CrossMark

Bilingualism: Language and Cognition 20 (2), 2017, 283–298 © Cambridge University Press 2015 doi:10.1017/S1366728915000127

Cross-language priming: A view from bilingual speech*

CATHERINE E. TRAVIS
Australian National University, ARC Centre of Excellence for the Dynamics of Language
RENA TORRES CACOULLOS
Pennsylvania State University, ARC Centre of Excellence for the Dynamics of Language
EVAN KIDD
Australian National University, ARC Centre of Excellence for the Dynamics of Language

(Received: April 15, 2014; final revision received: March 2, 2015; accepted: March 4, 2015; first published online 3 July 2015)

In the current paper we report on a study of priming of variable Spanish 1sg subject expression in spontaneous Spanish–English bilingual speech (based on the New Mexico Spanish–English Bilingual corpus, Torres Cacoullos & Travis, in preparation). We show both within- and cross-language Coreferential Subject Priming; however, cross-language priming from English to Spanish is weaker and shorter lived than within-language Spanish-to-Spanish priming, a finding that appears not to be attributable to lexical boost. Instead, interactions with subject continuity and verb type show that the strength of priming depends on co-occurring contextual features and particular [pronoun + verb] constructions, from the more lexically specific to the more schematically general. Quantitative patterns in speech thus offer insights unavailable from experimental work into the scope and locus of priming effects, suggesting that priming in bilingual discourse can serve to gauge degrees of strength of within- and cross-language associations between usage-based constructions.

Keywords: cross-language priming, spontaneous code-switching, Spanish–English bilingual community, variable subject expression

1. Introduction

The extent to which linguistic systems of bilingual speakers overlap is a perennial issue in bilingualism research. Psycholinguistic studies have argued for a large degree of integration in the lexical and grammatical systems of bilingual speakers (e.g., Brysbaert & Duyck, 2010; Hartsuiker & Pickering, 2008). Nevertheless, in community-based studies of spontaneous bilingual speech, there is scant evidence for contact-induced morphosyntactic change in long-standing close contact situations, such as Canadian French (Poplack & Levey, 2010), and in relatively recent immigrant communities, such as Los Angeles Spanish (Silva-Corvalán, 1994). Structural priming – the tendency for language users to reinvest in previously produced or processed structures in production and comprehension (e.g., Bock, 1986) – is a phenomenon particularly suited to the study of overlap in the grammatical systems of bilingual speakers.

In the current paper we explore structural priming in bilingual discourse and, through this, the question of 'overlap' in bilingual grammars, guided by three research questions on the relative strength of within- and cross-language priming. First, are within- and cross-language priming equally strong? Second, do within- and cross-language priming persist (or decay) to the same extent across intervening linguistic material? Third, is the relative strength of priming within and across languages accounted for by a lexical boost effect, that is, a strengthening of the priming due to the identity of form of the prime and target? By addressing these questions in spontaneous speech data and with a new structure type hitherto absent from the cross-language priming literature – subject expression – we offer evidence that the strength of priming varies with naturally co-occurring features of the discourse context and the particular construction in which the target occurs.

In all varieties of Spanish, pronominal subjects exist as an option alongside unexpressed (or null) subjects in most contexts, unlike in English where variation in subject expression is much more limited (Travis & Lindstrom, 2014; Torres Cacoullos & Travis, 2015, pp. 89–90). Here, we investigate priming of Spanish variable subject pronoun expression in a new corpus of New Mexican Spanish–English bilingual speech (Torres Cacoullos & Travis, in preparation). As illustrated in (1), the first person singular (1sg) subject is variably expressed as the pronoun *yo* (lines 2, 9, 12) or is what we will call unexpressed (lines

* Though the order is not alphabetical, all authors contributed equally to this work which was made possible by funding from the National Science Foundation (1019112/1019122; http://nmcode-switching.la.psu.edu) and by the NMSEB corpus teams at the University of New Mexico and the Pennsylvania State University. We thank the editors of the volume, Gerrit Jan Koostra and Pieter Muysken, and the three anonymous reviewers for their comments on the paper.

Address for correspondence:
Room W3.17, Baldessin Precinct Building (110), The Australian National University, Acton ACT 0200, Australia
Catherine.Travis@anu.edu.au

https://doi.org/10.1017/S1366728915000127 Published online by Cambridge University Press

11 and 12). Zeros (Ø) in the original and parentheses around *I* in the translation are used to ease identification of unexpressed subjects. In the translation, unambiguous code-switches – that is, multi-word sequences originally produced in English – appear in italics.

(1) Variable subject expression in Spanish

1.	Ivette	***I** wanted to go to the night clubs y=,*
2.		*.. **yo** era la única,*
3.		*de todas las que íbamos,*
4.		*. . . you know,*
5.		*que nos juntábamos en los weekends,*
6.		*to go dancing,*
7.		*or [whatever].*
8.	Rafael	*[mhm].*
9.	Ivette	*(H) **yo** era la única que no sabía arrear.*
10.		*. . . (0.7) y luego cuando,*
11.		*. . . **Ø** quería arrear,*
12.		*__Ø__ dije pues **yo** no quiero nomás arrear allá=.*
		[06 El Túnico, 0:51:54–0:52:09][1]
1.	Ivette	*'**I** wanted to go the night clubs and=,*
2.		.. **I** was the only one,
3.		of all of us who went,
4.		. . . you know,
5.		who would get together on the weekends,
6.		*to go dancing,*
7.		*or [whatever]*'.
8.	Rafael	'[mhm].'
9.	Ivette	'(H) **I** was the only one who didn't know how how to drive.
10.		. . . (0.7) and then when,
11.		. . . **(I)** wanted to drive,
12.		**(I)** said well **I** don't want to just drive there.'

1.1 Cross-language structural priming

Community-based studies of language variation have widely reported structural priming (also referred to as perseverance, persistence, parallelism). For example, Poplack (1980) found priming in Spanish variable plural (s) expression at the level of the Noun Phrase, where the presence of [s] in a preceding element in the phrase (such as a determiner) favored its presence on a subsequent element (such as the head noun). Scherre and Naro (1991, 1992) showed that priming in variable subject/verb agreement and subject/predicate adjective agreement in Brazilian Portuguese occurred both within the clause and across clauses: morphological plural marking was more likely to occur (on verbs and predicate adjectives) in cases where it had been marked on preceding elements in the same clause or in the preceding clause. Weiner and Labov (1983) demonstrated that, in English, if there was an agentless passive in the recent discourse, speakers were more likely to use another such passive than an active clause with a general (non-specific) subject pronoun (e.g., *Lower Merion's allowed to smoke in the halls* vs. *they allow Lower Merion to smoke in the halls*).

Experimental studies to date indicate that in bilingual speakers, priming occurs both within and across languages. This cross-language priming was first observed by Loebell and Bock (2003), who reported priming of the dative alternation both from German to English and from English to German. Subsequent lab-based studies have confirmed the effect in numerous language pairs and structure types. Priming of the passive has been observed in Spanish–English bilinguals (Hartsuiker, Pickering & Veltkamp, 2004) and in Dutch–English bilinguals (Bernolet, Hartsuiker & Pickering 2009). Although much of this past experimental work has concentrated on verb argument structure alternations (see also Kantola & van-Gompel, 2011; Meijer & Fox Tree, 2003; Salamoura & Williams, 2006, 2007; Schoonbaert, Hartsuiker & Pickering, 2007; Shin & Christanson, 2009), priming has also been reported for genitive noun phrases from Dutch to English (e.g., *the shirt of the boy / the boy's shirt*) (Bernolet, Hartsuiker & Pickering, 2013) and for adjective-noun/relative clause alternations from Dutch to German (e.g., *the red ball / the ball that's red*) (Bernolet, Hartsuiker & Pickering, 2007). In addition to variation between syntactic alternations, priming has been shown for relative clause attachments from Dutch to English (e.g., *someone shot the servants of the actress who was / were on the balcony*) (Desmet & Declercq, 2006), ambiguous relative clause interpretations from English to German (Kidd, Tennant & Nitschke, 2015), and the sentential location of code-switching in Dutch–English bilinguals (Kootstra, Van Hell & Dijkstra, 2012).

The experimental crosslinguistic priming literature has typically been interpreted within Hartsuiker et al.'s (2004) lexico-syntactic model of grammatical representation, itself an extension of Pickering and Branigan's (1998) 'monolingual' model. The model builds upon previous stage-based models of sentence production (e.g., Levelt, 1989), assuming discrete layers of representation for lexical concepts (e.g., GIVE [X, Y, Z]) and for lemmas (e.g., *to give*). Each lemma is connected to a unit (or *node*) denoting its grammatical class (e.g., verb), and 'combinatorial units' denoting the syntactic combinations in which it can occur (e.g., prepositional dative, double object dative). Cross-language priming is attributed to the interconnected nature of lexical

[1] All examples given are from the New Mexico Spanish–English Bilingual (NMSEB) corpus (Torres Cacoullos & Travis, In preparation). Examples are reproduced verbatim from the transcripts, with the addition of zeros in the original (and parentheses around *I* in the translation) to ease identification of unexpressed subjects, and bolding and underlining to mark the targets and primes. Transcription conventions are presented in the Appendix. Within brackets is the recording number, name and time stamp. For more information on the corpus, see Travis and Torres Cacoullos (2013) and Torres Cacoullos and Travis (forthcoming).

https://doi.org/10.1017/S1366728915000127 Published online by Cambridge University Press

and syntactic representations in the model. Lemmas for translation equivalents (e.g., *give, geven*, for an English–Dutch bilingual) are considered to be linked to a common conceptual representation and combinatorial nodes. Therefore, following the sentence *Simon gave Mary a book*, the connections between the conceptual, lemma and combinatorial nodes are activated for both English and Dutch, predicting a higher likelihood that a speaker will produce a double object dative in both languages.

Just as lexical overlap boosts priming within the one language, translation equivalent lemmas are argued to boost priming across languages because of greater network activation. For instance, *Simon gave Mary a book* is more likely to prime Dutch *Andrew gaf Emily een pen* 'Andrew gave Emily a pen' than *Andrew stuurde Emily een cadeau* 'Andrew sent Emily a present', since in the latter case only the combinatorial node is activated. At the same time, translation-equivalent priming has been reported to be weaker than within-language lexical boost priming (Schoonbaert et al., 2007), which may suggest that priming across languages is weaker than priming within the one language (see also Bernolet et al., 2013).

Utilizing a community-based study of spontaneous discourse, in this paper we provide new insights on bilingual priming unavailable in the lab.

2. A community-based corpus and a new morphosyntactic variable for the study of bilingual priming

2.1 Community-based spontaneous speech

Experimental studies of cross-language priming provide controlled conditions for the investigation of the degree to which bilingual speakers' linguistic systems overlap, but the data to date are limited in significant ways. First, the participants have typically been late sequential bilinguals, having acquired their second language via instruction in formal education settings. Second, while proficiency is tested, information is often lacking on participants' language use or linguistic experience and thus it is unknown to what extent, or even whether, they engage in code-switching outside the lab. Finally, by their very nature, experiments neutralize the effect of a host of contextual variables that impact on language use and with which priming may interact. Lacking to date is investigation of cross-language priming phenomena in the spontaneous speech of members of bilingual communities, bilingual speakers for whom code-switching is natural and common behavior.

In light of this gap in the literature, we base the current study on data drawn from the recordings comprising the New Mexico Spanish–English Bilingual speech corpus (NMSEB, Torres Cacoullos & Travis, in preparation). This corpus is made up of 31 sociolinguistic interviews of between 30 minutes and one hour long, for a total of approximately 29 hours of speech, or 340,000 words. The 40 participants are 23 women and 17 men, between the ages of 18 and 89 (born between the years of 1922 and 1993), whose occupations include miners, ranchers and schoolteachers. NMSEB participants are bilingual in Spanish and English, regularly using both with the same interlocutor in the same domain, as observed by the fieldworkers and subsequently confirmed in the recordings (cf., Poplack, 1993, p. 254). All participants are minimally third-generation New Mexicans from northern New Mexico, home to "arguably the oldest continually spoken variety of Spanish anywhere in the Americas that has not been updated by more recent immigration" (Lipski, 2008, p. 193).

Northern New Mexico was settled by the Spanish colonizers in the 16th and 17th centuries and, for some 200 years, Spanish speakers in this area had minimal contact with speakers of other varieties of Spanish (Gonzales Berry & Maciel, 2000, p. 4; Lipski, 2008, p. 195, 202), coming into contact with English from the mid 19th century when New Mexico became a territory of the United States. Today Hispanics represent as much as 80% of the population in some northern counties, though there is a trend to long-term language shift to English (Bills & Vigil, 2008) as just 60% of Hispanics in the state report speaking Spanish (U.S. census, 2010 American Community Survey one-year estimates).

All participants and fieldworkers in our study were northern New Mexican Hispanic Spanish–English bilinguals. With the aim of obtaining informal, everyday speech, the interviews were conducted by New Mexican students of the University of New Mexico, who were in-group members of these northern New Mexican communities (cf., Clyne, Eisikovits & Tollfree, 2001, pp. 235–236; Poplack, 1993, p. 260). The fieldworkers recorded conversations with family members and acquaintances, utilising a loosely structured set of topics to elicit narratives of personal experience, in which the monitoring of speech is minimized (Labov, 1984, pp. 32–42). The fieldworkers were instructed to speak in both English and Spanish, switching between the two as they naturally would, in accordance with "the appropriate code for the Hispano community" in New Mexico (Gonzales, 1999, p. 29). In this way, the language switching and linguistic structures that occur in the speech of the participants in the corpus arose spontaneously, not in response to direct elicitation from the fieldworkers.

2.2 The target structure: variable 1sg subject expression in Spanish

From this corpus, we extracted all tokens of finite Spanish verbs with (expressed and unexpressed) 1sg

https://doi.org/10.1017/S1366728915000127 Published online by Cambridge University Press

subjects produced by the participants (N = 3,766).[2] After circumscribing the envelope of variation and exclusion of tokens[3], there remain for analysis a total of 3,288 instances of variable 1sg subject expression, in which the overall rate of *yo* expression is 28%.

It is important to note that, despite New Mexican Spanish having been in intense and prolonged contact with English, we have found that the patterning of 1sg subject expression in NMSEB parallels that reported for non-contact varieties of Spanish (on New Mexican Spanish, see Torres Cacoullos and Travis (2011, pp. 250–251) and Torres Cacoullos and Travis (2015); on other varieties, see Bentivoglio (1983, p. 36, 60), Cameron (1994, p. 32, 38, 40), Otheguy and Zentella (2012, pp. 163–165), Silva-Corvalán (2001, pp. 154–169) and Travis and Torres Cacoullos (2012, p. 726)). Across dialects and genres of Spanish, the likelihood that speakers will use *yo* vs. an unexpressed 1sg subject is increased with cognition verbs (e.g., *yo creo* 'I think'). Also favorable to *yo* are contexts of lesser subject continuity, as in (2), where between the target verb (line 4) and the previous coreferential 1sg subject (line 1) there intervene two clauses (one of which, furthermore, has a subject with a distinct human referent, 'the doctor', in line 2).

(2)

1.	Trinidad	*y cuando **Ø** fui al doctor y,*
2.		*.. la doctora me dijo,*
3.		*tienes cáncer,*
4.		*ya **yo** me desesperaba.*
		[21 Demerits, 0:56:27–0:56:32]
1.	Trinidad	'and when (**I**) went to the doctor and,
2.		.. the doctor told me,
3.		you have cancer,
4.		**I** was already feeling desperate.'

The fact that these same constraints are operative in NMSEB provides evidence that, despite lexical borrowing from English and regular multi-word code-switching (Aaron, forthcoming; Wilson & Dumont, forthcoming), these speakers have made no accommodation to English-particular constraints in patterns of subject expression (Torres Cacoullos & Travis, 2015), and nor have they for subject position with respect to the verb (Benevento & Dietrich, forthcoming). Thus, there is no evidence for contact-induced morphosyntactic change in New Mexican Spanish. There is evidence, however, for cross-language priming.

2.3 Priming in subject expression

Priming in subject expression has been observed in several languages, including Spanish. In contexts where speakers have a choice between an expressed or an unexpressed subject, there is a tendency to repeat the same form used previously, or, as Cameron (1994, p. 40) put it, "pronouns lead to pronouns, and null subjects lead to null subjects". Such priming has been found in Australian Sign Language (Auslan; McKee, Schembri, McKee & Johnston, 2011, pp. 387–389), Tamambo (Vanuatu) (Meyerhoff, 2009, pp. 308–309), and, in experimental work, in Italian children (Serratrice, 2009).

The strongest priming effect for Spanish 1sg subject expression is Coreferential Subject Priming, that is, the effect of the realization of the previous coreferential 1sg subject produced by the same speaker (Cameron, 1994, pp. 39–40; Travis, 2005, pp. 343–345; Travis, 2007; Travis & Torres Cacoullos, 2012, pp. 729–733). Example (3) illustrates Spanish within-language priming: in lines 1–2, an unexpressed subject is followed by another unexpressed subject, and in lines 9–10, the pronoun *yo* is followed by another *yo*. Example (4) illustrates cross-language priming, where a previous *I* primes a subsequent Spanish *yo* (Torres Cacoullos & Travis, 2010, 2011, 2015).

(3) Coreferential 1sg subject priming: *yo* 'I' to *yo* 'I', unexpressed to unexpressed (Ø to Ø)

1.	Sandra	*nomás **Ø** lo vi,*
2.		*y **Ø** me solté llorando.*
3.		*. . . qué te pasa,*
4.		*dijo,*
5.		*que te pegó la maestra?*
6.		*no,*
7.		*Ø le dije,*
8.		*. . . (H) pero me dijo la maestra,*
9.		*que **yo** no podía venir a la escuela,*
10.		*porque **yo** no sé hablar inglés.*
		[03 Dos Comadres, 57:47:3–57:52:7]
1.	Sandra	'As soon as (**I**) saw him,
2.		and (**I**) burst into tears.
3.		. . . what happened,
4.		he said,
5.		did the teacher hit you?
6.		no,
7.		(I) said,

[2] We circumscribe the analyses here to first person singular subject pronoun (*yo*) expression, because it may be that conditioning factors for subject pronoun expression differ by grammatical person; first person differs from third person in that it refers to 'given' information (by virtue of the speaker as a discourse participant), and in that it can be expressed by pronouns or unexpressed subjects, while third person can also be expressed by lexical NPs.

[3] Excluded from the envelope of variation are tokens of post-verbal *yo*, which follows distinct constraints (cf. Benevento & Dietrich, forthcoming; Silva-Corvalán, 1982), a range of non-variable contexts (including fixed expressions and wh-interrogatives), and several contexts that could not be reliably analyzed (e.g., where the verb was unclear, where the candidate subject pronoun occurred in a different prosodic unit from the verb, unexpressed tokens involving repair).

https://doi.org/10.1017/S1366728915000127 Published online by Cambridge University Press

8. . . . (H) but the teacher said to me,
9. that **I** couldn't come to school,
10. because **I** don't know how to speak English.'

(4) Cross-language coreferential 1sg subject priming: English *I* to Spanish *yo*

1. Aurora *my mom thought I knew how t- --*
2. *.. had to know how to spell my name,*
3. *how to write it befo=re **I** went to school.*
4. *(H) pero cuando **yo**=fui,*
5. *no había kindergarten.*

[15 Las Cosas Viejas, 0:29:29–0:29:36]

1. Aurora *'my mom thought I knew how t- --*
2. *.. had to know how to spell my name,*
3. *how to write it befo=re **I** went to school.*
4. (H) but when **I** went,
5. there was no kindergarten.'

Target 1sg subjects were coded for the realization (form) of the previous coreferential 1sg subject (that is, a 1sg subject produced by the same speaker). In this bilingual corpus, a previous coreferential 1sg subject (what we refer to as the prime) was realized as an unexpressed subject with a Spanish verb (illustrated in line 1 in (3)), the Spanish pronoun *yo* (line 9 in (3)) or the English pronoun *I* (line 3 in (4)). (The fourth theoretical possibility, that of an English unexpressed subject, occurred only five times in the dataset used for the present study.)

Tokens occurring in contexts which could not be reliably analyzed for priming were set aside. Thus, excluded were tokens for which the previous coreferential mention occurred post-verbally, because it is unknown how position (pre- vs. post-verbal) interacts with the priming of realization (expressed vs. unexpressed); instances where the target or previous coreferential mention occurred in quoted speech (unless both target and prime occurred within the quoted speech, as in example (8)), pending investigation of how quotation may affect priming (N = 314); instances where a previous unexpressed token occurred with unclear speech or truncation (since it is unknown what the speaker intended); and instances where the previous subject pronoun is produced in a distinct prosodic unit (Intonation Unit) from the verb, since its status as a subject may be questioned (see Croft, 1995 and Croft, 2007 on Intonation Units and grammatical structure) (N = 96). We also excluded all tokens for which there was no coreferential 1sg subject in the preceding five clauses (N = 808), because of the dissipation of priming effects with increasing distance between target and prime (see Section 4).

Having delimited the relevant priming contexts, we now turn to consider further this observation of cross-language priming outside the lab (a first, as far as we know), in particular, how it can illuminate the question of the relative strength of within- versus cross-language priming.

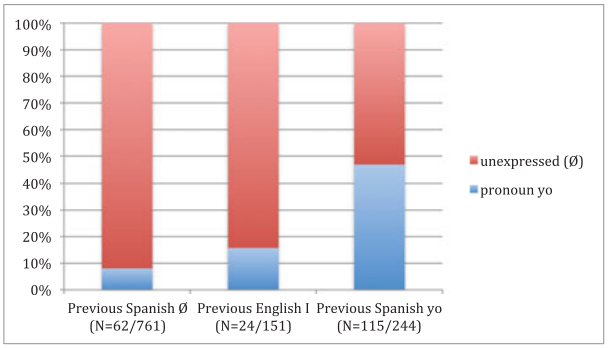


Figure 1. (colour online) Rate (%) of Spanish 1sg subject pronoun *yo* by realization of the previous mention in the context of maximal subject continuity (0 intervening clauses between coreferential mentions as subject) (N = 1,156).

3. Are within- and cross-language priming in bilingual speech equally strong?

Our first test of the relative strength of within-language (Spanish-to-Spanish) and cross-language (English-to-Spanish) priming focuses on the context of maximal subject continuity: namely, when the clause immediately preceding the target clause has a coreferential 1sg subject, as seen above in lines 1–2, and 9–10, in (3) and lines 3–4 in (4). Figure 1 presents the rate of *yo* expression in this maximally continuous context according to the realization of the previous coreferential subject. When the previous coreferential subject is unexpressed the subsequent rate of *yo* is lowest, at 8% (62/761); when the previous subject is English *I* the rate of *yo* is 16% (24/151); and when the previous subject is *yo* the rate of *yo* is as high as 47% (115/244).

The data were analysed using Generalized Linear Mixed Models (GLMM), which were calculated with the *lme4* package for Linear Mixed Effects (Bates & Mächler, 2010; R Core Team, 2012). A binomial distribution was specified because we were modelling the realization of 1sg Spanish subjects (i.e., Ø versus *yo*). Out of the four possible primes (unexpressed vs. pronominal Spanish and unexpressed vs. pronominal English), as noted above unexpressed English subjects are vanishingly rare, and therefore we treated the prime form as one variable with three levels (prime form: Ø, *I*, and *yo*, Reference level = Ø, as in Torres Cacoullos & Travis (2011, p. 258, Table 7)). This enabled us to determine whether speakers' choice of *yo* is significantly and differentially affected by these different prime types. A model containing the fixed variable of prime form was a significantly better fit to the data than a null model containing only the random

https://doi.org/10.1017/S1366728915000127 Published online by Cambridge University Press

Table 1. *Pairwise comparisons across prime forms in context of maximal subject continuity (N = 1,156).*

	β	*SE*(β)	*z*	*p*
Ø vs. *yo*	2.06	.29	7.17	<.001
Ø vs *I*	.73	.34	2.14	= .033
I vs. *yo*	1.33	.34	3.93	<.001

Overall *yo* expression = 17%, log likelihood = −432.7.

effect of speaker ($\chi^2 = 48.01$, df = 2, $p < .001$).[4] Table 1 reports the pairwise comparisons across the three primes, where β represents the log odds of *yo* expression in context A (e.g., a Ø prime) compared to context B (e.g., a *yo* prime) (cf., Jaeger, 2008). For instance, the log odds of *yo* expression following the previous expression of *yo* versus a Ø form (top row) is positive and significant (β = 2.06), which means that *yo* is significantly *more likely* than a Ø form following a *yo* prime than a Ø prime, thus confirming within-language priming. Cross-language priming is confirmed in the second row, which shows that *yo* is favored with a previous realization of *I* vs. an unexpressed subject. However, *yo*-to-*yo* priming is significantly higher than *I*-to-*yo* priming (bottom row).

This weaker effect for an English prime could be interpreted as a result of a genuinely weaker priming effect across languages, or alternatively, as a weaker effect for 'translation equivalents' (*I* to *yo*) than for identity priming (*yo* to *yo*). The latter would be consistent with a lexical boost account, whereby priming is enhanced due to identity of form. We return to this after considering the question of the persistence of priming, below.

4. Do within- and cross-language priming persist (or decay) to the same extent?

It is well known that priming can persist across intervening utterances in discourse. In sociolinguistic interviews, Weiner and Labov (1983, p. 52), for example, found that the occurrence of an agentless passive anywhere in the preceding five clauses favors a subsequent such passive (over a non-specific subject). Persistence of priming in spontaneous discourse has also been found for subject expression in Spanish (Travis, 2007, pp. 119–129; Travis & Torres Cacoullos, 2012, pp. 731–732). This result is consistent with the experimental work showing that priming can be retained over intervening linguistic material in both adult and child populations (Bock & Griffin, 2000; Savage, Lieven, Theakston & Tomasello, 2006; Serratrice, 2009). At the same time, both corpus and experimental studies have found that priming decays with distance (e.g., Branigan, Pickering & Cleland, 1999; Gries, 2005; Szmrecsanyi, 2005).

Here we explore such 'decay' and ask whether the coreferential priming from the immediately preceding clause that we observed in Section 3 is maintained over a degree of distance. Measures of distance between target and prime that have been used in corpus studies are elapsed time and number of words or parsing units (modeled logarithmically) (e.g., Gries, 2005, p. 120; Szmrecsanyi, 2005, p. 371). Following earlier variationist work (e.g., Scherre, 2001; Scherre & Naro, 1991; Scherre & Naro, 1992; Weiner & Labov, 1983), we base our distance measure on well-defined structural units – here, the clause – reasoning that other subjects and verbs affect the strength of a previous coreferential 1sg subject prime more than arbitrary linguistic material or mere time. An example can be seen in (2) above, where there are two clauses intervening between prime (in line 1) and target (in line 4).

Table 2 depicts the rate of *yo* in relation to priming by the previous coreferential 1sg subject according to distance in clauses between target and prime. We see here that *yo*-to-*yo* priming persists beyond the context of maximal continuity, or zero intervening clauses: across all rows, the rate of *yo* is higher with a previous *yo* than with a previous unexpressed 1sg subject. Table 2 also indicates, however, that this priming weakens with distance: at zero intervening clauses, the rate of *yo* is almost six times as high in the presence of a previous *yo* than with a previous unexpressed, 47% vs. 8% ($p < .0001$) (as we saw in Figure 1); at one intervening clause, it is almost five times as high, 49% vs. 10% ($p < .0001$); but at greater distances, it is just under twice as high (at two clauses, 42% vs. 24% ($p = .03$), at 3–4, 50% vs. 27% ($p = 0.007$) (p values from Fisher's exact test).

Now considering the cross-language, English-to-Spanish, results in Table 2, we find a similar tendency, but it is shorter lived than the corresponding within-language effect. At zero intervening clauses, the rate of *yo* expression is twice as high in the context of a previous *I* vs. a previous unexpressed subject (16% vs. 8%, $p = 0.006$). At a distance of one intervening clause, it remains almost twice as high (19% vs. 10%, marginally significant at $p = 0.06$), but at greater distances the tendency is clearly much weaker than the corresponding within-Spanish priming, since the rate of *yo* is not significantly higher with a previous mention as *I* than with a previous unexpressed subject (at two intervening clauses 26% vs. 24% ($p = 0.83$), at 3–4 intervening clauses 37% vs. 27% ($p = 0.27$). Thus, not only is cross-language coreferential 1sg subject priming weaker to start with (i.e., under maximal subject continuity), it is also shorter lived than the corresponding within-language effect.

[4] All models contained random effects for speakers, and random slopes for fixed effects if they significantly contributed to model fit (see Baayen, 2008).

https://doi.org/10.1017/S1366728915000127 Published online by Cambridge University Press

Table 2. *Rate (%) of Spanish 1sg subject pronoun* yo *by Coreferential Subject Priming (realization of the previous mention) and Distance in clauses from the previous mention (N = 2,723).*

	Previous Span Ø		Previous English *I*		Previous Span *yo*		Overall	
	%	N	%	N	%	N	%	N
0	8%	761	16%	151	47%	244	17%	1156
1	10%	206	19%	74	49%	76	20%	356
2	24%	106	26%	38	42%	43	28%	187
3–4	27%	114	37%	52	50%	50	35%	216
5+							45%	808
	12%	1187	21%	315	47%	413	28%	2723

Table 3. *Rate (%) of Spanish 1sg subject pronoun* yo *by Coreferential Subject Priming (realization of the previous mention), for maximally continuous subjects (0 intervening clauses) and discontinuous subjects (1–4 intervening clauses) (N = 1,915).*

	Previous Span Ø		Previous English *I*		Previous Span *yo*		Overall	
	%	N	%	N	%	N	%	N
0	8%	761	16%	151	47%	244	17%	1156
1–4	18%	426	26%	164	47%	169	26%	759
	12%	1187	21%	315	47%	413	21%	1915

4.1 Assessing priming within other variable constraints: interaction with subject continuity

Natural speech data allows – and demands – that priming effects be situated and interpreted within the wider set of constraints operative on variation. For variable Spanish subject expression, we must confront the interplay between priming and subject continuity. According to this well-established constraint (Section 2.2), the rate of *yo* is lower in continuous than in discontinuous contexts (Table 2, rightmost column), that is, when continuity of reference is broken by the presence of non-coreferential subjects, as in line 4 in (2) above. Table 3 and Figure 2 depict priming in the context of maximal subject continuity (zero intervening clauses) as compared with discontinuous contexts, here combining distances of 1–4 intervening clauses.

We find that the decay of priming is not a straightforward result of increasing distance between prime and target. Rather, there is interaction between coreferential 1sg subject priming and the effect of subject continuity on subject expression, such that priming is stronger in contexts of maximal subject continuity than it is in contexts of discontinuous subjects, as seen by the steeper slope of the line for 'maximally continuous subjects' in Figure 2. This reflects the fact that priming works antagonistically to subject continuity when the prime is a previous *yo* or *I* (which favors *yo*) and the context is one of maximal subject continuity (which favors the unexpressed variant), but synergistically when the prime is a previous unexpressed 1sg and the context is one of maximal subject continuity (both environments favoring unexpressed) (Travis, 2007, p. 124).

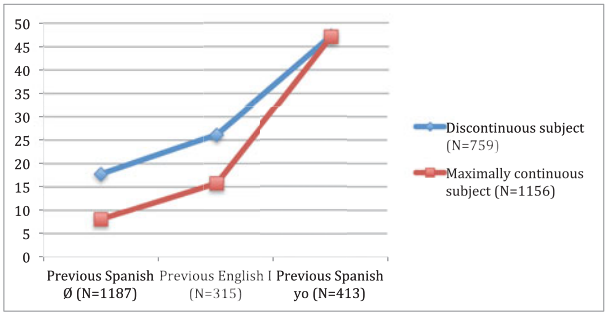


Figure 2. (colour online) Rate (%) of Spanish 1sg subject pronoun *yo* by realization of the previous mention for maximally continuous subjects (0 intervening clauses) and discontinuous subjects (1–4 intervening clauses) (N = 1,915).

For its part, the subject continuity constraint applies in the context of a previous unexpressed (seen on the left in Table 3 and in Figure 2): the rate of *yo* more than doubles, from 8% (62/761) in the context of maximum subject

https://doi.org/10.1017/S1366728915000127 Published online by Cambridge University Press

Table 4. *Final model parameter estimates for analysis predicting priming of* yo *by Prime Forms (Ø, I, yo) and Subject Continuity (0 versus 1–4 intervening clauses) (N = 1,915).*

	β	*SE*(β)	*z*	*p*
Intercept	−2.37	.18	−13.33	<.001
Prime form				
Ø vs I	.84	.27	3.07	.002
Ø vs. *yo*	2.12	.19	10.99	<.001
Subject continuity	.95	.19	5.00	<.001
Interaction				
(Ø vs *I*)∗Subj continuity	−.34	.35	−.96	.34
(Ø vs. *yo*)∗Subj continuity	−1.00	.28	−3.51	<.001

Overall *yo* expression = 21%, log likelihood = −832.96.

continuity to 18% (76/426) in discontinuous contexts. It is weakened, if not nullified, however, in the context of a previous *yo* (on the right in the table and figure): the rates are identical in both maximally continuous and in discontinuous contexts, at 47% (115/244 and 80/169, respectively). In contrast, in the context of a previous *I* (in the middle in the table and figure), the subject continuity effect is maintained: the rate of *yo* is over 1.5 times as high, from 16% (24/151) in the context of maximum subject continuity to 26% (43/164) in discontinuous contexts. Table 3 and Figure 2 thus provide further indication that cross-language *I*-to-*yo* priming is weaker than within-language *yo*-to-*yo* priming, as the former does not supersede the effect of subject continuity.

As above, these data were analyzed using GLMMs. The fixed variables of prime form (3 levels: Ø, *I*, and *yo*, Reference level = Ø) used in the previous analysis were entered into a factorial model with subject continuity (2 levels: 0 versus 1–4 intervening clauses, Reference level = 0). The final model parameters are shown in Table 4.

The two main effects (priming and subject continuity) were significant, as was their interaction. With respect to the main effect of prime, we observed the same differences across forms as were observed in the previous analysis restricted to cases of maximal subject continuity. This included a significant difference between *yo* and *I* primes ($\beta = 1.28$, $SE(\beta) = .27$, $z = 4.7$, $p < .001$), which shows that, when averaging over a distance of up to four intervening clauses, the within-language priming effect is stronger than the cross-language effect. The main effect for subject continuity shows that there is a greater likelihood that *yo* will be expressed when there are one or more intervening clauses between prime and target than when there are none. Finally, the significant interaction for prime form X subject continuity was driven by the fact that, whereas the (on average lower, or weaker) cross-language priming effect did not significantly change across discontinuous contexts, the (on average higher, or stronger) within-language priming effect did. Specifically, the weak cross-language priming effect appears to be maintained whereas the stronger within-language effect is weakened (Table 2). Since the within-language priming effect is very strong in continuous contexts, this interaction may be at least in part due to the fact that there is more room to significantly decay. The apparent persistence of the weaker cross-language priming effect in discontinuous contexts may be due to the fact that we have collapsed across distances of 1 to 4 clauses.

In summary, while coreferential 1sg subject priming persists with intervening material it decays with increasing distance in clauses from the prime. Furthermore, the priming effect is weaker across languages, and the decay more rapid. In these natural speech data, priming is shown to interact with features of the discourse context, such that it is stronger in contexts of maximal subject continuity than it is with discontinuous subjects.

Once again, these results could be taken as evidence of generally weaker priming across languages. Alternatively, if *yo*-to-*yo* priming is viewed as involving a lexical boost, these results could be seen to evidence a weaker effect for 'translation equivalents' than for identity priming, as has been reported in prior research (Bernolet et al., 2013; Schoonbaert et al., 2007). We confront this question in the following section.

5. Does lexical boost explain stronger within-language priming? Non-coreferential pronoun-to-pronoun priming

To tackle the question of whether the weaker effect for cross-language priming is attributable to a lexical boost effect of the identity of form, or whether it should be interpreted as due to genuinely weaker cross-language than within-language associations, we next ask

https://doi.org/10.1017/S1366728915000127 Published online by Cambridge University Press

Table 5. *Rate (%) of Spanish 1sg subject pronoun* yo *by realization of immediately preceding clause non-coreferential subject (produced by speaker) (N = 483).*

	Preceding Span Ø		Preceding English pronoun		Preceding Span pronoun		Overall	
	%	N	%	N	%	N	%	N
Rate of *yo*	33%	341	39%	76	52%	66	36	483

whether there is priming in the absence of form-to-form correspondence. Specifically, we investigate whether non-coreferential pronoun-to-*yo* priming occurs, that is, whether 1sg subject expression can be primed by other (non-1sg) Spanish and English pronominal subjects. If the weaker cross-language than within-language subject pronoun priming seen thus far is due to identity priming, then the difference between English-to-Spanish and Spanish-to-Spanish pronoun priming should disappear when the pronoun is other than a first person, a scenario in which identity priming is ruled out.

We examine non-coreferential priming by considering the realization of the immediately preceding clause subject produced by the same speaker, i.e., excluding potential primes produced by an interlocutor.[5] Furthermore, as well as the exclusions noted above (Section 2.3), we consider here only specific human referents, excluding tokens where the immediately preceding clause had a non-human or non-specific subject (which do not occur with a personal pronoun variant) and where the subject was a lexical NP, as well as, of course, all 1sg coreferential mentions. Table 5 provides the results.

There is a patently weaker effect for non-coreferential than coreferential subject pronoun priming, seen by comparing Table 5 with the first row of Table 3. This is due to the subject continuity effect, which results in an elevated rate of *yo* overall when the subject of the preceding clause is non-coreferential (be it expressed or unexpressed). The data were again analyzed using GLMMs. The main effect for prime type was not significant when compared to a null model containing only the random effect of speaker, $\chi^2 = 4.44$, df = 2, $p = .11$, confirming that overall, non-coreferential subject priming was weaker than coreferential 1sg subject priming (seen above in Tables 1 and 4). When we compare the individual prime forms, we find that there was a significantly higher rate of *yo* following a Spanish pronoun, as in (5), than when following a Spanish unexpressed subject, as in (6) [52% vs. 33%, $\beta = .62$, SE(β) = .3, $z = 2.1$, $p = .036$], evidence of a within-language non-coreferential subject priming effect. However, the rate of *yo* following an English pronoun, as in (7), was not significantly higher than when following an Spanish unexpressed subject [39% vs. 33%, $\beta = .25$, SE(β) = .29, $z = .85$, $p = .40$], suggesting no cross-language non-coreferential subject priming effect (though the within-language priming effect (rate of *yo* following a previous Spanish pronoun) did not significantly differ from the cross-language effect (rate of *yo* following a previous English pronoun) [$\beta = .36$, SE(β) = .38, $z = .99$, $p = .32$]).

(5) Non-coreferential subject priming: Spanish pronoun to *yo*

Ivette *estábamos en el blackboard,*
y luego estaban todas.
se iban para atrás del closet.
(H) .. y luego yo y ella pintábamos.
.. (H) pero pues ***ella*** *estaba aquí,*
yo *estaba aquí,*
y decían,
.. <VOX which one is better VOX>?
[04 Piedras y gallinas, 0:20:08–0:20:13]

Ivette ‘we were at the blackboard,
and then everyone was there.
they went behind the closet.
(H) .. and then she and I painted.
.. (H) but well **she** was here,
and **I** was here,
and they would say,
.. <VOX *which one is better* VOX>?’

(6) Non-coreferential subject priming: Spanish unexpressed to unexpressed

Rocío *el único lugar que yo iba,*
era para el cine.
. . . (0.7) mexicano con mi tía.
. . . (H) [porque ***Ø*** *me] llevaba a ver a= --*
Adriana *[mhm].*
Rocío *.. a Pedro In[2fante2]=,*
Adriana *[2fante2].*
Rocío *a --’*
(H) Jorge Negrete,
y=,

[5] Clauses immediately preceding the target 1sg verb were produced by an interlocutor in 14% (445/3,288) of the instances under study.

https://doi.org/10.1017/S1366728915000127 Published online by Cambridge University Press

. . . % % . . . María Félix,
(H) .. y um,
*(H) y **Ø** creo que por eso hoy,*
todavía Ø veo .. telenovelas en español.
[05 Las tortillas, 0:10:40–0:11:01]
Rocío ‘the only place that I went,
was to the cinema.
. . . (0.7) Mexican ((cinema)) with my aunt.
. . . (H) [because] (**she**) would take me
to see --’
Adriana ‘[mhm].’
Rocío ‘.. Pedro In[2fante2]=,
Adriana ‘[2fante2].’
Rocío to --
‘(H) Jorge Negrete,
and,
. . . % % María Félix,
(H) .. and um,
(H) and (**I**) think that that’s why today,
(I) still watch soapies in Spanish.’

(7) Non-coreferential subject priming: English pronoun to *yo*
Pedro *he didn’t know.*
***he** wasn’t there*.
Ricardo *. . . mhm.*
Pedro *pues,*
***yo** no conocía a --*
su esposa la Lynn,
I didn’t know who she <@ was @>.
[07 Basketball teams, 0:13:03–0:13:08]
Pedro *‘he didn’t know.*
***he** wasn’t there*.’
Ricardo ‘mhm.’
Pedro ‘well,
I didn’t know --
his wife Lynn,
I didn’t know who she <@ was @>.’

Once again, then, we see a tendency for a weaker cross-language than within-language priming effect. This cannot be dismissed as due to the workings of a within-language lexical boost as compared with translation counterparts, as there are no cases of identity of form here. We note, though, that more than one half of these tokens (270/483) occur with a recent coreferential 1sg subject mention, specifically within the preceding five clauses, as in (8). Thus it could be argued that this is in fact coreferential 1sg subject priming persisting over intervening clauses rather than a within-Spanish non-coreferential subject pronoun priming effect. If this is the case, then a within-language lexical boost remains as a potential explanation for the apparently weaker cross-language effect.

(8) Non-coreferential subject priming: Spanish pronoun to *yo*, with *yo* in the preceding five clauses
Aurora *he would just say,*
oh no más escríbalo,
***yo** lo firmo*.
Ricardo *.. yeah.*
Aurora *.. **tú** lo escribes,*
***yo** lo firmo*.
[15 Las Cosas Viejas, 0:27:42–0:27:45]
Aurora ‘*he would just say,*
oh just write it.
I’ll sign it.’
Ricardo ‘.. yeah.’
Aurora ‘.. **you** write it,
I’ll sign it.’

To determine whether lexical boost indeed accounts for weaker cross-language priming, we now limit our view to only those targets lacking a coreferential 1sg prime within the preceding five clauses, presented in Table 6. Here, precisely the same tendency for rates of expression obtains: 42% in the context of a preceding Spanish unexpressed, 44% in the context of a preceding English pronoun, and 62% in the context of a preceding Spanish pronoun. That is, again, the difference in *yo* rate is slightly greater between a preceding Spanish unexpressed and Spanish pronoun (marginally significant at $p = 0.06$) than that between a preceding Spanish unexpressed and an English pronoun ($p = 0.86$), suggesting that the patterning observed for non-coreferential priming (Table 5) is likely not (merely) attributable to the persistence of coreferential subject priming, and hence cannot be dismissed as (solely) a lexical boost effect.

What the existence of non-coreferential subject priming suggests is an association between the more specific [(*yo*) + Verb_{1sg}] target construction (with a variably expressed 1sg subject pronoun and a 1sg verb) and a more schematic [(pronoun) + Verb] construction (a variably expressed subject pronoun of any person and an agreeing verb), and could be taken to support an understanding of subject pronoun priming as operating on abstract syntactic structures, a point to which we return in the following section. In summary, the greater strength of non-coreferential subject priming within Spanish than from English to Spanish suggests that the weaker cross-language than within-language priming effect observed for Coreferential Subject Priming may not be dismissable as simply due to the formal identity that is obtained in *yo*-to-*yo* priming and absent in *I*-to-*yo* priming. Rather, cross-language subject pronoun priming is indeed weaker than within-language priming, both for coreferential 1sg subject priming (Section 3) (also evidenced in the fact that is decays more rapidly across than within languages, Section 4), and for non-coreferential subject priming.

https://doi.org/10.1017/S1366728915000127 Published online by Cambridge University Press

Table 6. *Rate (%) of Spanish 1sg subject pronoun* yo *by realization of immediately preceding clause subject (non-coreferential subject produced by same speaker, in the absence of a coreferential 1sg subject in the previous five clauses) (N = 213).*

	Preceding Span Ø		Preceding English pronoun		Preceding Span pronoun		Overall	
	%	N	%	N	%	N	%	N
Rate of *yo*	42%	143	44%	41	62%	29	45%	213

Table 7. *Rate (%) of Spanish 1sg subject pronoun* yo *by Coreferential Subject Priming (realization of the previous mention, 0–4 intervening clauses) and class of verb (N = 1,904).*

	Previous Span Ø		Previous English *I*		Previous Span *yo*	
	%	N	%	N	%	N
Cognition	33%	43/130	40%	25/63	62%	61/98
All other verbs	9%	94/1051	16%	41/251	42%	131/311
	12%	1181	21%	314	47%	409

6. Priming in interaction with particular constructions

Having answered our research questions on the relative strength of within-language versus cross-language priming and having been compelled by the data to dismiss a lexical boost account for the weaker cross-language subject priming effect, we now turn to the question of the grammatical units on which priming operates, namely the much-discussed issue of whether they are lexically-based or are abstract. Although priming effects have been found to be larger when prime and target sentences contain lexical overlap, priming persists in the absence of overlap. For example, for the dative alternation, *the girl sent her boyfriend an ultimatum* (as opposed to *the girl sent an ultimatum to her boyfriend*) primes *Stephen baked Anthony a cake* (over *Stephan baked a cake for Anthony*). This has been interpreted as evidence that priming acts on abstract linguistic representations (here, the double object dative vs. the prepositional object dative) independently of lexical form (Bock, 1986). On the other hand, it has also been shown that the frequency with which a particular verb occurs in prepositional or double object dative contexts moderates the priming effect (Gries, 2005) (see also Jaeger & Snider, 2013), a finding which has been experimentally supported (Bernolet & Hartsuiker, 2010), and been taken to indicate that priming is not independent from lexical form.

We noted above the interaction between priming and the subject continuity constraint on subject expression that can be observed in spontaneous discourse (Section 4.1).

Table 8. *Final model parameter estimates for analysis predicting priming of* yo *by Prime Forms (Ø, I, yo) and Verb Type (Cognition versus Other) (N = 1,904).*

	β	*SE*(β)	*z*	*p*
Intercept	−2.35	.16	−15.09	<.001
Prime form				
Ø vs I	.71	.23	3.13	.002
Ø vs. *yo*	1.68	.20	8.46	<.001
Verb type	1.93	.31	6.20	<.001
Interaction				
(Ø vs *I*)∗Verb type	−.38	.43	−.88	.38
(Ø vs. *yo*)∗Verb type	−.81	.36	−2.26	.023

Overall *yo* expression = 21%, log likelihood = −792.8.

Priming also interacts with the effect for verb class, whereby cognition verbs such as *creer* 'think' favor *yo* more than other verbs. Table 7 depicts the rate of *yo* by the class of verb and previous realization.

The data were again analyzed using GLMMs, and here, prime type (3 levels: Ø, *I*, and *yo*, Reference level = Ø) was entered into a factorial model with verb type (2 levels: Cognition versus 'Other', reference level = 'Other'). All model terms significantly contributed to model fit. The estimates for the final model are shown in Table 8.

The main effects of prime form and verb type were significant. The same pattern of priming effects as in the previous analyses was obtained, and once again

https://doi.org/10.1017/S1366728915000127 Published online by Cambridge University Press

further analyses revealed that the difference between *yo* and *I* primes was significant (β = .97, *SE*(β) = .23, *z* = 4.22, *p* <.001), supporting stronger within-language priming overall. The main effect of verb type showed that the rate of *yo* was highest when the target utterance contained a cognition verb. The significant prime form X verb type was driven by the fact that *yo*-to-*yo* priming was attenuated when target utterances contained a cognition verb, whereas this was not the case for the (generally already weaker) *I*-to-*yo* priming effect. What this demonstrates, then, is that the priming effect may be stronger or weaker depending on the construction in which the target occurs.

The lower susceptibility of 1sg cognition verbs to *yo*-to-*yo* priming follows from particular high frequency collocations. The vast majority (84%, 540/640) of the cognition verbs in these data are made up of four particular expressions (*yo creo* 'I think', with a *yo* rate of 90% (137/153), *yo sé* 'I know' with a *yo* rate of 80% (37/46), *yo no sé* 'I (don't) know', with a *yo* rate of 45% (93/206), and (*no*) *me acuerdo* 'I (don't) remember, with a *yo* rate of 13% (18/135)). These are good candidates for lexically-specific constructions by measures of token frequency and linguistic conditioning, and may be more likely to be accessed holistically rather than as combinations of subject and verb.[6] The lesser priming of subject expression with 1sg cognition verbs is evidence of lesser analyzability of the *yo* + cognition verb sequence as compared with other *yo*-plus-verb combinations, attributable to chunking (Bybee, 2010, p. 34, and references therein). Reduced analyzability leads to a weaker association between the 1sg subject of cognition verbs and other instances of the 1sg subject.

This lower susceptibility of more specified constructions to priming suggests that the mental representations of subject pronoun constructions may be of varying degrees of lexical specificity or schematicity (for example, [*yo* + *creo*] 'I think', [(*yo*) + Cognition Verb_{1sg}], [(*yo*) + Verb_{1sg}], [(pronoun) + Verb]). And, crucially, subject pronoun constructions may show varying strengths of association. Szmrecsanyi (2005, p. 139) has shown priming by a lexically-related construction that is not part of the alternation (e.g., priming of English future *be going to* vs. *will* by the motion verb *go*, what he calls β-persistence). Building on this and our current finding, we propose that priming speaks to the strength of associations between constructions and components of constructions (Torres Cacoullos, 2015). In this case, priming serves as a gauge of association between the *yo* in the cognition verb construction [*yo* + Cognition Verb_{1sg}], and the *yo* of the more general [*yo* + Verb_{1sg}].

Although there has been much discussion about whether priming functions on abstract syntactic representations or is more lexically-based, in our view drawing a sharp distinction between the grammar and the lexicon fails to capture the multiplicity of ways in which lexical and syntactic representations overlap. Rather than compartmentalizing lexicon and syntax, a usage-based approach allows us to view grammatical constructions on a continuum from the more lexically particular to the more schematically general. The question then becomes one of the strength of the associations between constructions. That is, the pertinent question is, to what degree is the more specific [(*yo*) + Verb_{1sg}] construction associated with a more schematic [(pronoun) + Verb] construction within Spanish and across Spanish and English?

As we have seen, the reduced strength of priming with cognition verbs points to a degree of independence of [(*yo*) + Cognition Verb_{1sg}] from the more general [(*yo*) + Verb_{1sg}] construction, while the differential robustness of coreferential and non-coreferential subject priming can be interpreted in terms of a closer association between instances of [(*yo*) + Verb_{1sg}] than between [(*yo*) + Verb_{1sg}] and [(Pronoun) + Verb]. In the bilingual case, in the same vein, the associations between [(*yo*) + $\text{Verb}_{1sg}]_{\text{Spanish}}$ and [*I* + $\text{Verb}_{1sg}]_{\text{English}}$ are less close than the corresponding within-language associations. This makes sense in a usage-based view of grammar, where grammatical constructions are understood to include information about linguistic as well as extralinguistic contexts of use (e.g., Bybee, 2010; Croft & Cruse, 2004; Goldberg, 2006).

7. Discussion

In the current study we have investigated within- and cross-language priming of variable 1sg subject expression in a bilingual corpus of spontaneous speech. Although in previous work we have shown cross-language priming in bilingual speech, in this paper we tested the limits of the effect, and in doing so provide the first detailed consideration of cross-language priming outside the lab. Guided by findings in the variationist and experimental literature, we explored the relative strength of within- and cross-language priming to ask (i) whether within-language and cross-language priming are equally strong, (ii) whether they persist to a similar extent and (iii) whether the notion of a lexical boost provides an account for their differential strength.

Our first finding was that, following previous research (Torres Cacoullos & Travis, 2011, 2015), both within-language *yo*-to-*yo* and cross-language *I*-to-*yo* priming were significant. Consistent with experimental

[6] Indices of the status of *yo creo* as a particular construction as outlined by Travis and Torres Cacoullos (2012, p. 741) based on a corpus of conversational Colombian Spanish, are: (1) the high token frequency of the *yo creo* string; (2) the high proportion it comprises of all occurrences both of the lexical type *creer* and of the pronoun *yo*; and (3) in independent multivariate analysis of (*yo*) *creo*, the absence of a subject continuity effect.

https://doi.org/10.1017/S1366728915000127 Published online by Cambridge University Press

studies (Bernolet et al., 2013; Schoonbaert et al., 2007) the cross-language effect was smaller in magnitude than the within-language effect (Table 1, Figure 1).

Second, in examining the persistence of priming in spontaneous discourse, we have had to take into account another contextual factor that conditions variable subject expression, namely subject continuity. Subject continuity moderates the priming effect, as expressed subjects are favored in discontinuous, and disfavored in continuous, contexts (Table 3, Figure 2). Nevertheless, we find that, while attenuated, coreferential 1sg subject priming persists across intervening material (up to 3–4 intervening clauses, within Spanish, Table 2). Furthermore, in comparison with within-language priming, the cross-language priming effect in discontinuous contexts is both weaker and decays more rapidly (after one intervening clause, Table 2, Table 4).

The present results are interesting in the context of lexical boost effects in experimental studies, which typically show catastrophic decay (Branigan et al., 1999; Malhotra, Pickering, Branigan & Bednar, 2008). One possible explanation for the amelioration but not total loss of the *yo*-to-*yo* (putative lexical boost) and *I*-to-*yo* (putative translation equivalent) priming effects in the presence of intervening material concerns the nature of the phenomena under study. In experimental studies, lexical boost effects result from overlap of primarily main verbs and head nouns, but 1sg pronouns may not undergo lexical boost in the same way; pronouns carry greater functional load, and therefore priming may be considered to be more 'abstract', even in the case of identity priming. Furthermore, the notion of translation equivalents is even less straightforward for grammatical morphemes such as subject pronouns than it is for nouns and verbs. The Spanish expressed 1sg pronoun *yo* (vs. an unexpressed mention) has been said to correspond to the English expressed pronoun *I* (vs. an unexpressed mention), or, alternatively, to an English stressed (vs. unstressed) *I* (Givón, 1983, p. 17; Payne, 1997, p. 43). But the inter-linguistic functional equivalence of grammatical morphemes or structure types is an empirical question (as noted by Otheguy, 2004, p. 177), and in fact the constraints on *yo* expression in Spanish, though similar, are not identical to those on expression and stress on *I* in English (Torres Cacoullos & Travis, 2015; Travis & Torres Cacoullos, 2014).

Thus, with our third question we sought to explore further the lexical boost account of priming, by testing for non-coreferential subject priming, in which identity of form is absent not only across languages, but also within the one language. Though weaker than Coreferential Subject Priming, *yo* was found to be primed by non-coreferential Spanish subject pronouns (Table 5, Table 6). This within-language non-coreferential subject priming effect indicates that we must go beyond a lexical boost *vs.* translation equivalent account of stronger within- than cross-language priming. How, then, may we account for the fact that *yo*-to-*yo* priming was stronger than *I*-to-*yo* priming?

We propose that the question may be answered by interpreting priming as a gauge of association between constructions. To further explore the interplay between pronominal priming and lexically-specific features of the grammar, we considered priming in interaction with particular constructions, and observed that the priming effect was smaller when targets contain cognition verbs, which favor expressed 1sg subjects more than other verbs (Table 7, Table 8).

Verb type or class effects suggest that we may think of the greater or smaller strength of priming as a measure of stronger or weaker postulated associations between the more particular constructions constituting a general structure type, such as [pronoun + verb]. That is, differential priming effects indicate degrees of association, such as between [(*yo*) + Cognition Verb_{1sg}] and [(*yo*) + Verb_{1sg}] or between [(*yo*) + Verb_{1sg}] and [(Pronoun) + Verb]. The less robust cross-language priming, we suggest, is just a special case of degrees of associations between constructions, here between [(*yo*) + $\text{Verb}_{1sg}]_{\text{Spanish}}$ and $[I + \text{Verb}_{1sg}]_{\text{English}}$.

8. Conclusions

We set out to examine here the question of 'overlap' in bilingual grammars, through an exploration of the relative strength of within- and cross-language priming. We have found that cross-language priming, compared with within-language priming, is narrower (applying only to coreferential pronoun *I*), less persistent (operative at one intervening clause at most between prime and target), and weaker (in contexts of maximal subject continuity, doubling the rate of *yo* expression vis-à-vis a previous unexpressed subject, whereas a *yo* prime increased it six-fold). We proposed that this is best interpreted as a reflection of the degree of association between subject pronoun + verb constructions, which, although existing both within and across languages, are attenuated in the latter case. Given this, it need not be surprising, then, that despite the overlap in their linguistic systems (here, associations between 1sg subject pronoun constructions indicated by cross-language priming), Spanish language-internal constraints on subject expression are not weakened in the Spanish of New Mexico bilinguals, as seen here for the subject continuity and cognition verb effects (Torres Cacoullos & Travis, 2015).

We have also demonstrated that priming is modulated by other factors conditioning variable subject expression in Spanish, namely subject continuity and usage patterns

https://doi.org/10.1017/S1366728915000127 Published online by Cambridge University Press

associated with particular classes (and tokens) of verbs. This suggests a need for a model of representation that accounts for variation in representation attributable to a speaker's experience with their language, or languages. Chang, Dell & Bock (2006), working with monolingual data, have developed a connectionist model, in which grammatical representations are acquired through experience, inducing abstract knowledge through analysis of the distributional properties of the input. The model successfully explains many priming effects in the experimental literature. From what we have seen here, such a model could extend to the bilingual data, provided that, as for monolingual data, the associations between constructions – form-meaning pairings grounded in their linguistic and extra-linguistic contexts of occurrence – are understood.

The results presented here highlight the value of naturalistic data as a means to challenge and sharpen psycholinguistic theory. Since studies of natural speech require consideration of the contextual distributions of targets and primes, they provide an opportunity to gauge whether and how priming interacts with co-occurring linguistic features in configurations as speakers encounter them in everyday language use.

Appendix I: Transcription Conventions (Du Bois, Schuetze-Coburn, Cumming & Paolino, 1993)

Carriage return	new Intonation Unit	[ ]	overlapped speech
.	final intonation contour	[2 2]	used to distinguish consecutive
,	continuing intonation contour		overlaps
?	appeal intonation contour	<@ @>	speech produced while laughing
--	truncated intonation contour	(H)	in breath
-	truncated word	(( ))	transcriber's comment
..	short pause (0.5 secs)	%	glottal stop
. . .	medium pause (0.5–0.7 secs)	<VOX VOX>	speech uttered with a
. . . ( )	timed pause (over 0.7 secs)		marked voice quality
=	lengthened syllable		

References

Aaron, J.E. (forthcoming). Lone English-origin nouns in Spanish: The precedence of community norms (Special issue, Gauging convergence on the ground: code-switching in the community, edited by Catherine E. Travis and Rena Torres Cacoullos). *International Journal of Bilingualism*.

Baayen, R.H. (2008). *Analyzing Linguistic Data: A Practical Introduction to Statistics using R*, Cambridge: Cambridge University Press.

Bates, D., & Mächler, M. (2010). lme4: Linear mixed-effects models using S4 classes. R package version 0.999375–33.

Benevento, N., & Dietrich, A. (forthcoming). I think, therefore *digo yo*: Variable position of the 1sg subject pronoun in New Mexican Spanish–English code-switching (Special issue, Gauging convergence on the ground: code-switching in the community, edited by Catherine E. Travis and Rena Torres Cacoullos). *International Journal of Bilingualism*.

Bentivoglio, P. (1983). Topic continuity and discontinuity in discourse: A study of spoken Latin-American Spanish. In T. Givón (ed.), *Topic continuity in discourse: A quantitative cross-linguistic study*, (pp. 255–311). Amsterdam: John Benjamins.

Bernolet, S., & Hartsuiker, R.J. (2010). Does verb bias modulate syntactic priming. *Cognition, 114(3),* 455–461.

Bernolet, S., Hartsuiker, R.J., & Pickering, M.J. (2007). Shared syntactic representations in bilinguals: Evidence for the role of word-order repetition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33(5),* 931–949.

Bernolet, S., Hartsuiker, R.J., & Pickering, M.J. (2009). Persistence of emphasis in language production: A cross-linguistic approach. *Cognition, 112,* 3000–3317.

Bernolet, S., Hartsuiker, R.J., & Pickering, M.J. (2013). From language-specific to shared syntactic representations: The influence of second language proficiency on syntactic sharing in bilinguals. *Cognition, 127(287–306)*.

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

Bock, J.K. (1986). Syntactic persistence in language production. *Cognitive Psychology, 18,* 355–387.

Bock, J.K., & Griffin, Z.M. (2000). The persistence of structural priming: Transient activation or implicit learning. *Journal of Experimental Psychology: General, 129(2),* 177–192.

Branigan, H.P., Pickering, M.J., & Cleland, A.A. (1999). Syntactic priming in written production: Evidence for rapid decay. *Psychonomic Bulletin & Review, 6,* 535–540.

Brysbaert, M., & Duyck, W. (2010). Is it time to leave behind the Revised Hierarchical Model of bilingual processing after 15

https://doi.org/10.1017/S1366728915000127 Published online by Cambridge University Press

years of service? *Bilingualism: Language and Cognition, 13,* 359–371.

Bybee, J. (2010). *Language, usage and cognition*, Cambridge: Cambridge University Press.

Cameron, R. (1994). Switch reference, verb class and priming in a variable syntax. *Papers from the Regional Meeting of the Chicago Linguistic Society: Parasession on variation in linguistic theory, 30(2),* 27–45.

Chang, F., Dell, G.S., & Bock, J.K. (2006). Becoming syntactic. *Psychological Review, 113,* 234–272.

Clyne, M., Eisikovits, E., & Tollfree, L. (2001). Ethnic varieties of Australian English. In D. Blair & P. Collins (eds.), *English in Australia*, (pp. 223–238). Amsterdam: John Benjamins.

Croft, W. (1995). Intonation units and grammatical structure. *Linguistics, 33,* 839–882.

Croft, W. (2007). Intonation units and grammatical structure in Wardaman and in cross-linguistic perspective. *Australian Journal of Linguistics, 27(1),* 1–39.

Croft, W., & Cruse, D.A. (2004). *Cognitive linguistics*, Cambridge: Cambridge University Press.

Desmet, T., & Declercq, M. (2006). Cross-linguistic priming of syntactic hierarchical configuration information. *Journal of Memory and Language, 54,* 610–632.

Du Bois, J.W., Schuetze-Coburn, S., Cumming, S., & Paolino, D. (1993). Outline of discourse transcription. In J. Edwards & M. Lampert (eds.), *Talking data: Transcription and coding in discourse*, (pp. 45–89). Hillsdale, NJ: Lawrence Erlbaum Associates.

Givón, T. (1983). Topic continuity in spoken English. In T. Givón (ed.), *Topic continuity in discourse: A quantitative cross-linguistic study*, (pp. 343–363). Amsterdam: John Benjamins.

Goldberg, A.E. (2006). *Constructions at work: The nature of generalization in language*, Oxford: Oxford University Press.

Gonzales Berry, E., & Maciel, D.R. (eds) (2000). *The contested homeland: A Chicano history of New Mexico*, Albuquerque, University of New Mexico Press.

Gonzales, M.D. (1999). Crossing social and cultural borders: The road to language hybridity. In L. Galindo & M. D. Gonzales (eds.), *Speaking Chicana: Voice, power and identity*, (pp. 13–38). Tucson: University of Arizona Press.

Gries, S.T. (2005). Syntactic priming: A corpus-based approach. *Journal of Psycholinguistic Research, 34(4),* 365–399.

Hartsuiker, R.J., & Pickering, M.J. (2008). Language integration in bilingual sentence production. *Acta Psychologica, 128(3),* 479–489.

Hartsuiker, R.J., Pickering, M.J., & Veltkamp, E. (2004). Is syntax separate or shared between languages. *Psychological Science, 15(6),* 409–414.

Jaeger, T.F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language, 59(434–446)*.

Jaeger, T.F., & Snider, N. (2013). Alignment as a consequence of expectation adaptation: Syntactic priming is affected by the prime's prediction error given both prior and recent experience. *Cognition, 127(1),* 57–83.

Kantola, L., & van-Gompel, R.P.G. (2011). Between- and within-language priming is the same: Evidence for shared bilingual representations. *Memory and Cognition, 39,* 276–290.

Kidd, E., Tennant, E., & Nitschke, S. (2015). Shared abstract representation of linguistic structure in bilingual sentence comprehension. *Psychonomic Bulletin & Review, DOI 10.3758/s13423-014-0775-2*.

Kootstra, G.J., Van Hell, J.G., & Dijkstra, T. (2012). Priming of code-switching in sentences: The role of lexical repetition, cognates, and language proficiency. *Bilingualism: Language and Cognition, 15,* 797–819.

Labov, W. (1984). Field methods of the project on linguistic change and variation. In J. Baugh & J. Sherzer (eds.), *Language in use: Readings in sociolinguistics*, (pp. 28–53). Englewood Cliffs, NJ: Prentice Hall.

Levelt, W. (1989). *Speaking: From intention to articulation*, Cambridge, MA: MIT Press.

Lipski, J.M. (2008). *Varieties of Spanish in the United States*, Washington, DC: Georgetown University Press.

Loebell, H., & Bock, K. (2003). Structural priming across languages. *Linguistics, 41(5),* 791–824.

Malhotra, G., Pickering, M.J., Branigan, H.P., & Bednar, J. (2008). On the persistence of Structural priming: Mechanisms of decay and Influence of word-forms, *Proceedings of the 30th Annual Conference of the Cognitive Science Society*, (pp. 657–662). Cognitive Science Society.

McKee, R., Schembri, A., McKee, D., & Johnston, T. (2011). Variable "subject" presence in Australian Sign Language and New Zealand Sign Language. *Language Variation and Change, 23,* 375–398.

Meijer, P.J.A., & Fox Tree, J.E. (2003). Building syntactic structures in speaking: A bilingual exploration. *Experimental Psychology, 50,* 184–195.

Meyerhoff, M. (2009). Replication, transfer, and calquing: Using variation as a tool in the study of language contact. *Language Variation and Change, 21(3),* 297–317.

Otheguy, R. (2004). Single-language and code-switching strategies in immigrant and heritage varieties: Spanish subject personal pronouns in Toribio's cross-modal hypothesis. *Bilingualism: Language and Cognition, 7(2),* 175–177.

Otheguy, R., & Zentella, A.C. (2012). *Spanish in New York: Language contact, dialectal leveling, and structural continuity* Oxford: Oxford University Press.

Payne, T.E. (1997). *Describing morphosyntax: A guide to field linguists*, Cambridge: Cambridge University Press.

Pickering, M.J., & Branigan, H.P. (1998). The representation of verbs: Evidence from syntactic priming in language production. *Journal of Memory and Language, 39(4),* 633–651.

Poplack, S. (1980). The notion of the plural in Puerto Rican Spanish: Competing constraints on (s) deletion. In W. Labov (ed.), *Locating language in time and space*, (pp. 55–67). New York: Academic Press.

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

https://doi.org/10.1017/S1366728915000127 Published online by Cambridge University Press

Poplack, S., & Levey, S. (2010). Contact-induced grammatical change: A cautionary tale. In P. Auer & J. E. Schmidt (eds.), *Language and Space: An international handbook of linguistic variation*, vol. 1: Theories and methods, (pp. 391–419). Berlin: Mouton de Gruyter.

R Core Team (2012). *R: A language and environment for statistical computing*, Vienna, Austria: R Foundation for Statistical Computing, http://www.R-project.org

Salamoura, A., & Williams, J.N. (2006). Lexical activation of crosslinguistic priming. *Bilingualism: Language and Cognition, 9,* 299–307.

Salamoura, A., & Williams, J.N. (2007). Processing verb argument structure across languages: Evidence for shared representations in the bilingual lexicon. *Applied Psycholinguistics, 28,* 627–660.

Savage, C., Lieven, E., Theakston, A., & Tomasello, M. (2006). Structural priming in language acquisition: The persistence of lexical and structural priming in 4-year-olds. *Language Learning and Development, 2,* 27–49.

Scherre, M.M.P. (2001). Phrase level parallelism effect on noun phrase number agreement. *Language Variation and Change, 13(1),* 91–107.

Scherre, M.M.P., & Naro, A.J. (1991). Marking in discourse: 'Birds of a feather'. *Language Variation and Change, 3(1),* 23–32.

Scherre, M.M.P., & Naro, A.J. (1992). The serial effect on internal and external variables. *Language Variation and Change, 4(1),* 1–13.

Schoonbaert, S., Hartsuiker, R.J., & Pickering, M.J. (2007). The representation of lexical and syntactic information in bilinguals: Evidence from syntactic priming. *Journal of Memory and Language, 56(2),* 153–171.

Serratrice, L. (2009). Carving up referential space: A priming account of pronoun use in Italian. In J. Chandlee, M. Franchini, S. Lord & G.-M. Rheiner (eds.), *Proceedings of the 33rd Annual Boston University Conference on Language Development*, (pp. 468–479). Somerville, MA: Cascadilla Press.

Shin, J.A., & Christanson, K. (2009). Syntactic processing in Korean-English bilingual production: Evidence from cross-linguistic structural priming. *Cognition, 112,* 175–180.

Silva-Corvalán, C. (1982). Subject expression and placement in Mexican-American Spanish. In J. Amastae & L. Elías Olivares (eds.), *Spanish in the United States: Sociolinguistic aspects*, (pp. 93–120). New York: Cambridge University Press.

Silva-Corvalán, C. (1994). *Language contact and change: Spanish in Los Angeles*, Oxford: Clarendon Press.

Silva-Corvalán, C. (2001). *Sociolingüística y pragmática del español* (Georgetown Studies in Spanish Linguistics), Washington, DC: Georgetown University Press.

Szmrecsanyi, B. (2005). Language users as creatures of habit: A corpus-based analysis of persistence in spoken English. *Corpus Linguistics and Linguistic Theory, 1(1),* 113–149.

Torres Cacoullos, R. (2015). Gradual loss of analyzability: Diachronic priming effects. In A. Adli, G. Kaufmann & M. García (eds.), *Variation in language: usage-based vs. system-based approaches* (pp. 267–289). Berlin: de Gruyter.

Torres Cacoullos, R., & Travis, C.E. (2010). Variable *yo* expression in New Mexico: English influence? In S. Rivera-Mills & D. J. Villa (eds.), *Spanish of the U.S. Southwest: A language in transition*, (pp. 185–206). Madrid: Iberoamericana.

Torres Cacoullos, R., & Travis, C.E. (2011). Testing convergence via code-switching: Priming and the structure of variable subject expression *International Journal of Bilingualism, 15(3),* 241–267.

Torres Cacoullos, R., & Travis, C. (2015). Foundations for the study of subject subject pronoun realization in Spanish and English: Assessing inter-linguistic (dis)similarity via intra-linguistic inherent variability. In A. M. Carvalho, R. Orozco & N. L. Shin (eds.), *Subject pronoun expression in Spanish: A Cross-dialectal Perspective*, (pp. 83–102). Georgetown: Georgetown University Press.

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

Torres Cacoullos, R., & Travis, C.E. (forthcoming). Gauging convergence on the ground: Code-switching in the community *International Journal of Bilingualism (Special issue edited by Catherine E. Travis and Rena Torres Cacoullos).*

Torres Cacoullos, R., & Travis, C.E. (2015). Two languages, one effect: Structural priming in code-switching. *Bilingualism: Language and Cognition (Special issue edited by Margaret Deuchar)*. http://journals.cambridge.org/repo_A965kR4kPSaNSo

Travis, C.E. (2005). The yo-yo effect: Priming in subject expression in Colombian Spanish. In R. Gess & E. J. Rubin (eds.), *Theoretical and experimental approaches to Romance linguistics: Selected papers from the 34th Linguistic Symposium on Romance Languages, 2004*, (pp. 329–349). Amsterdam / Philadelphia: John Benjamins.

Travis, C.E. (2007). Genre effects on subject expression in Spanish: Priming in narrative and conversation. *Language Variation and Change, 19(2),* 101–135.

Travis, C.E., & Lindstrom, A. (2014). Variable subject expression in two registers of spoken English: a non-null subject language? Paper presented at the Australian Linguistics Society Annual Conference, University of Newcastle, Australia.

Travis, C.E., & Torres Cacoullos, R. (2012). What do subject pronouns do in discourse? Cognitive, mechanical and constructional factors in variation. *Cognitive Linguistics, 23(4),* 711–748.

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

Travis, C.E., & Torres Cacoullos, R. (2014). Stress on I: Debunking unitary contrast accounts *Studies in Language, 38(2),* 360–392.

Weiner, E.J., & Labov, W. (1983). Constraints on the agentless passive. *Journal of Linguistics, 19(1),* 29–58.

Wilson, D.V., & Dumont, J. (forthcoming). The emergent grammar of bilinguals: The Spanish verb *hacer* 'do' with a bare English infinitive (Special issue, Gauging convergence on the ground: code-switching in the community, edited by Catherine E. Travis and Rena Torres Cacoullos). *International Journal of Bilingualism*.

https://doi.org/10.1017/S1366728915000127 Published online by Cambridge University Press