuando me me acuerdo es cuando estaba en el libro ocho (318.14)
'I don't remember, when I was in those grades, when I I remember is when I was in book [grade] eight'

Words that elsewhere may be classified as semantic extensions, such as colegio 'college,' grado 'school grade or mark', principal 'school principal', make up 8% of the single English-origin noun data.

The third kind of English origin noun is cognate nouns that were produced with English phonology, even though their etymology may not be English and their main meaning may be the same in Spanish and English. An example is separation. These add up to another 10% of the data.

The distribution of these noun types in the corpus is uneven. Some occur only once, whereas others are recurrent and spread over several speakers. Two different kinds of frequency, diffusion and use frequency, can be measured. Diffusion is based on the number of speakers using the lexical type, whereas use frequency is based on the number of occurrences. The two measures do not always coincide. For example, chansa 'chance' has a high diffusion (5 speakers), but a relatively low use frequency (6 occurrences), compared to grandma (5 speakers, 26 occurrences).

Assigning frequency counts to the noun types in the database entails the prior step of deciding which tokens to categorize as the same lexical type. Some tokens demonstrate phonological integration, others do not. For example, in (12), (13), and (14), the (a) tokens, with an added final vowel, are more phonologically integrated than their (b) counterparts.

(12) a. perdieron algo de del lenguaje español, pero hablan muy bien (020.8)
'they have lost something of the Spanish language but they speak very well'
b. están perdiendo su mismo language de ellos (117.39)
'they are losing their very language'

(13) a. malditos papeles *that will tell you* tú usas el sute y yo aquí ... (088.9)
'damn papers *that will tell you* you use the suit while I here ...'
b. pero *you just stood a chance* de ganar un suit de, digamos ... (190.8)
'but *you just stood a chance* to win a suit of, let's say ...'
(14) a. yo pongo el teléfono junto a la silla (147.6)
'I put the telephone by the chair'
b. estaba puesto el, el telephone mal (117.30)
'the telephone was set wrong'

Inter- and intra-individual variability in phonological adaptation has been found to depend on the degree of social integration (frequency of use and age of attestation) of the lexical item and also on speakers' English proficiency and neighborhood of residence in Canadian French (Poplack et al., 1988:70–75). Social factors in the phonological integration of English-origin items in New Mexican Spanish deserve a separate study. Nevertheless, we list these examples of "intralexical" variability to underscore the discordance between phonological and grammatical integration. In all three, the nonphonologically adapted nouns of English origin are preceded by Spanish determiners; in (12b) the possessive duplication construction (literally: their same language of theirs) is a feature of Mexican Spanish (Company, 1995); and in (14b) postverbal position of the subject (literally: was placed the telephone wrong) is contrary to English, but compatible with Spanish, word order.[5]

Forms with the same meaning were coded as the same lexical item, or lexical type, even if they demonstrated differing degrees or manner of adaptation to Spanish phonology, for example, lenguaje and language, sute and suit, teléfono and telephone. We also coded as manifestations of one lexical type all Spanish derived forms, for example, troca 'truck', troquita 'truck + -DIMIN', and troquero 'truck + -er, that is, truck driver' (see Poplack et al., 1988:99, note 9). On the other hand, forms with independent entries in English dictionaries, such as TV and television, including those involving English derivational morphemes, such as dad and daddy, were coded as different lexical types. Also coded as separate lexical items were identical forms with different meanings, such as yarda (lawn) and yarda (measurement).

In the total of 1071 tokens, we identified 461 different lexical types. Table 3 shows the spread and frequency of these single English-origin nouns. "Widespread" are those lexical types that were used by at least two speakers, whereas "singletons" are those that only occurred once in the entire corpus.[6] A third group is "idiosyncratic" types, those appearing more than once but used by only one speaker. As can be seen in Table 3, widespread single English-origin nouns comprise 19% of the lexical types, but 52% of all tokens. In contrast, the singleton nouns make up 61% of the lexical types, but only 26% of the data.[7]

Most occurrences of single English-origin nouns involve established loanwords. These are loanwords that have attained a certain level of social acceptance, as gauged by their attestation in dictionaries or other publications of word lists. Many of these are used in varieties of Spanish outside the United States. We consulted Mexican (Santamaría, 1959) and Peninsular (DRAE) dictionaries, re-

TABLE 3. *Diffusion and frequency of lone English-origin nouns*

	Lexical Types		Tokens	
	N	%	*N*	%
Widespread (used by two or more speakers)	87	19	560	52
Idiosyncratic (used by one speaker only)	91	20	228	21
Singleton (occurred once in corpus)	283	61	283	26
Totals	461	100	1071	100

gional dictionaries (Cobos, 1983; Galván & Teschner, 1989) and relatively early compilations of English-origin words in New Mexico (e.g., Espinosa, 1975; García, 1939; Gross, 1935; Kercheville & McSpadden, 1934), and other lists of anglicisms (e.g., Sala, 1982). In Table 4 we show that the great majority of widespread English-origin nouns, that is, those used by at least two speakers, are also established loanwords. Only one-fourth of widespread types, which add up to just 16% of the tokens, do not appear in any of the sources consulted. In contrast, 83% of the singleton lexical types is unattested. These results confirm a strong correlation between established loanword status and level of diffusion (Poplack et al., 1988:59).

Lexical need, here meaning the need to fill gaps in Spanish vocabulary for referents associated with the English-speaking context, has been cited as a motivation for lexical borrowing (Haugen, 1969; Weinreich, 1953). As reported in other studies of Spanish spoken in the United States Southwest, many English-origin nouns can be grouped into semantic fields such as those of the railroad, automobiles, and machinery (e.g., reque 'wreck'), food (e.g., espauda 'baking powder'), and institutions (e.g., social 'social security'). Nevertheless, among the highest frequency widespread lexical types are granma (5 speakers/26 occurrences), granpa (5/19), and daddy (4/16) (mom, in contrast, was used by only one speaker). These kinship terms are subject to much dialect and situational variation, such as abuelita or nana for 'grandmother'. Poplack et al. (1988:47) in their monumental study of Canadian francophones, concluded that "borrowing behavior is acquired, and not merely a function of lexical need." The widespread use of English-origin intimate vocabulary suggests that this may also apply to at least some aspects of the borrowing behavior of New Mexican Spanish–English bilinguals.

In the remainder of the article we will refer to widespread English-origin nouns as established loanwords, sometimes abbreviated as "established," based on the correlation between diffusion and dictionary attestation shown in Table 4. We will henceforth refer to singleton lone English-origin nouns as "nonce." The difference in the attestation rate is great enough to clearly demarcate these two groups. To maintain the sharp demarcation, we set aside the group of idiosyncratic types.[8]

TABLE 4. *Attestation of lone English-origin nouns in dictionaries by degrees of diffusion*

	Official Dictionaries		Regional Dictionaries		Unattested	
	Types (%)	Tokens (%)	Types (%)	Tokens (%)	Types (%)	Tokens (%)
Widespread	30	29	44	54	26	16
Idiosyncratic	9	9	21	23	70	68
Singleton	7	7	9	9	83	83

THE VARIATIONIST COMPARATIVE METHOD

The contentious items are the nonce English-origin nouns. Are they borrowings or code-switches? We implement the comparative method laid forth in Poplack and Meechan (1998a), drawing on four groups of comparison: unmixed Spanish, unmixed English, established loanwords, and code-switched multiword fragments. To resolve the status of single English-origin nouns, we rely on statistical parallels in the variability of determiner marking. The hypothesis on loanword integration states:

> If the constraints on variability of L_d-origin [donor language] forms are parallel to those constraining their L_r [recipient language] counterparts, the former can only be borrowings (Poplack & Meechan, 1998a:130).

Applying this to bare NPs:

> If the constraints on the occurrence of bare forms in nonce lone English-origin nouns in Spanish discourse, that is, those of a single occurrence in the corpus (the singletons of Tables 3 and 4), are the same as in established English-origin loanwords and in unmixed Spanish nouns, but different from nouns in code-switches to English and in unmixed English, then the nonce are behaving grammatically like Spanish, not English, nouns; hence, they are loans.

Prescriptive and formalist theoretical accounts of bare nouns, which are based on idealized notions of the language, cannot provide a benchmark for evaluating the behavior of the contentious items. For an empirical account of the variability in determiner marking, we extracted and coded samples of nouns from stretches of unmixed Spanish and unmixed English discourse from the same interviews. These monolingual samples were used to establish the baseline norms of determiner distribution in the Spanish and English of this speech community. It is only through examination of monolingual samples from the same speakers that we may determine conflict sites between the Spanish and English varieties in contact. Patterns of use of lone English-origin nouns at these Spanish–English conflict sites will then allow us to evaluate their status as either borrowed or code-switched.

Samples of the unmixed varieties were taken as follows. For Spanish, a monolingual stretch of Spanish discourse, as a rule the longest, was chosen in each

interview and the first 100 to 150 Spanish nouns in this section were extracted for each speaker.[9] Spanish intrasentential code switches and clauses surrounded by other-language-origin words were excluded, and a section was not considered monolingual if it presented more than a few instances of clause-length code-switching. For English, however, because of the dearth of monolingual English discourse in these interviews, the data come from only one-third of the speakers, with the number of tokens extracted per speaker ranging from 16 to 300. Most of the nouns occur in stretches of two or more English clauses; about 8% (63/772) of the total appears in a single clause surrounded by Spanish, or an intersentential switch, as in (15).

(15) le dije yo, 'oh tengo mucha familia *and they all want to come and pick piñon nuts, anyway*, que no te van a hacer nada en tu, tu rancho' (318.40)
'I said to him, "oh I have a lot of family *and they all want to come and pick piñon nuts, anyway*, they're not going to do you anything on your, your farm'"

We extend the multiway comparison to established loanwords, that is, the widespread lone English-origin nouns in the corpus, and to nouns in code-switches to English. The code-switching data were restricted to intrasentential switching, unevenly distributed among about half the speakers. Most are nouns in the interior of a multiword English fragment, as in (16). We expect these to pattern as in unmixed English. Coded separately were NPs occurring at the boundary between the alternating language fragments, as in (17). The equivalence constraint (Poplack, 2000) predicts switching between determiners and nouns for Spanish and English, given the homologous word order. But are patterns of occurrence of bare nouns at switch points congruent with Spanish, English, or both grammars? We return to these boundary NPs as well as two-word adjective + noun combinations later.

(16) y aquí los pobres *trying to make ends meet* (88.7)
'and here the poor *trying to make ends meet*'
(17) ¿Quieres *potato chips or something else*? (318.35)
'Do you want *potato chips or something else*?'

Once all tokens were extracted, the same exclusions were applied as for the single English-origin nouns. This means that proper nouns were excluded, as were invariable, "lexicalized" expressions, such as todo el tiempo 'all the time'. One-word responses were also set aside.[10] The data in each of the five groups of nouns is presented in Table 5.

We may summarize the predictions as follows. Established loanwords, which are both grammatically and socially integrated, are expected to exhibit similar patterns of determiner distribution to those of Spanish-origin nouns in monolingual Spanish. Nouns in code-switched multiword English fragments are expected to pattern with unmixed English. Placed in the middle of the table are the contentious nonce lone English-origin nouns. The nonce loan hypothesis (Sankoff et al., 1990) predicts that these will exhibit patterns matching Spanish and established loanword data and be at odds with English and code-switching data.

TABLE 5. *New Mexico bilingual data*

	Spanish	Established	Nonce	Code-Switch	English
Status	unmixed	loanwords	*ambiguous*	English fragment	unmixed
No. speakers	21	21	21	9	7
Tokens	1386	555	270	165	772
Predictions	same	same	different	same	

DETERMINER DISTRIBUTION

The distribution of determiners in the five sets of data is shown in Table 6. The most striking difference between Spanish and English, at the left and right extremes of the table, respectively, is the relative frequency of definite and indefinite articles (in bold). Spanish, at 38%, shows nearly double the rate of the definite article in English, at 22%. Conversely, the relative frequency of English indefinite *a(n)* (15%) is nearly double that of Spanish *un/a* (9%). These results are consistent with the greater degree of grammaticization of the indefinite article in English (Hopper & Traugott, 1993:117) and the generalization of the definite article in Spanish (Company, 1991).[11] As predicted by the nonce loan hypothesis, the nonce data, with 36% definite and 10% indefinite articles, line up with established loanwords and Spanish, on the one hand, and against code-switches and English, on the other.

Contrasts between the data sets become even more evident in the ratio of definite to indefinite NPs, grouped in Table 7. Definite NPs include those with a definite article, possessive, or demonstrative (Lyons, 1999:15–33). Added in to the definite count are complex expressions, which are either definite article or possessive-plus-numeral combinations, as in *los dos lugarcitos* 'the two places' or, more frequently, expressions with the universal quantifier *todo* or *all* and a form of the definite article, possessive, or demonstrative, as in *all those months*. Assembled under indefinite are nouns with indefinite article *un* or *a(n)*, numerals, or quantifiers, such as *unos, algunos* or *some* and *muchos* or *many, a lot of*. For English, tokens of proximal demonstratives *this* and *these* used to initiate indefinite NPs, as in (18), were added in to the indefinite count (Prince, 1981).[12]

(18) See, like these, like these ex-friends I had (117.15)

The ratio of definite to indefinite NPs, shown in the bottom row of Table 7, is at 2.9 and 2.8 for Spanish and nonce, respectively. This is nearly double the ratio for English and code-switches, at 1.6 and 1.5.

Possessives are largely responsible for the higher definite/indefinite ratio in the established loanword data (at 4.8).[13] The proportion of possessive-marked NPs across the five data sets (Table 6) may seem puzzling at first sight. Here the line-up appears to be Spanish and nonce with 8% and 6%, on one side, and established loanwords, code-switches, and English with 16%, 13%, and 14%, on

TABLE 6. *Distribution of determiners (relative frequencies) in New Mexico bilingual data*

	Spanish (*N* = 1386)		Established (*N* = 555)		Nonce (*N* = 270)		Code-switch (*N* = 165)		English (*N* =772)	
	n	%	*n*	%	*n*	%	*n*	%	*n*	%
Definite article	525	**38**	218	**39**	98	**36**	37	**22**	169	**22**
Possessive	114	8	91	16	15	6	21	13	105	14
Demonstrative	51	4	8	1	11	4	6	4	41	5
Complex	33	2	7	1	3	1	4	2	15	2
Indefinite article	122	**9**	49	**9**	27	**10**	25	**15**	115	**15**
Indefinite *this*	1						4	2	11	1
Quantifiable/number	127	9	19	3	18	7	17	10	83	11
Zero	413	30	163	29	98	36	51	31	233	30

TABLE 7. *Definite versus indefinite lexical NPs*

	Spanish (*N* = 1386)		Established (*N* = 555)		Nonce (*N* = 270)		Code-switch (*N* = 165)		English (*N* =772)	
	n	%	*n*	%	*n*	%	*n*	%	*n*	%
Definite	723	52	324	58	127	47	68	41	330	43
Indefinite	250	18	68	12	45	17	46	28	209	27
Ratio Definite/Indefinite		2.9		4.8		2.8		1.5		1.6

Note: Indefinite = marked by article *un/a—a(n)* in the singular, quantifier (*unos, algunos, muchos—some, a lot of*), or numeral in the plural.

the other. Close scrutiny reveals that the greater proportion of possessives in established loanwords is due to the frequency of kinship terms. Granma, grampa, daddy/dad, parents, baby, and boyfriend make up 15% (82/555) of the loanword data. These lexical types account for 70% (64/91) of all loanword possessives. Though family terms also account for the greatest portion of Spanish possessives, at 54% (62/114), these make up only 4% (62/1386) of the Spanish data. In contrast, only 37% (39/105) of English possessives are family relatives. Thus, although the greater proportion of possessive NPs in established loanwords in comparison to unmixed Spanish nouns can be attributed to the disproportionate number of kinship-term loanwords, possessives in English appear to be both more frequent and differently distributed than in Spanish.[14]

Most important for our purposes, the rate of bare NPs tells us nothing about the contentious nonce items. Although as we have seen, definite–indefinite space is

divided differently in Spanish and English, the relative frequency of zero (shown in the bottom row of Table 6) is virtually identical across the four comparison data sets, at approximately 30% in both the unmixed Spanish and unmixed English data, as well as in established loanwords and code-switches.[15] Settling the status of nonce English-origin nouns requires comparison of the linguistic conditioning of bare forms. That is, beneath the identical aggregate rate of zero in Spanish and English may be hidden different constraints on the occurrence of bare forms, or different distributions by contexts. We turn to Spanish–English conflict sites in the distribution of bare NPs. We come back later to the slightly higher rate of zero (36%) in the nonce group.

BARE NOUNS IN SPANISH VERSUS ENGLISH: PINPOINTING CONFLICT SITES

Information flow properties

Although not necessarily couched in these terms, all analyses of articles and their absence in Spanish appeal to information flow parameters. Information flow, as presented by Chafe (e.g., 1994), has to do with how speakers "package" ideas as they are talking, based in part on their model of the hearer: "Information flow ... has to do, not with the content of the ideas themselves, but with their status as, for example, *given* or *new*, *thematic* or *topical*, *foregrounded* or *backgrounded*, and the like" (Chafe, 1992:215). Information flow is implicated in determiners, word order, pronominalization, case roles, and other linguistic features. Four information flow categories proposed by Sandra Thompson and John Dubois are: identifiability, activation state, discourse function, and specificity–generality (Thompson, 1997:65–70).

Identifiability and activation state refer to the cognitive status of ideas. Identifiable noun phrases are those whose referent the speaker expects the hearer is able to identify, typically through previous mention, association with another identifiable element, shared background, or presence in the extralinguistic context (Dubois, 1980; Laury, 1997:18–22). Active (or given) NPs refer to ideas that are currently in a person's focus of consciousness, as opposed to the referents of new NPs, which have to be "newly activated at this point in the conversation" (Chafe, 1994:72). Identifiability and activation state are independent. While non-identifiable referents are inevitably new, identifiable referents can be given or new (Chafe, 1994:105; Thompson, 1997:68). In (19), the NP dominio público is new, being the first mention of the referent in the conversation, but identifiable, by virtue of shared background (Chafe, 1994:94–96).

(19) ya no pudieron pastear en el dominio público (270.7)
'they couldn't graze in the public domain any more'

When the phenomenon of interest, or dependent variable, is the choice between definite or indefinite forms, however, coding noun tokens to incorporate the information flow properties of identifiability and activation state directly as

factors, or independent variables, is difficult and easily susceptible to circularity. Identifiable noun phrases tend to be marked as definite in English (Chafe, 1994:93; Dubois, 1980:217).[16] Furthermore, presumption of shared background, extralinguistic presence, or even previous mention, is not always available to the analyst (see DuBois, 1987:811–812).

The other two information flow categories, discourse function and specificity, are distinct properties often conflated in discussions of 'referentiality' in the literature. Discourse function has to do, not with the nature of an NP's referent nor with the referent's cognitive status (as, say, identifiable or given), but with how the NP is used in a given point in the discourse. One of the main roles for NPs is to track participants. These are referential, or tracking, NPs. Dubois (1980:208) put it this way: "A noun phrase is *referential* when it is used to speak about an object as an object, with continuous identity over time." In Hopper and Thompson's (1984:711) terms, referential NPs are used for "manipulable" discourse participants. In contrast, nouns used nonreferentially serve to form predicates, to orient predications, or in a "classifying" function as predicate nominals (Thompson, 1997:69).

In his classic treatment of the article in Spanish, Amado Alonso distinguished between referring to objects *qua* objects as opposed to using nouns to characterize or classify:

> "el nombre con artículo se refiere a objetos existenciales y sin él a objetos esenciales. Con artículo, a las cosas; sin él, a nuestras valoraciones subjetivas y categoriales de las cosas"
>
> 'a noun with an article refers to existential objects and without it to essential objects. With the article, to things; without it, to our subjective and categorial evaluations of things' (Alonso, 1951:162, our translation).

This scholar illustrated with the following set of examples [in (a) and (b) *hombre* 'man' is marked with the definite article, in (c) it is bare]:

a. *El hombre parecía fatigado* 'The man appeared tired'

b. *El hombre es mortal* 'Man is mortal'

c. *Hombre no es lo mismo que caballero. No es hombre quien se porta así.* 'A man is not the same as a gentleman. He is not a man who behaves thus.'

He writes: "En a) con *hombre* me refiero a un individuo del género humano; en b) al género mismo ... ; en c) *hombre* no alude al individuo, ni tampoco al género cuantitativo, sino al rango categorial, al orden, a la clase considerada cualitativamente ..." 'In a) by *man* I am referring to an individual of the human genus; in b) to the genus itself ... ; in c) *man* refers neither to the individual nor to the genus as a quantity, but to the categorial status, the order, the class, considered as a quality ...' (Alonso, 1951:161–162, our translation and glosses).

Alarcos Llorach (1972:176/1967) reiterated the basic insight. In his example, *bebo vino* 'I drink wine' without the article classifies the entity drunk, whereas *bebo el vino* 'I drink the wine' with the article identifies a concrete reality of the

class. Without the article, *vino* 'wine' is not used to speak of an object as such but combines with the verb to represent, in the words of Lapesa (1995:129), a "signo valorativo, situación o categoría social, hábito, etc." 'evaluative sign, social situation or category, habit, etc.'. [17] *Vino* is no more referential in *bebo vino* than it would be as an incorporated noun in the constructed equivalent 'I engage in wine-drinking'.

In short, these analyses converge on the claim that, in modern information flow terms, nonreferential (nontracking) uses of NPs are marked by lack of determination in Spanish.[18] What is important at this point is that discourse function may be highly isomorphous with the presence or absence of determination in Spanish and thus difficult to operationalize independently, as we found for identifiability and activation state.

Specificity seems more promising as an operationalizable factor that can be coded for independently of determiner marking. Specificity has to do with the way the NP refers. Specific NPs are used to refer to specific people or things that are not considered to be interchangeable. Nonspecific NPs refer to any member of a class of entities (see Ashby & Bentivoglio, 1993:69–70). Examples of specific and nonspecific uses of NPs, appearing with the definite article, indefinite *un* 'a(n)', or bare, are shown in (20) and (21).

(20) Specific
  a. una vez que se me quemó el generador de la troca mía (311.20)
     'once when the generator of my truck burnt out'
  b. la pusieron sus hijos en un home en Cortez (318.46)
     'her sons put her in a home in Cortez'
  c. "grandma" me decía (219.16)
     ' 'grandma' he would say to me'

(21) Nonspecific
  a. ahí no siembres, se pudre la semilla (076.7)
     'don't plant there, the seed rots'
  b. que no me fuera a subir en un aeroplanito de esos chiquitos (318.13)
     'that I wouldn't get on an airplane of those little ones'
  c. yo arrié quince años troca (214.10)
     'I drove trucks for fifteen years'

Specificity and discourse function are distinct properties. For example, the bare noun casa in *Yo compré casa en Alamogordo* (156.2) 'I bought a house in Alamogordo' is specific, because it is a particular house, but nontracking (nonreferential), because the speaker does not go on to talk about the 'house' as such but about spending winters where it is located. Compré casa is a state of the owner and here the "object" casa is just contributing to encode this state, much as the predicate adjective pretty contributes to encode a state in, for example, the house is pretty (see Chafe (1994:111) on 'have' converting a referent into a state.[19] Also specific but nontracking are the time/place expressions in (22), which serve an orienting function. Nevertheless, in language use, the two most common configurations are specific tracking NPs and nonspecific nontracking NPs, as shown in

studies implementing the information flow parameters proposed by Dubois and Thompson (e.g., Ewing, 1999).

It is not surprising then, that bare specific nouns are rare in the Spanish data, around 4%. Zero-marked specific NPs are largely restricted to vocatives, as in (20c), kinship terms used as proper nouns, as in *Mamá murió muy joven* (144.11) 'Mother died very young', or time expressions, as in *vine pa'trás en noviembre* (102.10) 'I came back in November'. In English, bare NPs make up about 10% of specific uses. Most of these are time expressions and places, as in (22).

(22) a. that's like pretty much uh next month, I think (M01.5)
b. she got hurt over there at work (117.16)

In both Spanish and English, nonspecific uses favor the occurrence of bare nouns much more than specific uses.

The distribution of the marking of generic NPs, however, is different in these languages. Most clearly recognizable as generic nouns are those in subject position in gnomic situations, which hold for all time (Bybee, Perkins, & Pagliuca, 1994:126). This is Christophersen's (1939:33) "toto-generic sense": "the whole genus everywhere and at all times" (as in Alonso's "man is mortal" example). The difference between "toto-generic" and "parti-generic" is claimed to be marked by the opposition between presence and absence of determination in Spanish (Laca, 1999:902–903). In contrast, (toto-)generic mentions can take any form in English (Chafe, 1994:102–103; Dubois, 1980:224–225; Lyons, 1999:179–198), including bare NPs. To encapsulate the difference between Spanish and English, we coded as generic those nouns used to refer to "the whole genus" or an entire class of entities, including nonsubject position (see (25)). An example of a generic use is 'beauticians don't make money' in (5), repeated here in fuller context as (23).

(23) lo dejó *because* no hacen dinero las beauticians, *and they should have your own place*, pero las que trabajan para otra tienen que darle la mitad (318.37)
'she quit it *because* beauticians don't make money, *and they should have your own place*, but the ones who work for someone else have to give her half'

Generics are a subset of nonspecific NPs. However, unlike most nonspecific nouns, they are tracking (referential). This is shown in (23), where beauticians is talked about, or "tracked," in following mentions which are pronominal (*they*, *las* 'the ones'). In the quantitative analysis, we expect generics to favor bare forms in English but not in Spanish.

Spanish–English conflict sites in the conditioning of bare nouns

Determiners and their absence in Spanish have been treated by several linguists (see Bosque, 1995; Laca, 1999). This is the first study, as far as we know, of usage patterns in a corpus of spontaneous speech. In examining the distribution and conditioning of bare nouns, insights in the literature are considered, translated into testable hypotheses, and operationalized as factors in multivariate analyses. In addition to specificity (specific, nonspecific, generic), we coded nouns for a

number of syntactic and semantic factors hypothesized to influence variation between determinerless and determined NPs that might constitute loci of differences between Spanish and English: valency role, modification, polarity, string position, number, and semantic class. We use GOLDVARB, an application for the Macintosh (Rand & Sankoff, 1990), to discover the statistical significance of these factors, as well as their relative contribution and hierarchy of effect.

Table 8 shows the results of two independent variable rule analyses of the contribution of factors selected as significant to the probability of bare nouns in Spanish and English.[20] The configuration of the data shown, after several trial runs with exclusion and collapsing of factors, achieves correspondence between the two data sets so as to facilitate comparisons.

Specificity is the factor that contributes the most to the occurrence of bare, as opposed to determined, forms in Spanish. The magnitude of effect is indicated by the difference, or range, between the highest and lowest probability weight, at 57. Specificity is also significant for English, but the hierarchy of effect is patently different in the two languages. In Spanish, nonspecific uses, as in yo arrié troca.'I drove trucks' (see 21c), are most favorable to zero (with a probability weight of .73), whereas generic and specific uses line up together in disfavoring bare NPs (with .27 and .16, respectively). In English, generic lines up with nonspecific in favoring bare NPs. The contrast between Spanish and English generic marking is illustrated for a generic in subject position (24) and as an oblique (25).

(24) a. y **la** gente no quiere pagar por tu tiempo (102.3)
'and people don't want to pay for your time'
b. you know that **0** people kill other people (M01.3)
(25) a. me trujo mula, como para **los** viejos (219.12)
'she brought me a cane, like for old people'
b. they sell everything having to do with **0** strawberries (M01.5)

Valency role, or syntactic position, shows nearly an equal magnitude of effect as specificity for Spanish, with a range of 54. We hypothesized that the syntactic position of the noun should have an effect, given the interaction between valency roles and information flow parameters (Thompson, 1997). Subjects, which tend to be identifiable, are the least favorable to bare nouns in both Spanish and English.[21] Objects are more favorable to zero than subjects in both languages. However, the asymmetry between subjects and objects in the proportion of zero forms is much greater for Spanish, at 6% and 41%, respectively, than for English, at 12% and 34%. The syntax factor group has a smaller range with respect to other factor groups in English, indicating a smaller effect than in Spanish. This set of results is consonant with the diachronic path of generalization of the definite article in Spanish, which began in subject position (Alonso, 1951; Company, 1991).

Predicate nominals and arguments of existential verbs constitute an evident conflict site. This is revealed by the constraint hierarchy, in which this factor (probability weight .76) is ordered first in Spanish as most favorable to zero, but second to last, just above subjects (probability weight .36), in English. Attributes

TABLE 8. *Variable rule analyses of the contribution of factors selected as significant to the probability of bare nouns in Spanish and English (factor groups not selected as significant appear in brackets)*[a]

	Spanish			English		
Total *n*	1386			772		
Average	30%			30%		
Corrected mean	.17			.24		
	Probability	%	% Data	Probability	%	% Data
Specificity						
Nonspecific	.73	46	61	.67	41	61
Generic	.27	7	4	.81	50	3
Specific	.16	4	35	.21	9	36
Range	57			46		
Syntax						
Predicate Nominal/Existential	.76	49	8	.36	17	15
Object	.66	41	32	.57	34	32
Oblique	.47	27	40	.57	35	40
Subject	.22	6	20	.29	12	12
Range	54			28		
Modification						
Prenominal modifier	.78	59	4	[.48]	27	20
None	.52	30	78	[.52]	32	70
Postnominal modification	.37	21	19	[.42]	20	10
Range	41					
String position						
Second noun	.73	61	5	.80	63	3
First or only noun	.49	28	95	.49	29	97
Range	24			31		
Polarity (object)						
Negated	.65	65	7	[.59]	35	10
Nonnegated	.49	41	93	[.49]	28	90
Range	16					
Semantic class						
Occupation/status	[.65]	33	3	.25	11	6
Coincidence sites	[.50]	28	92	.50	30	90
Institution	[.37]	21	5	.80	53	5
Range				55		

[a]Boxes indicate conflict sites.

of copular verbs *ser* 'to be' (and *estar de*) are more likely to be bare in Spanish than in English (26) (see Lyons, 1999:104–105), as are also the single arguments of Spanish existential *haber* or *estar* than of English *there is/there are* (27), which most frequently are indefinite forms, as in the English examples (b).[22]

(26) a. no es **0** mentira ni eso, oiga (088.1)
'it's not a lie or anything, you know'
b. I know it's **a** word (270.15)
(27) a. cada mes hay **0** baile aquí (10.6)
'every month there's a dance here'
b. and then if there's **an** accident (M01.5)

The predicate nominals in (26) are used nonreferentially to classify or evaluate the subject (Lapesa, 1995:127). We will return to predicate nominals shortly, when we look at the semantic class of occupation/status nouns. With respect to existentials, *estar* and *haber* seem to function differently. On the one hand, all *estar* tokens occurred with an indefinite article, suggesting that *estar* is used to introduce tracking NPs, as in (28). On the other hand, all cases of zero occurred with *haber*. More than half the bare *haber* tokens are plurals (13/24), and of singular bare *haber* arguments, most were either mass nouns (e.g., *hierba* 'grass') or activities, like *baile* 'dance' (on the affinity between mass nouns and plural count nouns, see Garrido, 1995:270–280). More than half of negated *haber* arguments were bare (5/9), as in (29), in support of the assertion that negation favors bare nouns in Spanish existential constructions (Laca, 1999:920).

(28) está una mina en Cuesta, ahí trabaja mi esposo ahora (M01.13)
'there's a mine in Cuesta, my husband works there now'
(29) no hay industrias (102.10)
'there aren't any industries'

Semantic class shows the greatest magnitude, or range, of effect in English, at 55. We configured this factor group by collapsing into one factor, labeled "coincidence sites," all but two of the semantic classes, occupation/status and institutions. The most notable result is the similarity between Spanish and English, manifested in coincidence sites. Preliminary variable rule analyses with greater detail in the semantic class coding indicate that, in both languages, mass nouns are more favorable than count nouns,[23] inanimate are more favorable than human referents, and time and place expressions are unfavorable to zero (except for dates in English).[24] These coincidence sites make up about 90% of the data.[25]

It is in a few small areas of the grammar that differences are revealed. The occupation/status factor is ordered above institutions in the constraint hierarchy in Spanish; the order is reversed in English. This result corroborates statements in Spanish reference grammars concerning nouns designating occupations or social status, such as *bombero* 'fireman', *curandera* 'healer', *volunteer*, in predicate nominal constructions (see Laca, 1999:914; Lyons, 1999:104). The class of abstract nouns labeled institutions includes *escuela* 'school', *gobierno* 'government', *hospital* 'hospital', *iglesia* 'church', and *pinta* 'prison'. In Spanish, the former are more favorable to zero than the latter, which favor the definite article, with a rate of 71% (41/58). The following examples illustrate these two conflict sites.

(30) (a) era **0** bombero de, de floresta, de monte. (102.11)
'he was a fireman for the forests, for the mountains'
(b) he is **a** produce manager (088.8)
(31) (a) después que entramos a **la** escuela secundaria (190.18)
'after we started secondary school'
(b) when we started **0** school we already knew how to speak English (117.10)

Of the remaining factor groups considered, modification by a postnominal adjective, adnominal PP, or relative clause is most disfavorable to bare nouns in Spanish (see Lyons, 1999:104), as is postnominal modification in English, although the effect is not statistically significant in the latter. The difference between Spanish and English lies in prenominal modifiers, which we return to in the following section. String position shows the same hierarchy in both languages, that is, nouns second in a list or conjoined with another appear to be more favorable to zero than first or only nouns. The small number of second-noun tokens (61 Spanish, 19 English) did not allow us to ascertain finer differences, though in Spanish the pattern seems to be repetition of the determiner or its absence, as in (32) (Lapesa, 1995:125).

(32) no sé dónde se conocerían, **mi** mamá y **mi** papá no sé (088.1)
'I don't know where they might have met, my mother and my father I don't know'

Also included in earlier variable rule analyses was number, because it has been claimed that plural nouns may appear bare in object position whereas their singular counterparts would likely not (Lapesa, 1995:129, 132; see Lyons 1999:104). We also expected English generic plurals, as in *maybe you can remember, what word we use here for, for buzzards* (270.13), to favor zero more than Spanish. Although configurations involving number may present different marking patterns in the two languages, in the aggregate both data sets show identical rates of bare nouns for plurals (34%), as opposed to singulars (28% Spanish, 27% English).[26] Finally, polarity also appears to have similar effects in both languages, although the magnitude of effect is greater in Spanish, where it was selected as significant. We return to the effect of negation in the following section.

In summary, underlying the identical rates of bare nouns in Spanish and English is different linguistic conditioning in a few contexts of use, made evident by different constraint hierarchies. The hierarchy of constraints, or the ordering of factor weights within each factor group, yields "the detailed structure of the relationship between variant and context, or the 'grammar' underlying the variable surface manifestations" (Poplack & Tagliamonte, 2001:94). The analysis of variation between zero and determined nouns has pinpointed conflict sites in the marking of generic NPs, predicate nominal and existential constructions, and the semantic classes of occupation and institution nouns. We can now evaluate the status of nonce English-origin nouns as loans, if they show statistical parallels with Spanish, or as code-switches, if they pattern with English.

TABLE 9. *Variable rule analyses of the contribution of factors selected as significant to the probability of zero determiner in established loanwords, nonce lone English-origin nouns, and nouns in English code-switches (Factor groups not selected as significant appear in brackets)*

	Established			Nonce			Code-switch		
Total *N*	555			270			165		
Average	29%			36%			30%		
Corrected mean	.20			.29			.22		
	Probability	%	% Data	Probability	%	% Data	Probability	%	% Data
Specificity									
Nonspecific	.73	47	55	.74	54	60	.71	43	61
Generic	.43	11	7	.11	5	8	.69	43	4
Specific	.21	7	39	.19	9	31	.16	5	35
Range	52			55			55		
Syntax									
Predicate Nominal/ Existential	.70	54	11	[.59]	49	15	.15	8	9
Object	.63	36	35	[.59]	48	35	.75	56	31
Oblique	.49	27	33	[.47]	31	37	.41	21	51
Subject	.21	4	20	[.28]	9	13	KO	0	9
Range	49								
Semantic class									
Occupation/status	.70	52	4	[.43]	35	10		30	6
Coincidence	.50	29	90	[.52]	37	86		31	89
Institution	.32	24	6	[.36]	25	4		38	5
Range	38								

TESTING THE NONCE LOAN HYPOTHESIS: CONSTRAINTS ON ZERO IN ENGLISH-ORIGIN NOUNS

Single words versus multiword sequences

Table 9 shows the results of independent variable rule analyses of the contribution of factors selected as significant to the probability of zero forms in three groups of English-origin nouns: established loanwords (established), lone English-origin nouns of a single occurrence in the corpus (nonce), and nouns in multiword English fragments (code-switch) (see Table 5).[27] Factor groups included in the analyses were those comprising Spanish–English conflict sites identified in the analysis of the unmixed data: specificity, syntax, and semantic class.[28]

The most important general result is the parallelism between the nonce and established loanword data, in discord with the code-switching data. The former (single English-origin nouns) line up with unmixed Spanish and the latter (nouns in multiword English fragments), with English.

Specificity shows the largest range for the three data sets, which indicates that this factor group has the greatest effect. In the nonce and established loanword

TABLE 10. *Postverbal subjects*

	Spanish		Established		Nonce		English	
	N	%	*N*	%	*N*	%	*N*	%
Preverbal	138	54	50	51	17	52	86	100
Bare preverbal	7	5	3	6	1	6	10	12
Postverbal	117	46	49	49	16	48	0	
Bare postverbal	9	8	1	2	2	13	NA	

data, as in Spanish, generic uses of NPs pair up with specific uses in disfavoring zero. In code-switches, as in English, generic is highly favorable to zero, pairing up with nonspecific. The generic uses in (33) show a "minimal pair" from the same speaker involving the same syntactic structure and lexical item, with a definite article as a nonce loan, in (a), and without a determiner in unmixed English, in (b). The examples in (34) show generic use of (a) an established loanword, with a determiner, and (b) a code-switched noun, without.

(33) a. el nombre que la gente usaba aquí para los, para **los** buzzards (270.13)
'the name that people here used for buzzards'
b. what word we use here for, for **0** buzzards (270.13)

(34) a. yo le tengo miedo **a los** aeroplanos, no me puedo subir (318.12)
'I'm afraid of airplanes, I can't get on one'
b. los caminos que, que cruzaban por el range *to keep* **0** *people out* (156.4)
'the roads that crossed the range *to keep people out*'

In the syntax factor group, predicate nominals and arguments of existentials are ordered first in both nonce and established loanwords as most favorable to zero, but second to last in the code-switching data, where these nouns are marked preferentially with an indefinite article. The former replicate the Spanish hierarchy, the latter the English.

Evidence of the syntactic integration of nonce single English-origin nouns in Spanish is reinforced by scrutiny of subject–verb order. Spanish subjects are claimed to more readily appear bare when postverbal (e.g., Laca, 1999:895; see also Bull, Gronberg, & Abbott 1952; Hatcher, 1957). Table 10 shows the rates of bare subjects in the Spanish, established, nonce, and English data. In Spanish, a little fewer than half of all subjects are postverbal, with a rate of zero slightly higher than that for preverbal subjects (8% vs. 5%). One group of verbs stands out, intransitive motion verbs, most frequently *venir* 'come', *llegar* 'arrive', *entrar* 'enter', in which the ratio of post- to preverbal subjects is 3 to 1 (22/6). These may function as presentatives of new referents (Sánchez Ayala, n.d.), which tend to be postverbal (Silva-Corvalán, 2001:173). Established and nonce loans show subject position distributions and bare marking rates similar to Spanish nouns. In contrast, no subjects are postverbal in the English data (or among nouns internal

TABLE 11. *Adjective-noun order and overall frequency of adjectives*

	Spanish		Established		Nonce		Code-switching		English	
	N	%	*N*	%	*N*	%	*N*	%	*N*	%
Prenominal	49	45	8	24	6	75	32	100	153	98
Postnominal	61	55	26	76	2	25	0	0	3	2
Total *N* adjectives	110		34		8		32		156	
Total% adjectives	8%		6%		3%		19%		20%	

to a code-switched English fragment). Examples (35a) and (35b) show a speaker treating a Spanish and a nonce English-origin noun the same way with respect to word order as well as marking.

(35) a. cuando entraron los catálogos (144.11)
'when catalogues came in'
b. en ese tiempo entraron los, los uhm intertubes, you know, para los carros (144.7)
'at that time intertubes, you know, for cars came in'

Adjective–noun order provides another index of nonce loan integration. Table 11 shows the relative frequency of postnominal adjectivals in the five data sets. At the extremes of the table, Spanish shows about an even split between pre- and postnominal adjectives, whereas English shows virtually categorical prenominal placement (counted as postnominal were three tokens with comparative/demonstrative *that*, e.g., *a responsibility that big, a hook like that*). Prenominal adjectivals in Spanish are quantitatively concentrated in a near-closed class of words: *buen* 'good', *diferentes* 'different-PL', *mal* 'bad', *mismo* 'same', *otro* '(an) other', *primer* 'first', *puro* 'just/only', *último* 'last'. Of these, *otro*, *buen*, *puro*, and *diferentes* favor zero determiner. Both established and nonce loans show the same collection of prenominal modifiers, as illustrated in (36). Although the established loanword data appear to have a higher relative frequency of postnominal adjectives than unmixed Spanish, close to half (11/26) of these turn out to be numbers following grado. Example (37) illustrates the difference between code-switched grade (prenominal), on the one hand, and loanword grado and native libro (postnominal), on the other.

(36) a. está otro couple aquí que no querían que se casara su hija (318.10)
'there's another couple here who didn't want their daughter to marry'
b. ahora necesitas arriba de mil pesos para los puros utilities (318.12)
'now you need more than a thousand dollars for just utilities'

(37) a. al grado tres no más llegué (219.2)
'I only got to grade three'
b. íbamos como en el libro ocho, el *seventh or eighth grade*, ya no me acuerdo (318.18)
'we were like in book eight [grade eight], the *seventh or eighth grade*, I don't remember'

TABLE 12. Tener/have *object marking, negated versus nonnegated*

	Spanish		Established		Nonce		English	
	N	%	*N*	%	*N*	%	*N*	%
Negated								
Definite	0				2	33	0	0
Indefinite	0		1	25	1	17	4	80
Bare	5	100	3	75	3	50	1	20
Nonnegated								
Definite	20	26	13	34	4	27	10	20
Indefinite	37	48	12	32	6	40	23	46
Bare	20	26	13	34	5	33	17	34
Total *N*	82		42		21		55	
Overall bare		30		38		38		33

The greater overall frequency of adjectivals in English and code-switches, at 19%–20%, compared to Spanish and loanwords, at 3%–8%, is shown in the bottom row of Table 11. This difference is at least in part due to adjective–noun compounds, such as produce manager, heavy equipment operator, head cook in English. In Spanish, hence also in nonce loans, adnominal constructions, as in (38), are more productive than compounding.

(38) que aprendieran a leer y escribir y todos los subjetos de la escuela (142.10)
'that they learn to read and write and all the school subjects' [literally: the subjects of the school]

Let us now examine negation effects. Table 12 shows the distribution of definite (definite article, demonstrative, possessive), indefinite (indefinite article, quantifier), and bare objects of 'have'. Although the overall rates of zero-marked objects of Spanish *tener* and English *have* or *got* are nearly identical, at 30% and 33% respectively, polarity reveals a sharp difference. Differences in the rate of nonnegated 'have' objects are not statistically significant across the comparison groups ($p > .3308$), but the marking of negated objects (50–100% vs. 20% bare) constitutes a conspicuous conflict site.[29] Established and nonce line up with Spanish in favoring bare negated *tener* objects (39a). The preference in English is for indefinite NPs, as in *I don't have a* _____ (39b).

(39) a. estaba en la escuela yo no tenía **0** choice (MO2.7)
'I was in school I didn't have a choice'
b. I never had **a** chance to, to advance in anything (117.15)

Another small but usable area of conflict between Spanish and English is in the marking of 'have' objects designating an item of which individuals prototypically possess only one. Garrido (1986:350–351) stated that *tener* objects are

bare in the singular when the cultural expectation is 'having' one such object, for example, a wife or a husband as opposed to sons and daughters. Example (40) illustrates this with 'pick-up truck'. Assembling such objects in the lone English-origin noun data (appliances and vehicles including tv, teléfono, troca), we find a zero rate of 59% (10/17).

(40) tráeme mis muebles, como tienes **0** troca (214.4)
'bring me my furniture, since you have a (pick-up) truck'

Finally, in the semantic class factor group, occupation/status nouns favor zero more than nouns designating institutions in established and nonce, whereas the opposite holds for nouns in code-switches (Table 9). Example (41) shows an English occupation/status noun, marked with the indefinite article and, in the same construction by the same speaker, the corresponding Spanish noun, bare. Example (42) shows an occupation/status nonce loan, occurring in a Spanish code-switch repeating the content of the preceding English fragment, with substitution of indefinite a by zero. In contrast, (43a) is an example of an institution loanword used with a definite article, contrary to the corresponding English formulation with a bare form shown in (43b).

(41) y este, *he was* **a** *teacher, you know*, fue **0** maestro y todo (144.17)
and uhm, he was a teacher, you know, he was a teacher and everything'
(42) yo me acuerdo *when I was a* cuando era **0** teenager (318.27)
'I remember *when I was a* when I was a teenager'
(43) a. cuando fui **al** high school (M02.7)
'When I went to high school'
b. yet all the years that I went to **0** school (117.14)

Before concluding with the final verdict on the status of lone English-origin nouns, we present results for flagging. Tokens with adjacent hesitations, pauses, fillers, parentheticals, or metalinguistic comments, as in (44) and (45), were coded as flagged. The overall rate of such flagged NPs is 2% for the Spanish, 8% for the English, 4% for the established loanword, and 9% (24/270) for the nonce data.

(44) las pick-ups que les dicen (144.12)
'the pick-ups as they call them'
(45) cuando nos levanta el eh ese shuttle que va por nosotros (318.7)
'when we're picked up by the eh that shuttle que comes for us'

Flagged occurrences do not account, though, for the higher rate of zero in the nonce data, shown in Table 9 at 36%. Bare flagged nonce nouns are at 33%, and the distribution of determiners among flagged tokens does not deviate from the distribution for the nonce data overall, except for a higher proportion of demonstratives, at 17% (4/24) compared to 4% (Table 6).[30] These turn out to be distal ese, which may be functioning as a flagging device itself, for example, in (45). It is important that the relative frequency of flagging of nonce loans is uneven across speakers. More than half have no occurrences and one speaker alone (M02),

TABLE 13. *Bare nouns in Spanish and English: Conflict sites*[a]

Favoring Zero	Spanish unmixed	Established loanwords	Nonce ⇒ loans	Code-Switch English fragment	English unmixed
Generic	x	x	x	✓	✓
Predicate Nominal/Existential	✓	✓	✓	x	x
Negated 'have'	✓	✓	✓	NA	x
Occupation-status	✓	✓	✓	x	x
Institution	x	x	x	✓	✓

[a]✓ favors zero; x disfavors zero.

with an individual flagging rate of 20%, is responsible for 38% (9/24) of the cases. In any case, whatever the extralinguistic correlates of flagging, there are no effects on determination.

In summary, the comparison of constraint hierarchies and further detailed analyses of Spanish–English conflict sites show that when lone single-occurrence English-origin nouns in Spanish discourse surface bare, they are following Spanish grammatical patterns. Thus, they behave like established loanwords and Spanish nouns, and unlike nouns in English multiword code-switches and unmixed English nouns, in support of the nonce loan hypothesis. Table 13 summarizes the behavior of the five comparison groups (Table 5).

Thus, New Mexican speakers seem to behave like those in other bilingual communities in integrating—at the syntactic level—single words of English origin when using them in their Spanish discourse (see Poplack & Meechan, 1998a). While degree of phonological integration may depend on frequency and diffusion, syntactic integration occurs "instantly." The difference between established and nonce loans is not so much linguistic (grammatical integration) as extralinguistic (diffusion and degree of acceptance). As Poplack and Meechan (1998a:137) concluded, borrowings, including nonce ones, behave linguistically like native elements and not like those of the language to which they belong etymologically.

Two-word sequences and code-switch "boundary" NPs

We have dwelt on the divergent behavior of single English-origin nouns and nouns in multiword English fragments. Now, what of two-word English sequences? Table 14 shows the combinations making up the two-word data in the corpus. These are overwhelmingly adjective + noun units; 70% of all tokens appear either with a Spanish determiner or bare. Our concern here is whether these two-word sequences follow Spanish or English patterns in the distribution of zero determiner.

Patterns in consonance with Spanish, and at odds with English, for generic marking, occupation-status predicate nominal marking, and subject–verb order are illustrated in (46), (47), and (48). In a variable rule analysis of the contribution of specificity and syntax to the probability of bare nouns in the data shown in

TABLE 14. *Two-word English sequences* (N = *162)*

N	*%*	
66	41	Spanish determiner + English (Adjective + Noun) ahí hice **la** basic training (102.10) 'there I did basic training'
48	30	Zero + English (Adjective + Noun) este cuate era state police (156.3) 'this guy was a state police'
17	10	English (Number + Noun) me costó fifty dollars (214.4) 'it cost me fifty dollars'
14	9	English (Determiner + Noun) trabajan all week y luego el, el domingo (318.23) 'they work all week and then on, on Sunday'
11	7	English (Determiner + Adjective + Noun) pero your SAT test y todo está más alto (088.7) 'but your SAT test and everything is higher . . .'
6	4	Other English two-word (leer in Spanish, hacer un conspiracy there, you guys los vamos a mandar)

Table 13, generic uses line up with specific in disfavoring zero, as in Spanish and unlike English. Although syntax was not selected as significant, the constraint hierarchy is identical to Spanish, with predicate nominals and arguments of existentials ordered first as most favorable to zero. Also, one-third of subjects (6/17) are postverbal. Although such two-word sequences merit more analysis, these results suggest that they are treated overwhelmingly as compound borrowings rather than code-switched English fragments.

(46) yo estoy en contra del **los** state police (219.15)
'I'm against state policemen (the state police)'

(47) su mamá de ella enseña contra las dogras en la escuelas porque es **0** PE teacher y **0** Health teacher y ella les enseña (M02.10)
'her mother teaches against drugs in schools because she is a PE teacher and a Health teacher and she teaches them'

(48) yo tenía trece años cuando nacieron los, mis twin brothers (117.26)
'I was thirteen years old when [were born] the, my twin brothers were born'

A third alternative is constituent insertion. This is the embedding in a host language syntax of a word grouping with the internal structure of the lexifier language, as found in Moroccan speakers of Arabic and French (Nait M'Barek & Sankoff, 1988). Like code-switches, constituent insertions internally match the lexifier language, but unlike the former they are not restricted to points at which the word order of the languages is homologous. Although parallel cases in Fongbe–French bilingual discourse have been identified as constituent insertions (Poplack & Meechan, 1995:222–225), the bulk of English adjective + noun sequences in

Spanish–English bilingualism is more likely compound borrowing, with a smaller portion of code-switches. First, this kind of insertional language mixing is cross-linguistically rare (Sankoff et al., 1990:98). Second, although the adjective–noun structure is English, most cases are not complete "constituents" (full NPs or DPs): 41% occur with a Spanish determiner and another 30% appear bare. Third, it is likely that many of these adjective + noun pairs, though not allotted a dictionary entry, tend to co-occur in language use, for example, emergency room, test site, lab technician. In other words, these combinations may be more fixed and fused than we might think. In any case, patterns of zero marking in English adjective + noun sequences show that it is important that they be separated from multiword fragments when assessing code-switching theories.

Finally, one other kind of English-origin noun we coded separately occurred at the boundary between the alternating language fragments, as in (49). Boundary nouns, with or without a determiner, are not as frequent as nouns internal to a code-switched fragment, occurring at about a 1:3 ratio (54 vs. 165 tokens).

(49) pero el *hawk is a scavenger bird* también *in a way*, pero no tanto (270.14)
'but the *hawk is a scavenger bird* also *in a way*, but not so much'

Boundary NPs need not violate the "equivalence constraint" (Poplack, 2000), which predicts switching between determiners and nouns given homologous word order. However, even though there is no structural impediment to switching in terms of word order, we have seen that there are Spanish–English conflict sites in patterns of determiner absence, defined by discourse, syntactic, and semantic factors. What happens at the conflict sites we have pinpointed?

Basically, boundary nouns are shirked at conflict sites. It turns out that generic uses were limited to cases where both languages coincide, as with definite singular subjects (49) (i.e., we did not find examples like the constructed pero los hawks are scavenger birds). There were a scarce four predicate nominal tokens, two occupation-status nouns, and one noun designating an institution. These appear to be used in consonance with patterns of determiner distribution shared by both languages, as with the predicate nominals in (50).

(50) ya ahora ya es ***the*** *other way around* (190.3)
'now it's *the other way around*'
una persona no puede hacer nada sola, tiene, tiene que ser ***a*** *whole bunch of, you know* (318.46)
'a person can't do anything alone, it has to be *a whole bunch of, you know*'

Only two tokens, listed in (51) and (52), match Spanish and contradict English patterns, both occupation-status predicate nominals. These may well be cases of a nonce loan followed by a code-switch, but in any case such anomalous tokens are rare.

(51) y luego la, *my youngest daughter*, ella es **0** *supervisor at the state police* (190.10)
'and then the, *my youngest daughter*, she is a *supervisor at the state police*'

(52) pero sí debe ser **0** *direct descendant of, of the Clemente* (270.10)
'but she must be a direct descendant of, of the Clemente'

The general conclusion suggested by the dearth of occurrences is that boundary NPs are avoided at determiner conflict sites. This is important, because although nonce loans conflict with the donor language, as we have seen, NPs at switch points do not conflict with either language. Amalgamating these two classes of other-language elements undermines models of language mixing.

NONREFERENTIAL USES OF NONCE LOANS

Earlier we noted that the proportions of zero are identical across the comparison data sets (Table 6). We then showed that underlying identical rates are distinct linguistic constraints, that is, the configuration and direction of effect of factors affecting the occurrence of zero is different (Tables 8 and 9). In the multiway comparisons, nonce lines up with established loanwords and Spanish, and against code-switches and English. Nonce loans, however, present a higher rate of zero, at 36%, than both Spanish and English, as well as established loanwords and code-switches, all with zero rates around 30% (though the difference achieves statistical significance only with respect to the established loanword data ($p < .05$)). Might higher rates of bare forms indicate lack of grammatical integration?

We suggest that a higher rate of bare forms is due to nonreferential uses of English-origin nouns, which combine with Spanish verbs to form Spanish predicates. Not all or even most lexical NPs serve a referential, or tracking, function. A major nontracking function is that of a predicating NP: "Predicating NPs function as part of naming a type of event, activity, or situation" (Thompson, 1997:71). This is the case when the NP forms a unit with a semantically weak verb in what Ashby and Bentivoglio (1993:67–68) call support verb constructions, in which the verb just marks tense and aspect and the predicating noun carries most of the semantic content. Chafe (1994:111–113) calls high-frequency verbs such as *have, get, give, do, make* "low-content" verbs, which convert referents into states or events. Repeated in (53) are the examples we began this paper with:

(53) a. le daba wax (117.23)
give it wax = wax (the floor)
b. arreaba loaders (214.5)
drive loaders = be a loader driver, do loader driving
c. le puse complaint (219.16)
make a complaint against = accuse

These verb + noun combinations form unitary predicates: give wax = to wax, drive loaders = to be a loader driver, make a complaint against = to accuse. The nouns 'wax', 'loaders', and 'complaint' do not refer to discourse participants that are meant to be tracked; they are not even real arguments (Thompson, 1997:72).

A scrutiny of the so-called objects we coded in the syntax factor group suggests that many, if not most, have a predicating function. Cross tabulations of the syntax and specificity factor groups show that NPs coded as objects are overwhelmingly nonspecific (60/84 = 71% in nonce, 291/399 = 73% in Spanish).

Five verb types, likely candidates for low-content support verbs, take close to half of the "objects" (42/87 = 48%) in the nonce data and a similarly large proportion of the unmixed Spanish data (172/491 = 43%): *agarrar* 'get' (4–12), *dar* 'give' (8–16), *hacer* 'do, make' (7–50), *poner* 'put' (4–12), and *tener* 'have' (19–82). Examples of English-origin as well as native nouns, forming predicates with *hacer* and *tener* appear in (54) and (55).

(54) a. para que hagan drawings y pinten (M02.11)
'so that they make drawings and paint'
b. hago escultura, trabajo madera (102.3)
'I make sculptures, I work wood'

(55) a. teníamos mucho fun (076.4)
'we had a lot of fun'
b. luego tuvo chansa de comprar (190.8)
'then he had a chance to buy'
c. cuando una gente tiene interés (318.24)
'when somebody has an interest'

In appearing bare, predicating nonce-loan nouns are behaving like their Spanish counterparts.

Another major nontracking use is in a classifying function as a predicate nominal (Thompson, 1997:69). Here we find that the distribution of the nonce data deviates from the Spanish and established loanword data. On the one hand, the relative frequency of predicate nominals is 11% (29/270) in nonce as compared to 4% (60/1386) in Spanish and 7% (37/555) in established ($p < .05$). The English data show the same relative frequency for predicate nominals at 11% (82/772). However, a large portion of English predicate nominals, 38% (31/82), appear with *it* or *this/that* subjects, as in *that's a lot of money* (270.9) (see Scheibman, 2002:145–153). Only 17% (5/29) of nonce predicate nominals appear in this kind of construction, for example (56):

(56) yo no sé si eso será prejudice (318.10)
I don't know if that's prejudice'

On the other hand, nearly half of the nonce predicate nominals are occupation/status nouns, for example, *croppers*, *supervisor*, *treasurer*, at 45% (14/29). In comparison, occupation/status nouns make up only 15% (9/60) of Spanish, 22% (8/37) of established loanword, and 24% (20/82) of English predicate nominals. This semantic class of nouns comprises 10% of all the nonce data, compared with 3–6% in the other data sets (Tables 8 and 9). These distributions suggest that the disproportionate use of nonce occupation/status loans as predicate nominals in a classifying function, as in (57), contributes to the slightly higher rate of zero. It is important, however, that in appearing bare in the occupation/status predicate nominal construction, English-origin nouns match Spanish, not English, patterns, nicely illustrated in (42), and repeated in (58).

(57) y es **0** train master en Corpus Christi (190.19)
'and he's a train master in Corpus Christi'

(58) yo me acuerdo *when I was a* cuando era **0** teenager (318.27)
'I remember *when I was a* when I was a teenager'

In summary, a discourse approach to grammar in which we consider the information flow properties of nouns shows that nonce loans may be employed nonreferentially to serve a classifying function or as part of a predicate. Bilingual speakers borrow—and grammatically integrate by zero marking—English-origin nouns in forming Spanish predicates. It is not lack of grammatical integration, but these nonreferential uses that are manifested in bare nonce-loan nouns. We hope other studies of nonce borrowing will investigate other language pairs in regard to the hypothesis that bilinguals make use of nouns from a donor language to form predicates in the recipient language.

Linguistic conditioning configurations show that lone English-origin nouns in Spanish discourse of ambiguous appearance behave grammatically like Spanish, not English, nouns, at least as far as their occurrence as bare nouns is concerned. We conclude that when bilingual speakers use single English-origin nouns they make them grammatically Spanish; they are not inserting an English structure into Spanish. The results provide strong empirical support for the nonce loan hypothesis (Sankoff et al., 1990) and belie theories of language mixing that classify single other-language-origin words (and two-word compounds) as code-switches (e.g., Jake et al., 2002).

More broadly, the study shows that, even with typologically similar languages, variable rule analysis can reveal details of the grammar that constitute conflict sites, even when rates for variants are similar. Here we have applied a discourse-based approach to the uses of nouns and also have drawn on traditional and formalist accounts of determiners and their absence. But we stress that the comparative variationist method we have adopted is theory-independent and could be brought to bear on any linguistic theory, provided its predictions can be operationalized. We have thus shown that empirical, corpus-based research, which places primacy on speaker behavior and is accountable to all the variability in the data, is capable of testing competing models of bilingual data. Only on this basis can we confidently sustain claims about the outcomes of language contact. When it is possible, with a little effort, to undertake this kind of analysis, it seems untenable to evaluate the status of items like wax (1), loaders (2), and complaint (3), on the basis of superficial appraisal on a case-by-case basis.

NOTES

1. The numbers in parentheses following examples correspond to New Mexico–Colorado Spanish Survey interview and transcription page.

2. Included in the count of single English-origin nouns were clause-initial or clause-final tokens that were preceded or followed by English words, as in

no hacen dinero las beauticians, *and they should have your own place*, pero las que trabajan para otra tienen que darle la mitad ... (318.37)

3. There was a handful of tokens of a Spanish determiner surrounded by English words, as in

ahora tienen su carro, *they have their own car* los *teenagers they have everything* (318.27).

4. Words associated with U.S. institutions that we counted as English origin are *base* (baseball), *casino*, *receso*, *superintendente* (school), *cementerio*, *funeral*. None of these appear in the Mexico City *El habla popular de la Ciudad de México* (Lope Blanch, 1976) corpus, nor in Santamaría's *Diccionario de mejicanismos*, except for *receso*, with a meaning different from that in the present corpus.

5. The lack of phonological adaptation of telephone in (14b) might seem to point to code-switch status for this token, because it corresponds to an established loanword; two-thirds (5/8) of the speakers used a phonologically adapted form. Nevertheless, Poplack et al. (1988:94) found that it was highly bilingual speakers who "show the least tendency to shed their source-language phonology" [of loanwords]. The three speakers who used English phonology with telephone were also the ones with the highest incidence of intra-sentential code-switching (117, 270, 318), which indicates a high degree of bilingual ability (Poplack, 2000/1980). Most importantly, the statistical evidence compels recognition of (nonce) loan status for individual ambiguous tokens like this one.

6. Use by at least 2 of 21 speakers may not appear sufficient to qualify a word as widespread. Poplack et al. (1988:58), in their study of 120 francophones, showed a major break between one and three speakers in the percentage of unattested English-origin loanwords, at 80% and 36%, respectively. Although the percentage drops to 18% for words used by more than 10 speakers in that study, the demarcation between one and two or more speakers is sharp enough and appropriate to our relatively small sample size, as confirmed by percentages of attested loanwords in Table 4.

7. Relative token and type frequencies were similar for Spanish and English monolingual data, that is, widespread types (used by two or more speakers) make up about 60% of all noun occurrences (412/772 or 53% in English; 878/1386 or 63% in Spanish).

8. Another reason for setting aside the idiosyncratic types (Table 3) is that one speaker (318) accounted for 25% (58/228) of all tokens; three additional speakers (10, 117, 190) accounted for another 33% (76/228).

9. Included were nominal uses of adjectives, numbers and, in the English and single English-origin noun data, gerunds. Noun–noun compounds in the English data, like *food order*, were counted only once and coded by the second noun.

10. One-word responses made up 2% (24/1071) of all single English-origin nouns in these data.

11. Alonso (1951:159–161) stated that Spanish grammars first mention *un* as an indefinite article in the 19th century, influenced by grammars of other languages.

12. We found one token of Spanish *este* used as an indefinite, the following opening of a narrative: *Esa noche estaban velando a este, a este hombre en un lugarcito y* ... (142.1)

13. The low proportion of quantifiers (Table 6) also contributes to the higher definite/indefinite ratio in the established loanword data (Table 7). The low frequency of quantifiers-numerals reflects the semantic distribution of borrowed nouns: one third of quantifiers-numerals in unmixed Spanish occur with "time" words, such as *años* 'years' and *veces* 'times' (41/127), whereas there are no such words in the established loanword data. (The one nonce token is *Para allá vamos nosotros los weekends* (156.3).)

14. Possessive-marked body parts make up the same portion of possessive NPs in Spanish and English, at 7% and 6%, respectively.

15. The 30% bare NP rate for Spanish in this corpus (Table 6) is comparable to bare NP rates of 29% in a literary text and 33% in a news article (Iturrioz Leza, 1995:369).

16. Most NPs are identifiable; Ewing (1999:135, Table 4.2) showed a scarce 2% (4/171) nonidentifiable lexical nominals in a sample of Cirebon Javanese conversational data (no pronouns or unexpressed nominals were nonidentifiable).

17. The idea that the definite article converts common nouns, which are otherwise predicates denoting concepts, into designations of objects was proposed by Frege (1962) (see Laca, 1999:894, note 3).

18. We did attempt to code tracking NPs, independently of determiner marking, as those that speakers actually tracked with a second pronominal or null mention, but ran into the fact that speakers (seem to) present NPs as tracking and then do not track them. The proportion of tracking NPs that are not tracked might be higher in certain kinds of interview data (as opposed, perhaps, to naturally occurring conversation), in which speakers are interrupted or hustled onto another topic.

19. We thank Catherine Travis, who suggested these (constructed) examples illustrating nonspecific use: *Compraba casas en Alamogordo* 'S/he would buy houses in Alamogordo' or *Compra casa y la vende* 'S/he buys a house and sells it'. We also thank Ivo Sánchez, for much help in identifying tracking versus nontracking uses.

20. Weights for factor groups not selected as significant are shown to indicate the direction of effect; these are from an analysis in which all factors are "forced into the regression" (the first stepping-down run) (Poplack & Tagliamonte 1991:310). Also included but not selected was speaker group (speakers with code-switching vs. speakers without). For Spanish, significance = .022, log likelihood = −635.251, chi-square/cell = 1.6979; for English, significance = .015, log likelihood = −393.801, chi-square/cell = 1.2135.

21. We initially distinguished between the valency roles of agent (A), subject (S), and subject of copulative verb (X), based on Dubois's (1987) Preferred Argument Structure hypothesis and Ashby and Bentivoglio's (1993) finding that, in Spanish, lexical NPs that represent new information tend to appear in subject or object roles, but rarely in agent or X roles. In addition, X roles, which can be used in definitions, should be more likely than other subjects to appear bare (Suñer, 1982:221). The numbers were too small for detailed analysis, but preposed subjects of copulas *ser* and *estar* did show higher rates of zero determiner, at 11% (4/38), than all other preposed subjects, at 3% (3/100) (see Table 10).

22. English *there is/there are* arguments were marked either with *a(n)* (7/21) or a quantifier, for example, *some, a lot of, many*, or numeral (10/21). Only one token was negated (*there was no way*, M01:3).

23. The count versus mass coding was based on lexical types, because coding by use would entail considering the presence of a determiner, our dependent variable, and number, another factor group. In the absence of an empirical study, we coded for this property by our intuitions on the primary, or most frequent, use.

24. There were too few tokens of dates ($N = 7$) in the Spanish data to test the claim that (nonabbreviated) years generally appear bare (Laca, 1999:921). Years make up a relatively high percentage (4%) of the nonce data, with a zero rate of 67%. The following illustrates the variation: *Yo me comencé en el nineteen fifty-two- fifty three ... y mi esposo nai- comenzó en 0 nineteen forty-nine* (076.1).

25. Prior to collapsing different semantic class factors into "coincidence sites," mismatches between the ordering of probability weights and marginal percentages involving occupation/status nouns showed up in GOLDVARB runs of the English data including the syntax and specificity factor groups. This is because occupation/status nouns that are predicate nominals are overwhelmingly nonspecific and none are bare. Semantic class was selected as significant in Spanish in earlier runs before collapsing count nouns, mass nouns, abstract nouns, and time and place expressions into a coincidence-sites factor.

26. Including number in GOLDVARB runs of the Spanish data, when factors were collapsed in the semantic class factor group, resulted in mismatches between weights and percentages due to interactions with specificity and syntax. It turns out that there were no bare generic or specific subjects that were plural; for singular, there was one bare generic subject in a definition (*salubridad es la persona que* ...: (A01.6) (see Suñer, 1982:221)) and one bare specific subject, a kinship term used as a proper noun (*mamá murió muy joven* (144.11)).

27. For established, significance = .043, log likelihood = −261.022, chi-square/cell = 1.1782; for nonce, significance = .000, log likelihood = −141.814, chi-square/cell = .8499; for code-switching, significance = .000, log likelihood = −79.850, chi-square/cell = 1.0976.

28. Inclusion of polarity in the nonce data revealed that negated objects were all nonspecific. Semantic class was not included in the analysis of the code-switch data because no occupation/status or institution nouns were generic or subjects.

29. Unfortunately, for the sake of comparison, we had no tokens of negated 'have' objects in the code-switching data.

30. Distribution of determiners in flagged nonce data ($N = 24$): definite article 25%, possessive 4%, demonstrative 17%, indefinite article 13%, quantifier/number 8%, zero 33%.

REFERENCES

Alarcos Llorach, Emilio. (1972/1967). El artículo en español. In E. Alarcos Llorach (ed.), *Estudios de gramática funcional del español*, 2nd ed. Madrid: Gredos. 166–177.

Alonso, Amado. (1951). Estilística y gramática del artículo en español. In A. Alonso (ed.), *Estudios lingüísticos, temas españoles*. Madrid: Gredos. 151–194.

Ashby, William J., & Bentivoglio, Paola. (1993). Preferred argument structure in spoken French and Spanish. *Language Variation and Change* 5:61–76.

Benveniste, Émile. (1974). *Problèmes de linguistique générale, II*. Paris: Gallimard.

Bills, Garland D., Hernández Chávez, Eduardo, & Hudson, Alan. (1995). The geography of language shift: Distance from the Mexican border and Spanish language claiming in the Southwestern U.S. *International Journal of the Sociology of Language* 114:9–27.

Bills, Garland D., & Vigil, Neddy A. (1999). Ashes to ashes: The historical basis for dialect variation in New Mexican Spanish. *Romance Philology* 53:43–67.
Blas Arroyo, José Luis, & Tricker, Deborah. (2000). Principles of variationism for disambiguating language contact phenomena: The case of lone Spanish nouns in Catalan discourse. *Language Variation and Change* 12:103–140.
Bosque, Ignacio (ed.). (1995). *El sustantivo sin determinación. Presencia y ausencia de determinante en la lengua española*. Madrid: Visor.
Bull, W., Gronberg A., & Abbott, J. (1952). Subject position in contemporary Spanish. *Hispania* 35:185–188.
Bybee, Joan, Perkins, Revere, & Pagliuca, William. (1994). *The evolution of grammar: Tense, aspect, and modality in the languages of the world.* Chicago: University of Chicago Press.
Chafe, Wallace. (1992). Information flow. In W. Bright (ed.), *Oxford international encyclopedia of linguistics*. Oxford: Oxford University Press. 215–218.
——— (1994). *Discourse, consciousness, and time*. Chicago: University of Chicago Press.
Christophersen, P. (1939). *The articles: A study of their theory and use in English*. Copenhague–London: Munksgaard–Oxford University Press.
Cobos, Rubén. (1983). *A dictionary of New Mexico and Southern Colorado Spanish*. Santa Fé, NM: Museum of New Mexico Press.
Company Company Concepción. (1991). *La frase sustantiva en el español medieval. Cuatro cambios sintácticos*. México: Universidad Nacional Autónoma de Mexico.
——— (1995). Old forms for new concepts: The recategorization of possessive duplications in Mexican Spanish. In H. Anderson (ed.), *Historical linguistics 1993*. Amsterdam: Benjamins. 77–93.
Corominas, Joan. (1980–1991). *Diccionario crítico etimológico castellano e hispánico, por Joan Corominas; con la colaboración de José A. Pascual*. Madrid: Gredos.
Diccionario de la Real Academia Española, www.rae.es
Dubois, John 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*. Norwood, NJ: Ablex. 203–274.
——— (1987). The discourse basis of ergativity. *Language* 63:805–855.
Espinosa, Aurelio M. (1914–1915). Studies in New Mexican Spanish. Part III: The English elements. *Revue de Dialectologie Romane* 6:241–317.
——— (1975/1917). Speech mixture in New Mexico: The influence of the English language on New Mexican Spanish. In E. Hernández-Chávez, A. D. Cohen, & A. F. Beltramo (eds.), *El lenguaje de los chicanos*. Arlington, VA: Center for Applied Linguistics. 99–114. (Originally published in H. Morse Stephens and H. E. Bolton (eds.), *The Pacific Ocean in history*. New York: Macmillan. 408–428).
Ewing, Michael C. (1999). The clause in Cirebon Javanese conversation. Doctoral dissertation, University of California, Santa Barbara.
Frege, Gottlieb. (1962/1892). Ubre Begriff und Gegenstand. In G. Frege (ed.), *Funktion, begriff, bedeutung. fünf logische studien*. Göteborg: Vandenhoeck.
Galván, Roberto A., & Teschner, Richard V. (1989). *El diccionario del español chicano/The dictionary of Chicano Spanish*. Lincolnwood, IL: National Textbook Company.
García, Trinidad. (1939). *A vocabulary of New Mexican Spanish words of English origin from southwestern New Mexico*. M.A. thesis, Stanford University.
Garrido, Joaquín. (1986). Sobre el número nominal y el artículo en español. In D. Kremer (ed.), *Actes du XVIIIe Congrès Internacional de Linguistique et de Philologie Romanes*, Tome II. Tübingen: Max Niemeyer Verlag. 346–359.
——— (1995). Sintagmas nominales escuetos. In I. Bosque (ed.), *El sustantivo sin determinacíon*. Madrid: Visor. 269–338.
Gross, Stuart Murray. (1935). *A vocabulary of New Mexican Spanish*. MA thesis, Leland Stanford Junior University.
Hatcher, Anna. (1957). Casos se han dado. *Hispania* 40:326–329.
Haugen, Einar. (1969). *The Norwegian language in America*. Bloomington: University of Indiana Press.
Hopper, Paul, & Thompson, Sandra A. (1984). The discourse basis for lexical categories in universal grammar. *Language* 60:703–752.
Hopper, Paul, & Traugott, Elizabeth Closs. (1993). *Grammaticalization*. Cambridge: Cambridge University Press.

Iturrioz Leza, José Luis. (1995). Los artículos y la operación de determinación. In I. Bosque (ed.), *El sustantivo sin determinacíon*. Madrid: Visor. 339–389.
Jake, Janice L., Myers-Scotton, Carol, & Gross, Steven. (2002). Making a minimalist approach to codeswitching work: Adding the matrix language. *Bilingualism: Language and Cognition* 5(1):69–91.
Kercheville, Francis M., & McSpadden, George E. (1934). Preliminary glossary of New Mexican Spanish. *University of New Mexico Bulletin* 5(3).
Laca, Brenda. (1999). Presencia y ausencia de determinante. In I. Bosque & V. Demonte (eds.), *Gramática descriptiva de la lengua española* (vol. 1). Madrid: Espasa-Calpe. 891–928.
Lapesa, Rafael. (1995/1975). El sustantivo sin actualizador en español. In R. Lapesa, Dos estudios sobre la actualización del sustantivo en español. *Boletín de la Comisión Permanente de Academias* XXI: 50–67. Reprinted in I. Bosque (ed.). 121–137.
Laury, Ritva. (1997). *Demonstratives in interaction. The emergence of a definite article in Finnish*. Amsterdam: Benjamins.
Lipski, John M. (1994). *Latin American Spanish*. London: Longman.
Lope Blanch, Juan M. (1976). *El habla popular de la Ciudad de México. Materiales para su estudio*. México: Universidad Nacional Autónoma de Mexico.
Lyons, Christopher. (1999). *Definiteness*. Cambridge: Cambridge University Press.
Muysken, Pieter. (2000). *Bilingual speech: A typology of code-mixing*. Cambridge: Cambridge University Press.
Myers-Scotton, Carol. (1993). *Duelling languages*. Oxford: Clarendon.
Nait M'Barek, M., & Sankoff, David. (1988). Le discours mixte arabe/français: Des emprunts ou des alternances de langue? *Revue canadienne de linguistique* 33:153–154.
Otheguy, Ricardo. (1993). A reconsideration of the notion of loan translation in the analysis of U.S. Spanish. In A. Roca & J. M. Lipski (eds.), *Spanish in the United States. Linguistic contact and diversity*. Berlin: Mouton de Gruyter. 21–45.
Poplack, Shana. (2000/1980). Sometimes I'll start a sentence in Spanish *y termino en espanol*: Toward a typology of code-switching. In L. Wei (ed.), *The bilingualism reader*. London: Routledge. 221–258.
——— (1993). Variation theory and language contact. In D. Preston (ed.), *American dialect research*. Amsterdam: Benjamins. 251–286.
———(2001). Code-switching (linguistic). In N. Smelser & P. Baltes (eds.), *International encyclopedia of the social and behavioral sciences*. Elsevier Science Ltd. 2062–2065.
Poplack, Shana, & Meechan, Marjory. (1995). Patterns of language mixture: Nominal structure in Wolof-French and Fongbe-French bilingual discourse. In P. Muysken & L. Milroy (eds.), *One speaker, two languages*. Cambridge: Cambridge University Press. 199–232.
——— (1998a). Introduction: How languages fit together in codemixing. *International Journal of Bilingualism* 2(2):127–138.
——— (eds.). (1998b). *International Journal of Bilingualism 2(2), Special Issue: Instant loans, easy conditions: The productivity of bilingual borrowing*.
Poplack, Shana, Sankoff, David, & Miller, Christopher. (1988). The social correlates and linguistic processes of lexical borrowing and assimilation. *Linguistics* 26:47–104.
Poplack, Shana, & Tagliamonte, Sali. (1991). There's no tense like the present: Verbal *-s* inflection in Early Black English. In G. Bailey, N. Maynor, & P. Cukor-Avila (eds.), *The emergence of Black English: Text and commentary*. Amsterdam: Benjamins. 275–324.
——— (2001). *African American English in the Diaspora*. Oxford: Blackwell.
Prince, Ellen. (1981). On the inferencing of indefinite *this* NPs. In A. K. Joshi, B. L. Webber, & I. A. Sag (eds.), *Elements of discourse understanding*. Cambridge: Cambridge University Press. 231–250.
Rand, David, & Sankoff, David. (1990). GoldVarb: A variable rule application for the Macintosh, Version 2. Montreal: Centre de recherches mathématiques, Université de Montréal.
Sala, Marius (coord.). (1982). *El espanol de America* (Tomo I. Lexico, Parte Primera). Bogota: Publicaciones del Instituto Caro y Cuervo.
Sánchez Ayala, Ivo. (n.d.) The Spanish presentative lexicon. Unpublished manuscript, University of California, Santa Barbara.
Sankoff, David. (1988). Variable rules. In U. Ammon, N. Dittmar, & K. J. Mattheier (eds.), *Sociolinguistics: An international handbook of the science of language and society* (vol. 2). Berlin: Waltter de Gruyter. 984–997.
——— (1998). A formal production-based explanation of the facts of code-switching. *Bilingualism: Language and Cognition* 1:39–50.