"*So* respetamos un tradición del uno al otro"

So and *entonces* in New Mexican bilingual discourse*

Jessi Elana Aaron

This paper examines the use of two discourse markers, English-origin *so* and Spanish-origin *entonces*, in New Mexican bilingual speech. Both forms appear in the mixed speech and in the otherwise monolingual English and monolingual Spanish of bilingual speakers in New Mexico. Through a quantitative examination of the 413 uses of *so* ($n=289$) and *entonces* ($n=124$) in a 204,000-word corpus, it is found that both perform the same discourse functions with the same relative frequency, thus showing no evidence of specialization. It is also shown that *so* occurs with code-switches significantly more often than *entonces*, and therefore may function as a "trigger" for code-switches (cf. Clyne 1997). This switching is not preferred in certain contexts, but rather follows the same patterns as in monolingual discourse. Lastly, it is found that the use of *so* in monolingual Spanish and monolingual English shows no significant differences: it is used in the same way in both modes.

Keywords: bilingualism, discourse markers, Spanish, variation

1. Introduction

New Mexican Spanish is the oldest variety of Spanish in what is today the United States. Spanish was first brought to the region in the sixteenth century by Spanish colonizers before pilgrims landed at Plymouth Rock. Until the twentieth century, New Mexico remained isolated from other Spanish-speaking communities, with Santa Fe lying 1500 miles from Mexico City. This isolation allowed a unique New Mexican Spanish to develop (Bills and Vigil 1999). At the same time, this variety has been in contact with English since the introduction

Spanish in Context 1:2 (2004), **161–179**.
ISSN 1571-0718 / E-ISSN 1571-0726

of the railroad and public school system in 1880–90s, and shift to English has been documented in a number of studies (Bills et al. 1995).

Because of this unique situation, New Mexican Spanish has been one of the most studied varieties of Spanish in the United States, beginning as early as the beginning of the twentieth century with the pioneering work of Aurelio M. Espinosa (Espinosa 1914–15, 1975). One of the most noted characteristics of New Mexican Spanish is the use of both Spanish- and English-origin words in the same utterance, i.e. code-mixing.[1] Despite the widespread interest in this variety, however, much of the previous work done on New Mexican Spanish has depended on lists of words. Only very recently has there been any quantitative investigation of speech patterns in this community (for example, Torres Cacoullos and Aaron 2003b).

The empirical investigation of code-mixing in this variety, however, is still in its initial stages. One as of yet untouched phenomenon in New Mexican Spanish is the coexistence of both English- and Spanish-origin discourse markers. Though the designation of an item as a 'discourse marker' can at times be debatable, Torres (2002:65) notes that "at a basic level most linguists would agree that discourse markers contribute to the coherence of the discourse by signaling or marking a relationship across utterances". Discourse markers may have grammatical functions, as well. This paper will look at the use of two well-recognized discourse markers, English-origin *so* and Spanish-origin *entonces*, in New Mexican bilingual speech. Both of these forms appear in the language-mixed speech and in the otherwise monolingual English and monolingual Spanish of bilingual speakers in New Mexico.

Though this is the first study to examine this aspect of bilingual speech in New Mexico, the issue of discourse markers in the speech of bilingual speakers has been explored, though not extensively, in other varieties and language pairs. Torres (2002), for example, looks at the use of English-origin discourse markers in Puerto Rican Spanish. She argues that, while the uses of English- and Spanish-origin discourse markers do overlap, there is no evidence that the Spanish-origin discourse markers are losing ground in favour of the English-origin discourse markers. Instead, she finds that level of bilingualism is a factor in the relative frequency of English-origin discourse markers. Solomon (1995), in contrast, argues for separate discourse functions for semantically similar *ka(p)* (a Yucatec Maya term) and Spanish-origin *entonces* in the discourse of bilingual Yucatec Maya speakers. Brody (1995) offers an account of the borrowing of Spanish discourse markers into indigenous American languages, and suggests that the borrowing of these and other particles may be an indicator of

language change, with the donor-language forms replacing and leading to the demise of the recipient-language forms, which are lost. Similarly, Salmons (1990) finds that the use of English-origin discourse markers in German discourse of German-American bilinguals leads to a nearly categorical loss of the native German discourse-marker system.

Other studies have argued that other-language-origin discourse markers (e.g. English-origin *so* in Spanish discourse) can serve as triggers for code-switching (Pfaff 1982; Brody 1987). Also, discourse markers remain one among many convergence sites, or places in which two (or more) languages overlap or converge (either due to similarities in related languages or to established borrowing), in trilingual mixing for Clyne (1972, 1997; Clyne and Cassia 1999), such that speakers associate them equally with two (or more) languages.[2] The use of discourse markers themselves, however, remains out of focus and is not the primary object of Clyne's studies.

Overall, then, there are two prevalent hypotheses in studies on discourse markers in bilingual speech: first, that semantically similar discourse markers develop specialized discourse functions and thus are maintained in complementary distribution (for example Solomon 1995), or, conversely, that native and other-language-origin discourse markers are in variation, with the non-native item either co-existing in stable variation with (Torres 2002) or eventually replacing (Salmons 1990) the native one; and second, that other-language-origin discourse markers can trigger code-switches (Pfaff 1982; Brody 1987; Clyne 1997: 107; Clyne and Cassia 1999: 68; Torres 2002: 67). If we apply this to the examination of *so* and *entonces*, we are left with the following hypotheses:

1. *So* is used preferentially to perform only some of the discourse functions performed by *entonces* in monolingual discourse, while *entonces* is still preferred in others. In other words, these two forms are acquiring specialized discourse functions. If this is false, there are two alternatives: a. *so* and *entonces* perform the same discourse functions and are in a state of stable variation (thus not indicating a change in progress), or b. *so* and *entonces* perform the same discourse functions, and *so* is in a process of replacing *entonces*.
2. The use of *so* triggers code-switches more often than the use of *entonces*.

The first hypothesis will be explored through a detailed analysis of the semantics and discourse functions of these two markers. *Entonces* and *so*, etymologically Spanish and English, respectively, are often considered to be similar semantically and to perform similar functions in discourse (Travis forthcoming; cf. also Schiffrin 1987: 202; Sweetser 1990; Rendle-Short 2003). The functional analysis

will be based on Travis's (forthcoming) four-part definition of *entonces*. She argues that discourse marker *entonces* has four separate yet interrelated meanings: "introducing a result, introducing speech acts based on a conclusion, following speech acts based on a conclusion, and marking discourse progression" (Travis forthcoming). Similarly, Torres (2002: 69, 74) notes that *so* can be used in cause or result contexts, to introduce an evaluation, or to move a narrative along.

The second hypothesis, which is that other-language-origin discourse markers (i.e. discourse markers that originate etymologically in a language other than Spanish) or "items from an overlapping area between two languages" (i.e. items which are shared by and thus the same in two languages), in this case, the English-origin *so*, can trigger a code switch (Pfaff 1982; Brody 1987; Clyne 1997: 107; Clyne and Cassia 1999: 68) will be tested. In order to test this hypothesis, patterns of switching with both *so* and *entonces* will be explored, as well as their relationship to the semantics of the markers.

2. Corpus and data

The data for the present study were taken from 11 interviews of the *New Mexico Colorado Spanish Survey* (NMCOSS) and 8 interviews from the University of New Mexico Barelas Corpus (Barelas 2001). The 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 (for a detailed description of the corpus, cf. Bills and Vigil 1999). The speakers are primarily residents of the counties of Bernalillo, Mora, Taos, and Río Arriba. The Barelas corpus consists of a total of 11 sociolinguistic interviews collected by graduate students in 2001 in a course in Hispanic sociolinguistics at the University of New Mexico, in the historic, predominantly Hispanic community of Barelas in Albuquerque, New Mexico. Each student conducted an hour-long conversational interview with a bilingual adult living and/or working in Barelas who had spoken some variety of Mexican Spanish, which includes New Mexican Spanish, since childhood. For both corpora, interviews were only included if speakers used both the discourse markers *so* and *entonces* in their speech. Exclusion of speakers without this use assured that the speakers who were included did, indeed, have the option of using both *so* and *entonces* as a part of their linguistic repertoires. The sample from NMCOSS is made up of 131,527 words, and the eight Barelas interviews comprise a total of 72,193 words, thus making the entire sample size around 203,700 words.

Since the purpose of both corpora was to collect a sample of New Mexican Spanish, the majority of the data are from segments of principally monolingual Spanish. Nevertheless, some speakers use single English-origin nouns in their discourse,[3] code-switch intra- or intersententially, or speak in stretches of primarily monolingual English. None of these speech styles was disregarded, and all were included in the sample. The relationship of these different speech styles to the use of *so* and *entonces* will be discussed below.

For this study, the function of interest is the discourse marker, shown in (1) and (2).

(1) …todo entendían y lo hablaban. *So* los míos
all understand-3P-IMP and 3S-ACC speak-3P-IMP *so* the 1S-POSS
siempre han habla'o las dos idiomas.
always have-3P speak-PART the two languages
'… they understood everything and they spoke it. *So* mine [kids] have always spoken both languages.' (NMCOSS 117)

(2) Dice, "No, ésas no son las tortillas." Le dije,
say-3S NEG those NEG be-3P the tortillas 3S-DAT say-1S-PRET
"*Entonces*, ¿qué son?"
entonces what be-3P
'He says, "No, those aren't the tortillas." I said to him, "*Entonces* what are they?' (Barelas JA)

What can be noted about discourse-marker functions is that they are syntactically and semantically independent, i.e. they can be removed from the utterance and it is left both syntactically intact and with its propositional content unchanged. Both *so* and *entonces*, however, have other, non-discourse-marker functions, in which they function as different parts of speech. The uses of *so* excluded were its use as an intensifier (3) and as a subordinating conjunction (4), as well as its adverbial use, such as *I don't think* ***so***.

(3) *Pero, los* grades are *so* bad *también, oiga.*
but the grades are *so* bad also hear-IMPER
'But the *grades are so bad*, too, hey.' (NMCOSS 88)

(4) …but he sent them over there *so* they would learn better English. (NMCOSS 270)

Other non-discourse-marker uses of *entonces* are its use as a temporally deictic noun (5) or temporal adverb (6). For the purposes of this paper, I will refer to both of these as "temporal" uses.

(5) …todos éramos mexicanos y hablábamos en español. Desde
all be-1P-IMP Mexicans and speak-1P-IMP in Spanish since
entonces ha comenzado una, una actitud de que oh de qué
entonces have-3S begin-PART a a attitude of COMP oh of what
vergüenza no poder hablar inglés.
shame NEG be.able-INF speak-INF English
'We were all Mexicans and we spoke in Spanish. Since *entonces* an attitude has started that, oh, that how embarrassing not to be able to speak English.' (NMCOSS 117)

(6) Cuando venga el tiempo, va a estar bien. *Entonces*
when come-3S-SUBJ the time go-3S PREP be-INF well *entonces*
ya va a estar bien.
already go-3S PREP be-INF well
'When the time comes, it'll be okay. *Entonces* it'll be okay.' (Barelas MA)

Since the use of discourse markers is optional, any evaluation of their use in bilingual speech must be made based not on case-by-case analysis, but on the general distributional patterns of the items in question. Because of this, all occurrences of the discourse marker *so* and all occurrences of *entonces*, both as a discourse marker and in temporal uses (cf. Garcés Gómez 1994; Travis forthcoming), were extracted, resulting in a total of 289 instances of *so* and 179 instances of *entonces*. The 55 temporal uses of *entonces*, as in (5) and (6), were then separated, leaving a new total of 124 tokens of *entonces*. Though many of the discourse-marker uses of *entonces* may also permit a temporal reading, a use was only considered temporal if it (a) was used as a noun, or (b) if it was used in an obviously deictic comparison, as in (6). Though the non-discourse-marker use of *entonces* will not be central to this study, its presence is worth noting: an overwhelming 31% (55/179) of all uses of *entonces* in these data are as a temporal adverbial, in stark contrast to the results obtained by Travis (forthcoming), who, out of 201 tokens, found no such use in conversational monolingual Spanish discourse in Colombia. This point will be returned to later.

3. Results

English-origin discourse marker *so* is widespread in the speech of New Mexican bilinguals: of a sample of 21 speakers from the NMCOSS, 11 speakers used it, and in the Barelas corpus, 8 of 11 speakers used it. Not only is *so* widespread, it is also frequent among those speakers who use it. In fact, as shown in Table 1,

Table 1. Relative frequency of *so* and *entonces* in each corpus

	Barelas		NMCOSS		Total	
	N	%	N	%	N	%
so	100	52	189	85	289	70
entonces	91	48	33	15	124	30
Total	191	100	222	100	413	100

for both corpora, taking into account all occurrences of *so* and *entonces* in discourse-marker functions — and excluding the other, non-discourse-marker functions mentioned above —, *so* is relatively more frequent among those who use it than *entonces*. Again, speakers who did not use both *so* and *entonces* were excluded from this study.

As we can see, Barelas speakers use *so* and *entonces* at nearly equal rates: in this corpus, *so* occurred 100 times as a discourse marker, while *entonces* as a discourse marker occurred 91 times. Of these two uses combined, then, *so* makes up a little over half, at 52%. NMCOSS speakers, on the other hand, use *so* much more often, with a relative frequency of 85%. In total, *so* is used 40 percentage points more often than *entonces* (70% vs. 30%).

3.1 Specialized discourse functions?

In order to begin to explore the question of the status of these two discourse markers in New Mexican bilingual speech, however, we must look beyond simple frequency counts. Instead, it is important to look at the functions each item has in the discourse of bilingual speakers. As mentioned above, Travis (forthcoming) proposes four basic functions for *entonces*: (1) introducing a result, (2) introducing a conclusion, (3) following a conclusion, and (4) marking discourse progression. Since similar functions have been proposed for *so* in monolingual English (Rendle-Short 2003) and bilingual English/Spanish (Torres 2002), and since these were found to be applicable to my data, these same categories were also adopted for the classification of *so*. The third use, as defined, however, is only relevant to *entonces*, since *so* is not used after a conclusion. Here *entonces* is being used in a function that does not overlap with the use of *so*. In addition, there is a somewhat common use of *so* not found with *entonces* in the data: a *so* that ends an utterance, leaving an (often implicit) utterance unexpressed, as in (10).[4] An example of each function can be seen in

(7)–(11). Example (7) shows what in this work will be called the Resultative use, (8) the Conclusion use, (9) the Final *entonces* use, (10) the Unexpressed utterance use, and (11) the Discourse progression use.

These discourse markers can be used to introduce a result. In (7a), the speaker talks about how her younger brother began to call her "Mom" and her husband "Daddy" *as a result of* hearing her own children call her that. In (7b), the speaker says that people warned her that she would be more afraid *as a result of* travelling on a small airplane.[5] In (7), then, the discourse markers are used to introduce a *direct, real-world consequence* of the situation described in the preceding utterance.

(7) introducing a result

a. Y él le, le decía papá, porque oía los
and he 3S-DAT 3S-DAT say-3S-IMP dad because hear-3S-IMP the
demás, estaban chicos todavía mis hijos, oía,
rest be-3P-IMP small still my children hear-3S-IMP
oía que ellos a mí me llamaban Mamá, y
hear-3S-IMP COMP they PREP 1S-DAT 1S-DAT call-3P-IMP Mom and
a, a él le decían Daddy. *So* él también se
PREP PREP he 3S-DAT say-3P-IMP daddy *so* he also 3-REFL
puso a decirlo
put-3S-PRET PREP say-INF+3S-ACC
'And he would call him dad, because he would hear the others, my kids were still little, he would hear, he would hear that they would call me Mom, and they would call him Daddy. *So* he started saying it, too' (Barelas IN)

b. Bueno también me dijieron que no me
well also 1S-DAT say-3P-PRET comp NEG 1S-REFL
fuera a subir en un aeroplanito de esos
go-1S-IMP-SUBJ PREP ascend-INF in one airplane-DIM of those
chiquitos, porque *entonces* me iba a dar más
little-DIM because *entonces* 1S-DAT go-3S-IMP prep give-INF more
miedo.
fear
'Well they also told me not to get on one of those little airplanes, because *entonces* I would be more scared.' (NMCOSS 318)

These markers may also be used to introduce a conclusion. In (8a), the speaker arrives at the conclusion that she has lived in the town for 47 years based on

what she has just said. In (8b), the speaker comes to the conclusion that her interlocutor is cold-blooded based on what the interlocutor just said.

(8) introducing a conclusion

a. mi hija tenía seis años cuando nos
POSS daughter have-3S-IMP six years when 1P-REFL
vinimos pa'cá y ahora tiene cincuenta y tres *so*
come-1P-PRET over.here and now have-3s fifty and three *so*
tenía cuaren-, hace cuarenta y siete años que
have-3S-IMP for- make-3s forty and seven years COMP
vivo aquí…
live-1s here
'… my daughter was six when we came over here, and now she's fifty-three, *so* she was for-, I have lived here for forty-seven years…' (NMCOSS 318)

b. A: Hace frío.
A: do-3s cold
B: ¿Hace frío? [Risas] *Entonces* eres friolenta como yo.
B: do-3s cold [Laugh] *entonces* be-2s cold-blooded like I
A: It's cold.
B: It's cold? [Laughter] *Entonces* you're cold-blooded like me. (Barelas JA)

In (8), then, unlike in (7), there is no *result* being described. In (8a), the speaker has not lived in the town for forty-seven years *because* her daughter is fifty-three, but rather she can conclude that she has lived there for forty-seven years based on this information. Her length of time living in the town does not result directly from her daughter's age. In (8b), A is not cold-blooded *because* it is cold; instead, B can conclude that A is cold-blooded based on A's previous turn, in which she states that it is cold. A is not cold-blooded as a result of the cold, but would remain cold-blooded even if it were warm outside. While many authors collapse examples such as (7) and (8) (e.g. Pons Bordería 1998), they are fundamentally different. Examples such as those in (7) introduce real-world results, while examples such as those in (8) require inferential processes. Since it is, in fact, quite possible to distinguish these two contexts, I have kept them separate for this analysis.

Similar to the introduction of a conclusion, *entonces* may be used following a conclusion. In (9), *entonces* follows the conclusion that they missed something based on new information provided by the interlocutor.

(9) final *entonces*
Oh, sí? Ah, nos perdimos eso *entonces.*
oh yes ah 1P-REFL lose-1P-PRET that *entonces*
'Oh, really? We missed that, *entonces.*' (NMCOSS 10)

In (10), the cut-off *so*, which ends this speaker's turn in the discourse, leads up to an unexpressed utterance. This utterance could be a conclusion, a result, or a myriad of other possibilities, and so examples such as (10) cannot be collapsed with other categories discussed thus far.

(10) preceding an unexpressed utterance
de una vez fui al doctor y no me hizo mal
of one time go-1S-PRET to.the doctor and NEG 1S-DAT do-3S-PRET bad
la azúcar que comí, *so* —
the sugar COMP eat-1S-PRET *so*
'… I went to the doctor right away and the sugar I ate didn't hurt me, *so* —'
(NMCOSS 318)

In (11), both discourse markers make no special link (i.e. conclusive or resultative) to the previous discourse, but instead function to move the discourse along. This category includes a wide range of uses, including the reintroduction of an old topic or the introduction of a new topic, and is the most diverse of all of the categories.

(11) marking discourse progression
a. Y Angélica es muy alta y güera también. *So*
and Angélica be-3s very tall and light-skinned also *so*
dice Angélica,
say-3s Angélica,
'Mom, I felt so bad when I walked out because they all said to each other, "Since when do we let whites join our club?"
"And Angélica is really tall and light-skinned, too. *So* Angélica says, "Mom, I felt so bad when I walked out because they all said to each other, "Since when do we let whites join our club?"' (Barelas PC)
b. pero si alguien quiere vivir a gusto en Albuquerque,
but if someone want-3s live-INF PREP pleasure in Albuquerque
es en Barelas o en South Valley. *Entonces*, ¿tienes una pregunta
be-3s in Barelas or in South Valley *entonces* have-2s a question
o algo que no, no haya dicho?
or something COMP NEG NEG have-1S-SUBJ say-PART
'… but if somebody wants to live happily in Albuquerque, it's in

> Barelas or in the South Valley. *Entonces*, do you have a question or something that I haven't, haven't said?' (Barelas JG)

In order to successfully compare the behaviour of these two discourse markers, it is essential that the same categories be maintained for both. As shown above, this is, for the most part, possible, and allows us to examine distributional patterns for each, shown in Table 2. As Table 2 shows, the distributions of *so* and *entonces* are remarkably similar. Indeed, chi-square tests show that there are no significant differences in the relative frequencies of *so* and *entonces* for any of the shared categories ($p > .01$) (i.e. conclusion, resultative, discourse progression, and indeterminable). It may further be noted that the uses that do not overlap, i.e. final *entonces* and introducing an unexpressed utterance make up only 2% of the data.

Table 2. Distribution by discourse function

	Conclusion		**Resultative**		**Disc. Prog**		**Fin. *entonces***		**Unexp. utt.**		**Indet.**		**Total**	
	N	%	N	%	N	%	N	%	N	%	N	%	N	%
so	55	1	104	36	109	38	0	–	5	2	16	6	289	100
entonces	20	16	31	25	63	51	4	3	0	–	6	5	124	100
Total	75	18	135	33	17	42	4	1	5	1	22	5	413	100

For all categories, $p > .01$.

This is evidence that, as far as their discourse-marker uses go, *so* and *entonces* have not taken on specialized discourse functions. In fact, they appear to be in free variation. Is this evidence, however, that *so* is encroaching upon, and perhaps will eventually replace, the discourse-marker territories of *entonces*, reminiscent of the way English discourse marking has replaced the German system in the speech communities studied by Salmons (1990)? This is not so clear. While evidence does show that *so* is more frequent than *entonces*, as Torres (2002) notes, this is not necessarily support for the view that *so* will eventually replace *entonces*. It could equally show that these two forms find themselves in stable variation. There are two possible pieces of evidence for the argument that discourse marker *so* will replace discourse marker *entonces*. First, *so* occurs 40 percentage points more often than *entonces* (see Table 1). Second, 31% (55/179) of all *entonces* uses are temporal, not discourse-marker uses. These results offer only clues, however. Only diachronic studies, which will

soon be possible as the record of spoken New Mexican Spanish ages, will reveal the fate of *entonces.*

3.2 *So* as a trigger for code-switching

As mentioned in the Introduction, it has been proposed that other-language-origin discourse markers can act as triggers for code-switches (Clyne 1972, 1997, 1999; Pfaff 1982; Brody 1987). In order to test this hypothesis, I counted, first, whether the discourse marker occurred at the border of a switch, and second, the nature of the switch. For this study, I have considered a code-switch to be any multi-word fragment, inter- or intrasentential, preceded by a different code. I have distinguished these from single other-language-origin nouns, or singletons, which may or may not be phonologically integrated into the language that surrounds them. Nonetheless, many recent studies have shown that, phonologically integrated or not, singletons tend to be syntactically integrated and behave grammatically like the recipient language, not the donor language, and thus should be classified as borrowings, not code-switches (Poplack et al. 1988; Sankoff et al. 1990; Poplack and Meechan 1998; Torres Cacoullos and Aaron 2003a). For example, words like *complaint* in the sentence *Y le puse* ***complaint*** *a ese chota* "And I made a complaint against that cop" (NMCOSS 219), would be considered to be "nonce borrowings" (Poplack and Meechan 1998), not code-switches, and as such would, as a group, behave grammatically like Spanish, not English. This is an important distinction, for this means that these words should be treated not as items in the donor language, but rather as borrowings in the recipient language.

Given this distinction, there were eight combinations that appeared in the data, shown in Figure 1.[6]

If we are to consider singletons to be borrowings, as evidence from various studies suggests we should, then the only possibilities listed above that would be considered switches from one grammar to another would be 6, 7, and 8. The others would not constitute a code-switch. Due to the contentious nature of singletons, however, I kept these different combinations separate in the original coding of the data.

In the speech of New Mexico, as can be seen in Table 3, *so* occurs not only in English and in code-switches, but it also occurs 50% of the time in monolingual Spanish. In this sample, which has a dearth of monolingual English, it occurs less, at only 26%, in monolingual English discourse. *Entonces*, on the other hand, only occurs once in monolingual English in this speech community,

1. Spanish → Spanish (e.g. *estaba pequeño todavía, **so** como quien dice aquí me crié*, NMCOSS M2)
2. English → English (e.g. *I went up to second grade, **so** I really don't have much of a memory*, NMCOSS M1)
3. English singleton → Spanish (e.g. *ni uno de ellos tiene los ojos de azules como los tenía el grandpa, **so** es muy…*, NMCOSS 10)
4. Spanish singleton → English (e.g. *the people from the outside are coming into the fiestas. **So** like us, we didn't come…*, NMCOSS M1)
5. Spanish → English singleton (e.g. *luego me quitaron el pie. **So** disability me dieron*, Barelas IN)
6. English → Spanish (e.g. *if our kids get out of line, they can get paddled, **so** cuando a mí me traen ese papel mis hijas…*, NMCOSS M1)
7. Spanish → English (e.g. *y ellos decían que la mandaban. **So** they gave her various tests…*, NMCOSS 190)
8. English singleton → English (e.g. *participan en uhm, en sports, **so** they're always in practice*, NMCOSS M1)

Figure 1. Language-mixing and monolingual combinations that occurred in the data

Table 3. Types of language switches co-occurring with *so* and *entonces*

	so		***entonces***		**Total**	
	N	%	N	%	N	%
Total monolingual	234	81	113	91	347	84
Spanish → Spanish	144	50	104	84	248	60
English → English	74	26	1	1	75	18
Eng. singleton → Spanish	14	5	6	5	20	5
Span. singleton → English	1	< 1	0	–	1	< 1
Spanish → Eng. singleton	1	< 1	2	2	1	< 1
Total switched	45	**16**	3	**2**	48	12
English → Spanish	16	6	2	2	18	4
Spanish → English	25	9	1	1	26	6
Eng. Singleton → English	4	2	0	–	4	1
Indeterminable	10	3	8	6	18	4
Total	289	100	124	100	413	100

$p = .0001$ for difference between words for total mono/switched
Note that even if singletons are considered to be in their language of origin, $p = .0027$.

and occurs in monolingual Spanish 84% of the time. *So* occurs at the border of a code switch 16% of the time. In contrast with *so*, *entonces* occurs at the border of a code-switch only 2% of the time.

These results, which show a significant difference ($p < .01$) in the frequency of code-switches co-occurring with each discourse marker, suggest that *so* may, in fact, function as a trigger for code-switching. Given its high frequency in

monolingual Spanish (50%), it seems likely that *so* is an established borrowing in New Mexican Spanish (cf. Salmons 1990 for a similar argument on this classification), and thus is not considered by speakers in this community to be English. If this is the case, then *so* would constitute a convergence site at which switching would be facilitated (Clyne 1997:107; Clyne and Cassia 1999:68), since *so* would be considered to be both English and Spanish. This is not the case with *entonces*, which is not associated with English (note, again, its nearly complete absence from monolingual English), and thus could not be considered a convergence site.

Here we must consider the possibility that *so* triggers code-switching only in certain environments. Is a certain discourse context more favourable to a switch with *so*? Table 4 shows how each function was used within both monolingual and switched discourse for *so* and *entonces*.[7]

As we can see in Table 4, switches with *so* are not preferred any particular context. While there is 20% conclusion use in monolingual *so* data, there is 14% in switched *so*, and while there is 39% resultative use in monolingual *so*, there is 31% in switched *so*. Similarly, the discourse progression use is at 39% in monolingual *so*, compared to 47% in switched *so*.

Though the numbers do not match exactly, chi-square tests show no significant differences in the use of *so* in monolingual or switched discourse, nor in the difference between these two modes in the data as a whole ($p > .01$ for all categories within *so* and total). What we may comment on, however, is the fact that the only three switches present with *entonces* occurred in discourse progression use. Since this use is by far the most common and broadest, though, and since switches with *entonces* are so rare, this may be more of an artifact of the numbers themselves than an indication of a preference for certain contexts for switching with *entonces*.

3.3 *So* in monolingual discourse

One final question I wish to address here is the use of *so* within monolingual English and monolingual Spanish. Is this form used the same way in both monolingual modes, or do English and Spanish uses of *so* pattern differently? Table 5 shows the use of *so* within each monolingual discourse mode.[8]

Once again, while the numbers do not match exactly, we see no significant differences ($p > .01$) between the use of *so* in monolingual English and its use in monolingual Spanish. Even in the use of *so* to introduce a conclusion, which is used more in English at a difference of 13%, this difference does not turn out to

Table 4. Switches within each discourse function

	so		*entonces*	
	N	%[a]	N	%
Total monolingual	224	100	109	100
Conclusion	46	20	18	17
Resultative	88	39	31	28
Discourse prog.	88	39	58	53
Final *entonces*	0	–	2	2
Unexp. utterance (*so*…)	2	1	0	–
Total switched	41	100	3	100
Conclusion	6	14	0	–
Resultative	15	31	0	–
Discourse prog.	20	47	3	100
Final *entonces*	0	–	0	–
Indeterminable[b]	24	100	12	100
Total	289	70	124	30

For all categories in *so* and total, $p > .01$.

[a] The percentage columns for *so* and *entonces* represent the percentage of data that were used in each function within each switch type, e.g. 20% of the monolingual *so* uses introduce a conclusion.

[b] Tokens were 'indeterminable' if (1) there was unintelligible speech that obscured the meaning of the utterance, or (2) if the meaning of the utterance could not be determined even though it was intelligible, including cut-off utterances and false starts.

Table 5. The use of *so* in monolingual speech

	Conclusion		**Resultative**		**Disc. Prog**		**Unexp. utt.**		**Indet.**		**Total**	
	N	%	N	%	N	%	N	%	N	%	N	%
Spanish	26	16	62	39	63	39	2	1	7	4	160	100
English	20	27	26	35	25	34	0	–	3	4	74	100
Total	46	20	88	38	88	38	2	1	10	4	234	100

$p > .01$ for all categories

be statistically significant at the .01 level ($p = .0178$), though a larger sample may prove this to be a site of divergence. These results suggest that *so* used in both monolingual Spanish and switched discourse patterns along with the *so* in the monolingual English of this speech community, and is not semantically constrained or specialized.

4. Conclusions

This paper has looked at the use of *so* and *entonces* in the Spanish/English bilingual speech of New Mexico. Two hypotheses have been tested: (1) that semantically similar *so* and *entonces* have developed specialized discourse functions in the discourse of New Mexican bilinguals, and (2) that *so* triggers code-switches more often than *entonces*. With regards to the first hypothesis, it was found that both *so* and *entonces* perform the same discourse functions with the same relative frequency, thus showing no evidence of specialization. It was further argued that this is not, however, strong evidence that *so* is replacing *entonces*: they could be in stable variation. The only evidence in favour of the argument that *so* will replace *entonces* is its higher relative frequency, as well as the elevated temporal use of *entonces* in these data, at 31%, compared to a complete absence of this use in monolingual spoken Colombian Spanish (Travis forthcoming). However, while the latter could perhaps indicate that *entonces* is specializing to its non-discourse-marker use, this elevated use may simply be an artifact of the nature of the data. Since many of the New Mexico interviews dealt with how things were "in the old days" (unlike Travis's data, which is spontaneous conversation), the temporal use of *entonces* is to be expected, and thus does not lend support to the argument of discourse marker replacement.

With regards to the second hypothesis, it was found that *so* does indeed occur with code-switches significantly more than *entonces*. Furthermore, it was found that this switching is not preferred in certain contexts, but rather that it follows the same patterns as in monolingual discourse. Lastly, it was found that the use of *so* in monolingual Spanish and monolingual English shows no significant differences: it is used in the same way in both modes.

The evidence provided here shows that the English-origin discourse marker *so*, which is used in the same way in monolingual English, monolingual Spanish, and switched discourse, overlaps in function with the Spanish-origin discourse marker *entonces*. Furthermore, its semantic distribution coincides with that of *entonces*, showing no indication of specialization. The only special role *so* seems to have is its ability to serve as a bridge between two codes, facilitating the switch from one grammar to another. In other words, while *so* may be relatively new to the Spanish language, its functions are not.

Notes

* I would like to thank Catherine Travis, Salvador Pons Bordería, the *Spanish in Context* editors, and two anonymous referees for their suggestions on this paper. Any faults remain my own.

1. Following Poplack and Meechan (1998), I use 'code-mixing' as an umbrella term that encompasses both code-switching, or "the juxtaposition of sentences or sentence fragments, each of which is internally consistent with the morphological and syntactic (and, optionally phonological) rules of its lexifier language" (Poplack 1993: 255), and borrowing, or the syntactic and morphological (and, optionally phonological) integration of an other-language form into the recipient language (Sankoff et al. 1990).

2. A simple example of such a convergence site for Spanish and English would be *hacienda* 'hacienda', which is an established borrowing from Spanish into English, and thus is a point at which the two languages meet and are, barring minimal phonological differences for some speakers, indistinguishable. By 'established borrowing', I refer to other-language-origin words that are widely used within a speech community (including monolingual speakers), and which may appear in regional or other monolingual dictionaries. Examples of such words in English would be *garage* (from French), *coup d'état* (from French), or *hinterland* (from German). Examples of English-origin words that are established borrowings in New Mexican Spanish include *jaiscul* 'high school' and *troca* 'truck'.

3. There are also a few occurrences of English-origin verbs, adjectives, articles, and other parts of speech. These uses are relatively uncommon, however.

4. Though *entonces* may also be used in this way in discourse, it was not found in these data.

5. Note that (7b) can also be read temporally. Whenever a discourse-marker reading was possible, however, I considered it as such, thus taking a maximally narrow definition of "temporal" uses. This is a conservative move on my part, since if we are to argue that *entonces* as a discourse marker is disappearing from this community, we would expect to find a great deal of temporal uses, and few discourse-marker uses. To consider these more ambiguous uses as principally discourse marker assures that any evidence of a rise in temporal use is not cushioned by the inclusion of possible discourse-marker uses.

6. Combinations that are not mentioned in Figure 1, such as Spanish singleton → Spanish or English → Spanish singleton, did not occur in the data.

7. The reader should note that figures in Tables 4 and 5 differ slightly from those in Table 3 due to the fact that some tokens could not be coded for function.

8. Since *entonces* only occurred once in monolingual English speech, (ellos no más hizo una, "you burned me — *entonces*, pay back time", Barelas MK), a comparison of this sort is not relevant.

Abbreviations

1P	first-person plural	IMP	imperfect
1S	first-person singular	IMPER	imperative
2S	second-person singular	INF	infinitive
3P	third-person plural	NEG	negation
3S	third-person singular	PART	past participle
ACC	accusative	POSS	possessive
COMP	complementizer	PREP	preposition
DAT	dative	PRET	preterit
DIM	diminutive	SUBJ	subjunctive

References

Barelas corpus. 2001. Directed by Rena Torres Cacoullos, Department of Spanish and Portuguese, University of New Mexico. Reference codes refer to researchers' initials, as follows. JA: Jessi Elana Aaron; MA: Matt Alba; PC: Patricia Cano; JC: Jens Clegg; JG: Jessica Grime; PG: Philip Guengerich; MK: Margarita Keller; IN: Irena Nezic; LO: Lis Ovando; DW: Damian Wilson.

Bills, Garland D. and Neddy A. Vigil. 1999. "Ashes to Ashes: The Historical Basis for Dialect Variation in New Mexican Spanish". *Romance Philology* 53.43–67.

Bills, Garland D., Eduardo Hernandez Chavez and Alan Hudson. 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.

Brody, Jill. 1987. "Particles Borrowed from Spanish as Discourse Markers in Mayan language". *Anthropological Linguistics* 29.507–521.

Brody, Jill. 1995. "Lending the Unborrowable: Spanish Discourse Markers in Indigenous American Languages". *Spanish in Four Continents: Language Contact and Bilingualism* ed. by Carmen Silva-Corvalán, 132–147. Washington, D.C.: Georgetown University Press.

Clyne, Michael and Paola Cassia. 1999. "Trilingualism, Immigration and Relatedness of Languages". *Review of Applied Linguistics* 123–124.57–77.

Clyne, Michael. 1997. "Some of the things trilinguals do". *International Journal of Bilingualism* 1:2.95–116.

Clyne, Michael. 1972. "Perception of Code-Switching by Bilinguals: An Experiment". *Review of Applied Linguistics* 16.45–48.

Espinosa, Aurelio M. 1914–1915. "Studies in New Mexican Spanish. Part III: The English Elements". *Revue de Dialectologie Romane* 6.241–317.

Espinosa, Aurelio M. 1975. "Speech Mixture in New Mexico: The Influence of the English Language on New Mexican Spanish". *El lenguaje de los chicanos* ed. by Eduardo Hernández-Chávez, Andrew D. Cohen and Anthony F. Beltramo, 99–114. Arlington, Virg.: Center for Applied Linguistics.

Garcés Gómez, María Pilar. 1994. "Funciones y valores de entonces en español". *Estudios para un corpus del español* ed. by Manuel Alvar Ezquerra and Juan Andres Villena Ponsoda, 217–229. Málaga: Universidad de Málaga.

Pfaff, Carol. 1982. "Constraints on Language Mixing: Intrasentential Code-Switching and Borrowing in Spanish/English". *Spanish in the United States* ed. by Jon Amastae and Lucía Elías-Olivares, 264–297. Cambridge: Cambridge University Press.

Pons Bordería, Salvador. 1998. *Conexión y conectores: Estudios de su relación en el registro informal de la lengua.* Valencia: Universidad de Valencia.

Poplack, Shana. 1993. "Variation Theory and Language Contact". *American Dialect Research* ed. by D. Preston, 251–286. Amsterdam: John Benjamins.

Poplack, Shana and Meechan, Marjory. 1998. "Introduction: How Languages Fit Together in Codemixing". *International Journal of Bilingualism* 2:2.127–138.

Poplack, Shana, David Sankoff and C. Miller. 1988. "The Social Correlates and Linguistic Processes of Lexical Borrowing and Assimilation". *Linguistics* 26.47–104.

Rendle-Short, J. 2003. "So what does that show us?: Analysis of the Discourse Marker 'so' in Monologic Talk". *Australian Review of Applied Linguistics* 26:2.46–62.

Salmons, Joe. 1990. "Bilingual Discourse Marking: Code Switching, Borrowing, and Convergence in Some German-American Dialects". *Linguistics* 28.453–480.

Sankoff, David, Shana Poplack and Swathi Vanniarajan. 1990. "The case of the nonce loan in Tamil". *Language Variation and Change* 2:1.71–101.

Schiffrin, Deborah. 1987. *Discourse Markers.* Cambridge: Cambridge University Press.

Solomon, Julie. 1995. "Local and Global Functions of a Borrowed/Native Pair of Discourse Markers in a Yucatec Maya Narrative". *Berkeley Linguistics Society* 21.287–298.

Sweetser, Eve. 1990. *Fom Etymology to Pragmatics: Metaphorical and Cultural Aspects of Semantic Structure.* Cambridge: Cambridge University Press.

Torres, Lourdes. 2002. "Bilingual Discourse Markers in Puerto Rican Spanish". *Language in Society* 31.65–83.

Torres Cacoullos, Rena and Jessi Elana Aaron. 2003a. "Bare English-origin nouns in Spanish: rates, constraints, and discourse functions". *Language Variation and Change* 15:3.289–328.

Torres Cacoullos, Rena and Jessi Elana Aaron. 2003b. "Determiner Variation with English-Origin Nouns in New Mexican Spanish: Borrowing Bare Forms". *NWAV 31*. Pennsylvania Working Papers in Linguistics.

Travis, Catherine E. Forthcoming. *Discourse markers in Colombian Spanish: A study in polysemy*. Cognitive Linguistics Research Series. Berlin: Mouton de Gruyter.