languages

MDPI

Article

On the Functional Convergence of Pragmatic Markers in Arizona Spanish

Brandon Joseph Martínez

Department of Spanish & Portuguese, University of Arizona, Tucson, AZ 85721, USA; brandonjmartinez@arizona.edu

Abstract: Tags, compared to other types of pragmatic markers (PMs), are typically considered as separate yet related phenomena and are usually differentiated by their syntactic positions and discourse functions, among other factors. The current work explores this differentiation utilizing 36 sociolinguistic interviews with Spanish-English bilinguals in southern Arizona, USA. Standard language variation and change (LVC) methodologies were used in the extraction, coding, and statistical analyses of this dataset (n = 591), with four PM variants identified for study through an exploratory methodology: the tags *no* and *qué no* and the discourse markers (DMs) *you know* and *saber*. The results of our analyses indicate that, while utterance position, self-reported gender, and length of residence were all significant in the multivariate analysis, discourse function was dropped from the statistical model. Therefore, we interpret this finding as an indication that functional differences between these two pragmatic resources have been levelled through grammaticalization, demonstrating that for Arizona Spanish, tags and DMs belong in the same functional category of PMs. Furthermore, an analysis of codeswitching behavior triggered by the incoming variant *you know* demonstrates that it is becoming incorporated into the Spanish pragmatic system, patterning similarly to its counterpart *saber* in terms of function and position, without attrition of the native variant.

Keywords: pragmatics; discourse markers; tag questions; US Spanish; codeswitching; grammaticalization; language variation and change (LVC); language contact

check for updates

Citation: Martínez, Brandon Joseph. 2024. On the Functional Convergence of Pragmatic Markers in Arizona Spanish. *Languages* 9: 148. https://doi.org/10.3390/languages9040148

Academic Editors: Katherine O'Donnell Christoffersen, Ryan M. Bessett and Ana Maria Carvalho

Received: 18 December 2023
Revised: 8 April 2024
Accepted: 12 April 2024
Published: 17 April 2024

Copyright: © 2024 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Pragmatic markers (PMs) are common in spoken language, with functions and frequency of use determined at the community-level (Childs 2021). Past studies have established several key features of PMs: (i) they are the result of grammaticalization, where PMs lose referential meaning, and take on discursive functions (Martín Zorraquino and Portolés Lázaro 1999); (ii) there is a potential variability of form, a side-effect of grammaticalization (Martín Zorraquino and Portolés Lázaro 1999; Silva-Corvalán and Enrique-Arias 2017); (iii) they have the ability to diffuse throughout the syntactic structure, previously described as optionality to the syntax (Clayman and Raymond 2021; Kern 2019; Pichler 2013; Portolés Lázaro 1998; Silva-Corvalán and Enrique-Arias 2017); and (iv) they are multifunctional, with one overall dominant function in any given context (Allerton 2009; Gómez González 2012; Kern 2019; Palacios Martínez 2014; Travis 2005). Furthermore, previous investigation of PMs in bilingual varieties (Aaron 2004; Andersen 2022; Carvalho and Kern 2019; Palacios Martínez 2014; Pichler 2021; Said-Mohand 2007; Sankoff et al. 1997) show that intense contact between languages can potentially cause changes within the pragmatic system of the recipient language, sometimes accelerating change and its diffusion (Pichler 2021), and at other times diversifying the system without the attrition of native forms (Aaron 2004).

The current work analyzes PM usage among Spanish-English bilinguals in Tucson, Arizona (AZ), through a variationist lens (Tagliamonte 2012). Specifically, we review the distribution of tags *no* and *qué no* (a basic PM subtype), compared to the discourse markers (DMs) *you know* and *saber* (a separate PM subtype, following Travis 2005). Examples of each variant are seen in examples (1)–(4):

(1) *Y,* ***no****, ahora están muy mal los chamacos.* (CESA031) 'And, yeah, now the kids are really bad'.[1]
(2) *Sigue el Spring Fling hasta el lunes,* ***¿qué no****?* (CESA073) 'The Spring Fling is going until Monday, right?'
(3) ***You know****, eso era lo que me gustaba a mí*. (CESA013) 'You know, that's what I liked'.
(4) *Y pues,* ***sabes qué****, estuvo mal de mi parte haberme enojado por eso.* (CESA021) 'And well, you know, it was bad on my part to have gotten mad for that'.

Utilizing 36 sociolinguistic interviews from the *Corpus de Español en el Sur de Arizona* (CESA; Carvalho 2012-), this analysis explores the effects of four independent variables in conditioning the realization of these four variants, two linguistic (function and position), and two extralinguistic (gender and length of residence). This study provides a comprehensive analysis proposing the possible convergence of tags with DMs, alongside compelling evidence of their grammaticalization path. Furthermore, the study demonstrates the permeability of non-native PMs, such as *you know* in US Spanish, adding valuable insight to the contact literature. Our main research questions were as follows:

(i) What discourse functions (shared or otherwise) do different subtypes of PMs, including tags and DMs, fulfill in Arizona Spanish?
(ii) Following Pichler (2013), among others (Kluge 2011; Palacios Martínez 2014; Schleef and Mackay 2022), are certain discourse functions correlated with certain syntactic positions in Arizona Spanish?
(iii) What role, if any, does the incoming English loan *you know* have in conditioning codeswitching behavior in Arizona Spanish?

2. Pragmatic Markers

To begin, DMs and tags have oftentimes been analyzed as two separate yet related phenomena in the discourse/pragmatic literature. That is, while they are both the result of grammaticalization processes that, over time, reduce their referentiality and cause them to take on pragmatic features (Gómez González 2012, 2014; Martín Zorraquino and Portolés Lázaro 1999), tags are usually analyzed as a separate subtype of PM due to their originally addressee-oriented functions, as well as what was thought to be their fixed clause-final position in both Spanish and English (Gómez González 2012, 2014). However, several authors have analyzed the grammaticalization cline of tags, which allows them to take on other discourse functions, as well as migrate to other syntactic positions within the clause (Palacios Martínez 2014; Pichler 2013; among others), making them more DM-like over time.

2.1. Pragmatic Markers: A Synopsis

Due to a plethora of terms being used in the literature, such as discourse markers (Said-Mohand 2007), *muletillas* 'pet phrases' (Molina 2005), or fillers (Clayman and Raymond 2021), it is hard to precisely define what the term 'PM' refers to. Due to the pragmatic need which creates them, PMs do not form a uniform natural class, but rather a loose pragmatic category (Martín Zorraquino and Portolés Lázaro 1999; Travis 2005). They are drawn from a range of different lexical categories (Bybee 2015; Croft 2022), becoming *pequeñas palabras* 'little words' (Aschenberg and Loureda Lamas 2011, p. 12), which have abstracted away from their original referential meanings to have scope over discourse (Aschenberg and Loureda Lamas 2011; Degand et al. 2022; Silva-Corvalán and Enrique-Arias 2017). PMs contribute little to nothing to an utterance's propositional meaning (Aaron 2004; Andersen et al. 2017; Gómez González 2014); however, they help to shape syntactic and pragmatic structures (Aschenberg and Loureda Lamas 2011; Bybee 2015). Furthermore, Portolés Lázaro (1998) provides a rather succinct description of PMs as "…invariable linguistic units which do not exercise a syntactic function in the predicational framework of the sentence (that is, they are outside of the sentence's syntax) and which possess a purpose in discourse: guiding inferences made in communication. In other words, they signal relations which exist between units of discourse" (pp. 25–26).[2]

As in many instances of variation driving change (Bybee 2015), PM development can be attributed to grammaticalization, small changes occurring in repeated usage over time (Bybee 2011). Through increased frequency, what were once referential items can gain structural autonomy, demonstrating greater formal and syntactic flexibility (Bybee 2011; Childs 2021; Gómez González 2014; Kern 2019; Kluge 2011; Martín Zorraquino and Portolés Lázaro 1999; Pichler 2013, 2021; Travis 2005). Furthermore, PMs along the cline of grammaticalization (Traugott and Dasher 2001) exhibit functional change, gaining functions non-conducive to a response (Carvalho and Kern 2019; Gómez González 2014; Palacios Martínez 2014; Pichler 2013, 2021). That said, PMs tend to be generally multifunctional, but with one dominant function overall (Aschenberg and Loureda Lamas 2011; Crible and Degand 2019; Croft 2022; Degand et al. 2022; Denis and Tagliamonte 2016; Kluge 2011; Schleef and Mackay 2022; Torres and Potowski 2008; Travis 2005).

Following Bybee (2011), grammaticalization here is conceived of as small changes in frequency and usage over time. However, it is not a single, unitary process, but rather a number of related changes typically moving in the same direction to accomplish similar ends; that is, to transform lexical items and constructions into grammatical elements. There are, however, certain universal features of grammaticalization, including semantic bleaching, decategorialization, pragmatic extension, phonetic reduction, autonomous divergence, subjectification, and generalization (Bybee 2011; D'Arcy 2005). Furthermore, it is typically conceptualized as gradual and unidirectional, with variation in usage at the individual and community levels (Bybee 2011, 2015; Martín Zorraquino and Portolés Lázaro 1999; Silva-Corvalán and Enrique-Arias 2017; Tagliamonte 2012).

However, not all researchers agree that PMs arise through grammaticalization, for various reasons, including the need to (i) demonstrate that change is occurring diachronically, and (ii) show that items are being recategorized into new grammatical classes. To the first point, although we do not at the time of writing have longitudinal data, we believe that the results from other dialects of Spanish (Aschenberg and Loureda Lamas 2011; Kern 2019; Martín Zorraquino and Portolés Lázaro 1999; Palacios Martínez 2014; Said-Mohand 2007) and English (D'Arcy 2005; Pichler 2013, 2021; Schiffrin 1987) are sufficiently indicative of the fact that certain diachronic changes are occurring for both tags and DMs, as analyzed in this study. Following the above literature, it does seem to point to the fact that they are undergoing grammaticalization. To the second, a key feature of grammaticalization posits that the grammaticalized item will take on a new class membership and thus new grammatical functions, due to decategorialization. In that vein, D'Arcy (2005), in her study of *like,* convincingly argues that taking on pragmatic functions *is* sufficient grounds for considering PM development as an example of class reanalysis in itself (p. 58). Furthermore, the fact that PMs can be grouped together loosely as a pragmatic class (Martín Zorraquino and Portolés Lázaro 1999) would indicate that as PMs decategorize, they are potentially reanalyzed into a common grammatical category.

Finally, we hedge this discussion by admitting that there does remain the possibility that PM development *cannot* in fact be accounted for underneath even an extended definition of grammaticalization. Turning to the literature, Kluge (2011), for example, in their analysis of English *you know*, French *tu sais*, and Spanish *tú sabes*, offers us the possibility that this change from verbs-of-knowing to PMs may best be analyzed as pragmaticalization, as these PMs retain a remnant of their original semantics, and only gain a discourse/interactional role rather than a new grammatical category. Furthermore, in the case of Spanish *claro*, Ocampo (2006) argues that while there are similarities to grammaticalization, DM development can better be explained as discursivization, in that it is a move from the lexicon/grammar → discourse, while grammaticalization is characterized in their model as only movement from the lexicon → grammar. Finally, Denis and Tagliamonte (2016), in their discussion of the displacement of *you know* by *right* in Canadian English, posit that this is an instance of lexical replacement rather than grammaticalization, because *right* entered the system already fulfilling the same functional roles as *you know*, and is only replacing it in terms of frequency-of-use rather than restructuring the entire

pragmatic system. In any case, based on previous proposals and the current data analysis, we follow Onodera's (2011) suggestion that, even given counterexamples which may point to pragmaticalization or discursivization rather than grammaticalization, "the similarities between the development of pragmatic markers and grammaticalization prevail over the differences" (p. 618).[3]

2.2. Tag "Questions": A Misnomer?

To quote Gómez González (2012), we move now to "the question of tag questions" (p. 59). While tags themselves are considered a subset of PMs, the term tag question (TQ) is misleading, creating bias through inferring that all tags serve a response-seeking function. While that may be the case sometimes, where tags are used to elicit confirmation or information (Allerton 2009; Croft 2022; García Vizcaíno 2005; Gazdik 2022; Gómez González 2014), tags themselves do not always need or want an answer (Allerton 2009; Gómez González 2014; Kimps et al. 2014a, 2014b; Pichler 2013). Thus, we concur with Kimps et al. (2014a) that, rather than TQs, we should refer to this PM subset simply as "tags".

Not unlike other PMs, tags are referred to as "appendages" (Allerton 2009, p. 309) with "little to no propositional meaning" (Gómez González 2014, p. 95), which are attached to an anchor clause (Allerton 2009; Childs 2021; García Vizcaíno 2005; Gómez González 2012, 2014; Palacios Martínez 2014; Pichler 2013, 2021). In English, due to grammaticalization, they include full-form variant tags such as *isn't it*, which can undergo various stages of phonetic reduction to become invariant coalesced forms such as *innit*, thus gaining greater structural flexibility in the process (Allerton 2009; Childs 2021; Denis and Tagliamonte 2016; Gómez González 2012, 2014; Palacios Martínez 2014; Pichler 2013, 2021). In languages such as Spanish, however, only invariant forms are permitted, i.e., *no*, *verdad* (Gómez González 2012, 2014), with some possibility for variation in terms of phonetic form. Reflecting on the variable nature of tag usage cross-linguistically, Kimps et al. (2014a) demonstrate that tags are more common in British than in American English, and Gómez González (2012, 2014) shows that they are used less frequently in English than in Spanish.

As mentioned previously, referring to tags or any kind of PM as "fillers" does them a disservice (Allerton 2009; Clayman and Raymond 2021). In their grammaticalization cline from lexical items to PMs, tags become multifunctional (Gómez González 2012; Pichler 2021), gaining a range of functions, from managing turn-taking (Allerton 2009; Gómez González 2012, 2014) to serving functions that are conducive or not to a response (Pichler 2013). In terms of position, while the utterance-final position is considered canonical for tags in English and Spanish (Allerton 2009; Gómez González 2012, 2014), both utterance-initial and -medial positions are also possible (Croft 2022; Gómez González 2012, 2014; Kimps et al. 2014a, 2014b; Palacios Martínez 2014). This is the opposite of DMs, which are claimed by some scholars as being preferred in an utterance-initial position (D'Arcy 2005; Onodera 2011), but which can also occur in other clause positions (Martín Zorraquino and Portolés Lázaro 1999). Importantly, it has been noted that different positions are correlated with certain discourse functions (Kluge 2011; Palacios Martínez 2014; Pichler 2013; Schleef and Mackay 2022).

2.3. Pragmatic Systems in Contact

Turning now to our community of study, Tucson, AZ, is a medium-sized city located within the US-Mexico Borderlands region. Therefore, because our participants are necessarily Spanish-English bilinguals to varying degrees, we must consider how language contact can shape and impact pragmatic systems. Previous work demonstrates that lexical items are easily borrowed in situations of language contact, whether local or remote. Moving beyond simple borrowing, Andersen et al. (2017) detail the role that pragmatics plays in contact situations:

> "…contact is seen not as an external factor that triggers change, but as one that is internal to the processing and use of language itself in the multilingual

> speaker's repertoire of linguistic structures. Accordingly, speakers are seen as creative communicators who draw on their entire repertoire in order to make communication more efficient, and the functional value of linguistic categories is a factor that plays a role in the speaker's ability to select structures within their repertoire" (p. 74).

Due to their interconnected grammars, the bilingual speaker draws from a greater number of linguistic features than their monolingual counterparts (Otheguy et al. 2015), being able to deploy a number of different pragmatic resources at any given time. As borrowings become more frequent, they may lead to systematic changes, through filling in functional gaps, acquiring new functions, or bringing functions over from the donor language (Andersen et al. 2017; Bybee 2015).

In exploring the effects which contact has in the development of pragmatic systems, we turn first to US Spanish in contact with English. For example, Aaron (2004) and Torres and Potowski (2008) both analyzed bilingualism's effects for *so* and *entonces* in this variety. Aaron (2004), in her analysis of Spanish-English bilinguals in New Mexico, found no significant difference in usage between either variant (although *so* is more likely to trigger a codeswitch), possibly due to semantic similarity (p. 163). Rather than displacing *entonces*, however, *so* co-exists with the native variant in a state of stable variation, sharing functional distribution. In Chicago, Torres and Potowski (2008) similarly find that these same variants also share functions, although *so* is used more frequently and is correlated with language proficiency. While in both cases *so* and *entonces* co-exist within their respective systems, due to a reduction in Spanish proficiency through subsequent generations, it is possible that *so* will replace *entonces* entirely (Aaron 2004; Torres and Potowski 2008). Turning to Arizona, Kern's (2019) analysis of *like*, *como*, and *como que* in English-Spanish bilinguals showed that they were used in similar syntactic positions and for similar functions. However, there was little to no convergence between these languages, as no Spanish PMs were used in English, and *like* was only minimally used in Spanish. Finally, looking at *tú sabes* vs *you know* in Florida, Said-Mohand (2007) finds that they share similar functions, with use of each variant correlated with language proficiency.

Pragmatic contact has been analyzed across other language pairs as well. First, in their analysis of Spanish in contact with Portuguese, Carvalho and Kern (2019) evaluate the permeability of tags such as *no*, *no es*, and *né* in a bilingual border community in Uruguay, demonstrating that while there is evidence for a small degree of borrowing between languages, their results do not indicate large-scale convergence. Second, in evaluating how language contact affects the grammaticalization of *innit* in British English, Pichler (2021) hypothesizes that L2 speakers are responsible for the spread of this tag due to the speakers' origin in a number of ethnically-diverse communities, reflective of a range of L1 languages, positing that these speakers are accelerating a change already underway—the grammaticalization of *isn't it*—by taking qualities of invariant tags in the speakers' L1 and transmitting them to English. Third, Sankoff et al. (1997), in their classic study of English-French bilinguals in Canada, find that the use of PMs was affected by the degree of integration into the Francophone community and correlated with L2 French proficiency. Finally, in an example of remote contact, Andersen (2022) demonstrates the incorporation of English PMs such as *please* and *sorry* in monolingual Norwegian reality shows, showing that these loans are incorporated into the speakers' pragmatic systems alongside their native counterparts without displacement, and with only a difference in illocutionary strength.

With the above examples from several different language pairs in mind, we may now ask whether pragmatic borrowings are simply diversifying the bilingual system in the short-term, or whether they can lead to the restructuring of the entire system in the long-term. Although these studies demonstrate that pragmatic borrowing does occur in both local (Aaron 2004; Carvalho and Kern 2019; Kern 2019; Pichler 2021; Said-Mohand 2007; Sankoff et al. 1997; Torres and Potowski 2008) and remote (Andersen 2022) contact situations, the trend seems to be the incorporation of borrowed variants into a recipient language without the attrition of native PM variants. Instead of drastic systemic restructuring, these

studies indicate stable synchronic variation with the possibility for future change. Turning now to the study at hand, it may be fortuitous to hypothesize that our Arizona speakers will display interlingual effects, representative of convergence, shifting, or acceleration of an already-existing change. Conversely, we may also hypothesize that there will be no evidence of convergence between the bilinguals' two languages, even given the presence of the English loan *you know*.

3. Methodology

3.1. The Data: Corpus del Español en el Sur de Arizona

This study is based on 36 interviews taken from CESA (Carvalho 2012-), a corpus of 76+ sociolinguistic interviews with Spanish-English bilinguals from Tucson. For our dataset, we initially focused on those born-and-raised in Tucson before expanding our selection criteria to all those who utilized *you know* at least once in discourse, to explore the extent of this variant's diffusion within the community. Our final sample included 24 female and 12 male speakers, the vast majority of whom are college-aged millennials. For that reason, neither age, education-level, nor social class were taken into account in this analysis. Following an exploratory analysis of the data, four possible variants were identified for study, grouped into two categories based on their lexical origin as negative adverbs (*no* and *qué no*) or verbs-of-knowing (*saber* and *you know*).

For the first group, that is, tags, *no*, as used in Spanish, is well-documented (Carvalho and Kern 2019; Degand et al. 2022; García Vizcaíno 2005; Gazdik 2022; Gómez González 2012, 2014; Uclés Ramada 2020), usually analyzed as canonically invariant and seeking an interlocutor's response (Gómez González 2012, 2014). It is by far the most well-studied Spanish PM analyzed here, with negative-polarity tags being more common than positive ones (Degand et al. 2022; Gómez González 2014), although a majority of past studies analyzed Peninsular rather than Latin American Spanish. To our knowledge, this is the first study to include the tag *qué no*, which, from our experience, is very frequent in Arizona, New Mexico, and Sonora.

For the second group, DMs, both variants are well-documented. Kluge (2011) claims that *saber* is already on its way to becoming a full-fledged PM, if it is not one already, and it has in the past been investigated in both Old and New World dialects (Bybee 2015; Kluge 2011; Molina 2005; Said-Mohand 2007; Silva-Corvalán and Enrique-Arias 2017). Given its retention of an element of its referential meaning (Schiffrin 1987) and its uncertain location along the cline of grammaticalization, it retains some formal qualities of referential verbs; that is, an ability to be inflected for person and number, with or without a pronoun or adverb (subvariants include *sabes*, *sabe*, *tú sabes*, *ya sabes*, etc.). For the English-origin *you know*, it has also been investigated as a PM/tag, which is clearly in the cline of grammaticalization (Bybee 2015; Clayman and Raymond 2021; Degand et al. 2022; Denis and Tagliamonte 2016; Kluge 2011; Sankoff et al. 1997; Schiffrin 1987; Schleef and Mackay 2022). To our knowledge, this is the first study systematically looking at use of *you know* alongside *saber* in the Spanish of the US Southwest.

In variationist study, it is important to circumscribe the envelope of variation, every possible place where a variable varies within the grammar (Tagliamonte 2006, 2012). However, due to the nature of discourse/pragmatic variables, this can be difficult. Per Denis and Tagliamonte (2016), in circumscribing this variable context, we must look for derivational, functional, and semantic equivalence between the variants (that is, weak complementarity; Tagliamonte 2012), with a reliance on context, to establish pragmatic meaning and to differentiate referential from pragmatic uses of the variable. In doing so, a large number of possible tokens had to be excluded for the following reasons: (i) the PM occurred as a backchannel, within quoted speech, or was a false start; (ii) the token seemed to be lexical rather than pragmatic; (iii) for *no*, it was used in a *sí-no*/*no-sí* or *ay no* construction, was clearly a negation, or was proximally-located to another negative-polarity word, which made its discursive meaning ambiguous; (iv) for *qué no*, it was not confused with a complement phrase (CP) boundary, that is, *que no* (without the orthographic accent);

and (v) for *you know*, it was not a part of the larger chunk *you know what I mean?*, which carries similar pragmatic functions and is likely the syntagmatic origin of this PM (Onodera 2011), but was excluded due to lack of structural similarity to the *saber* tokens.

Upon identifying participants and defining the envelope of variation, each interview was methodically reviewed, utilizing transcripts and audio recordings to identify tokens. Once found, utilizing acoustic cues to locate its position in relation to the main clause, tokens were extracted in context, tabulated in a separate spreadsheet, and coded for discourse function, utterance position, self-reported gender, and length of residence.[4] When completed, our dataset consisted of n = 591 coded tokens available for analysis.

3.2. Coding

As mentioned previously, four predictors were coded for and utilized in the statistical analysis, including discourse function, utterance position, self-reported gender, and length of residence, which will be outlined in the following paragraphs.

First, there are a number of problems which come with determining discourse/pragmatic function, the foremost being that nearly every study utilizes a different set/subset of functions, given the fact that function is idiosyncratic to the community in question (Childs 2021; Degand et al. 2022; Denis and Tagliamonte 2016; Kluge 2011; Molina 2005; Pichler 2013; Said-Mohand 2007; Sankoff et al. 1997; Torres and Potowski 2008). Even if we utilize the same coding system, we then run into the issue that coding for function is subjective and dependent on the coder and context (Denis and Tagliamonte 2016). Finally, due to their multifunctionality (Allerton 2009; Gómez González 2012; Kern 2019; Palacios Martínez 2014; Travis 2005), we must rely on the principle of "weak complementarity" (Tagliamonte 2012) to tease out a dominant function appropriate for a given context and to ensure comparability across variants. With that in mind, we adapted the functional typologies proposed for the study of tags by both Gómez González (2012, 2014) and Pichler (2013), extending said typologies to other types of PMs for the first time, with coding established at three levels. First, all PMs were coded for broad and narrow discourse functions (Gómez González 2012, 2014), with "broad" referring to one of three general functional categories, addressee-, speaker-, or exchange-oriented, which can be further separated into seven "narrow" categories. Next, each PM function was coded for whether it was conducive to a response (Pichler 2013), with addressee-oriented functions considered as conducive, and speaker- and exchange-oriented functions as non-conducive (Carvalho and Kern 2019). Example tokens for each of the seven narrow discourse functions, as well as a short descriptor of each category, can be seen in examples (5) to (11) below:

Information/confirmation-seeking: seeks a (usually affirmative) response.

(5) *Sigue el Spring Fling hasta el lunes, ¿**qué no**?* (CESA073)
'The Spring Fling is going until Monday, right?'

Action-seeking: seeks an action from an interlocutor, real or imagined, through commands/offers.

(6) *Todos juntos vamos a cambiar todo, ¿**no**?* (CESA052=
'We're going to change everything together, aren't we?'

Attitudinal/stance-taking: provides a subjective evaluation or position in relation to the topic.

(7) ***You know**, eso era lo que me gustaba a mí.* (CESA013)
'You know, that's what I liked'.

Challenging/confrontation: lends illocutionary force to a negative speech act on the part of the speaker.

(8) *Pero, **you know**, si no te adaptas, pues...* (CESA037)
'But, you know, if you don't adapt, well...'

Focusing: highlights or re-emphasizes a specific element within the discourse.

(9) *Cuando XY se estaba criando, **tú sabes**, XY tiene autism...* (CESA016)

'When XY was growing up, you know, XY has autism. . .'

Phatic/alignment: ensures that an interlocutor is following the stream of thought (differentiated from information/confirmation-seeking above by not needing a response).

(10) *Dices que es más informal, ¿**no**?, más, bueno sí. . .* (CESA052)
You say it's more informal, right, more, well yeah. . .

Discursive: brackets or structures speech itself.

(11) *Jamás los vas a conocer, al saber que, **you know**, que hay un niño.* (CESA037)
'You're never gonna know them, to know that, you know, that there's a kid'.

A visual breakdown of the relationship between conduciveness, broad, and narrow discourse functions can be seen in Figure 1 below:

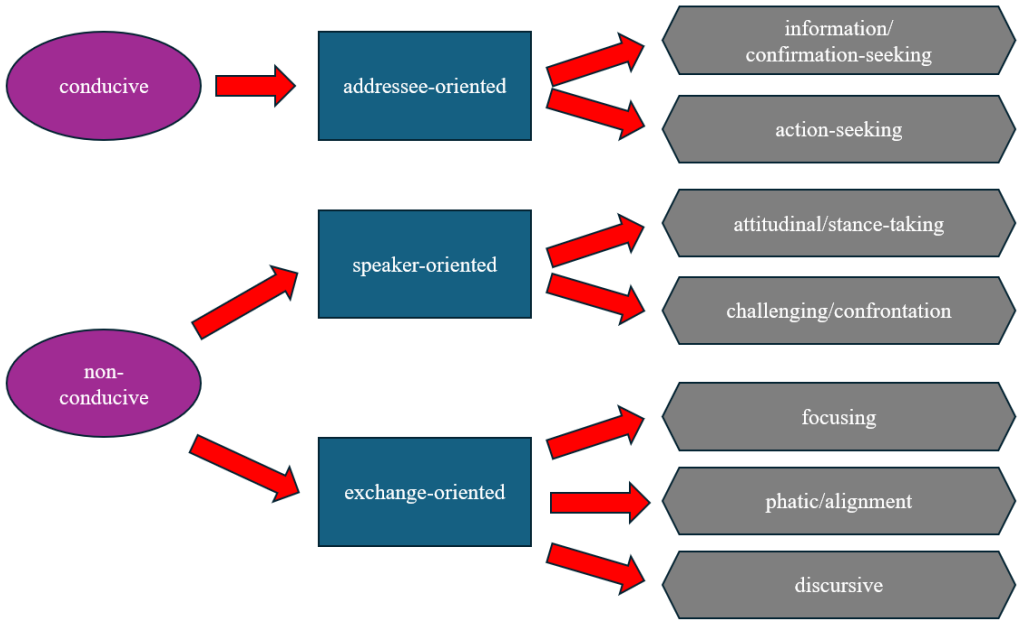

Figure 1. Discourse-function breakdown by conduciveness (purple), broad discourse function (blue), and narrow discourse function (gray). Adapted from Gómez González (2012, 2014) and Pichler (2013).

Due to collinearity between conduciveness and broad discourse function, each was submitted to separate statistical analyses to determine which provides a better fit for interpreting the results.

Moving now to our second predictor, PMs have been traditionally considered as optional to the syntax (Aaron 2004; Aschenberg and Loureda Lamas 2011; Bybee 2015; Degand et al. 2022; Kern 2019; Onodera 2011; Palacios Martínez 2014; Portolés Lázaro 1998; Sankoff et al. 1997; Silva-Corvalán and Enrique-Arias 2017; Travis 2005) and able to occur in any turn/utterance position (Clayman and Raymond 2021; Denis and Tagliamonte 2016; Martín Zorraquino and Portolés Lázaro 1999; Sankoff et al. 1997; Schiffrin 1987; Schleef and Mackay 2022), whether at the edge of clauses or within them (D'Arcy 2005; Kern 2019). Furthermore, different positions are correlated with certain functions (Kluge 2011; Palacios Martínez 2014; Pichler 2013; Schleef and Mackay 2022). Here, PMs were coded for their position in an intonation unit/utterance, whether initial, medial, or final, utilizing impressionistic cues to assist in the coding process.

Third, the effects of gender are mixed in the PM literature, demonstrating either men or women leading changes or utilizing full-formed vs reduced/coalesced variants (Childs 2021; Sankoff et al. 1997; Tagliamonte 2012) or, conversely, showing no effects (Andersen 2022; Kern 2019; Kimps et al. 2014b; Said-Mohand 2007). Although our participant sample is unevenly split at a ratio of 1:2 male/female, and therefore should be taken with a grain of salt, we believe that the results of the analyses should provide indication as to specific usage patterns by both genders for these bilinguals.

Finally, due to a similarity in age for most of our participants (year of birth 1980–1995/millennials), the factor utilized as a proxy was the participants' length of residence in Tucson. This may be indicative of language accommodation, in which speakers over time approximate a local standard for a given variable (Aaron and Hernández 2007; Esparza 2017). Furthermore, because Sankoff et al. (1997) and Torres and Potowski (2008) found PM use to be dependent on the degree of community integration, it is possible that the same might be seen here. We coded the data for three categories: "Since birth", "1–19 years", and "20+ years"; later, the "since birth" and "20+ years" categories were collapsed due to a similarity in patterning for the variables.

3.3. Statistical Analysis

Following standard variationist procedure (Tagliamonte 2006, 2012), after extraction and coding were completed, all data were tabulated in distributional analyses by predictor, to get a raw frequency distribution. Then, utilizing the statistical software Rbrul (Johnson 2009, 2023), two separate multivariate analyses were run with DM as the input value, reflecting the lexical origin of the tags *no* and *qué no* as negative adverbs, compared to the origin of the DMs *you know* and *saber* as verbs-of-knowing. The two runs contained either broad discourse function or conduciveness as a predictor, to avoid the aforementioned collinearity between them.

4. Results and Discussion

From our dataset of n = 591 tokens of PMs in bilingual Arizona Spanish, we found that the tags *no* and *qué no* are approximately twice as frequent as the DMs *you know* and *saber*. The specific breakdowns by variant are shown in Table 1 and Figure 2.

Table 1. Raw distribution of pragmatic markers in Arizona Spanish.

	Tags		**DMs**		
	no	***qué no***	***you know***	***saber***	**Total**
Tokens	370	13	147	61	591
%	63%	2%	25%	10%	

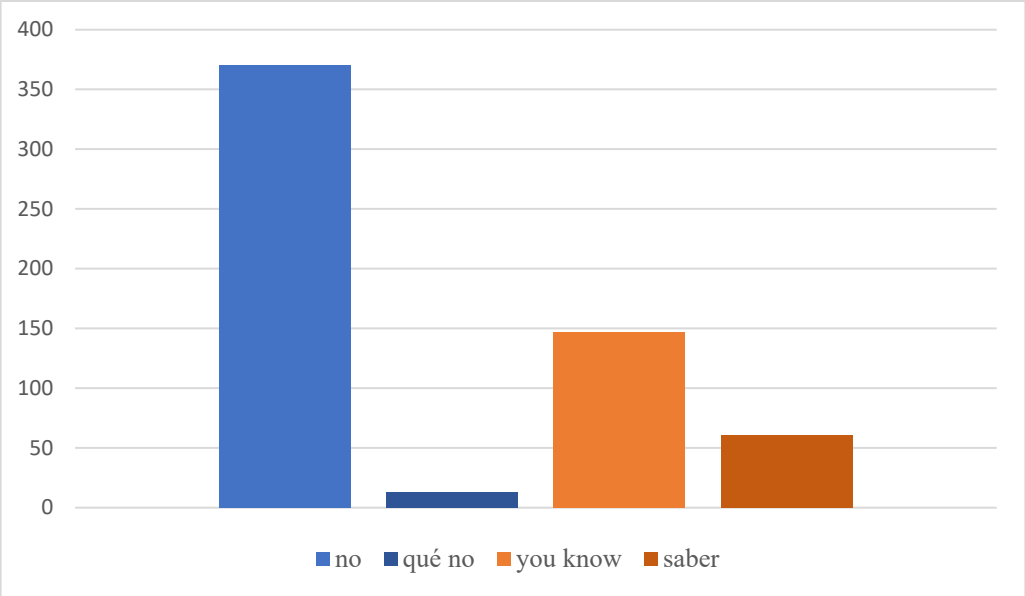

Figure 2. Raw distribution of pragmatic markers in Arizona Spanish.

Individual usage rates are important in looking at the extent of diffusion of a variable; however, these numbers alone fail to give us a full idea of the distribution of the variants

throughout the community. While *no* may be considered as being the *de facto* PM for Arizona Spanish analyzed here, used by 94% of all participants, the other variants are not as widely spread, with *saber* being used by 50%, *you know* by 42%, and *qué no* by 14% of speakers. A specific breakdown by variant can be seen in Table 2 below:

Table 2. Participant usage of pragmatic markers in Arizona Spanish.

	Tags		DMs		
	no	***qué no***	***you know***	***saber***	**Total**
Participants	34	5	15	18	36
% of Total	*94%*	*14%*	*42%*	*50%*	

Reflecting the fact that use of discourse/pragmatic variables is considered as highly individualized behavior, we take these numbers as an indication that while *no* is already well-established as a PM in most varieties of Spanish (Carvalho and Kern 2019; Degand et al. 2022; García Vizcaíno 2005; Gazdik 2022; Gómez González 2012, 2014; Uclés Ramada 2020), the DMs *you know* and *saber* are representative of only an intermediate degree of diffusion, as they are used by half of all speakers. As for *qué no*, due to its heretofore understudied status in Borderlands Spanish, further exploration of its usage in Arizona, New Mexico, and Sonora is required to establish how widespread its use is.

Following the distributional analyses, we submitted the dataset to two separate multivariate analyses. The difference between the runs was the use of broad discourse function or conduciveness as one of the factor groups. However, these were dropped from the model in both cases, indicating that, when all else is the same, function is not significant in the selection of one type of PM over the other for Arizona Spanish. That said, utterance position ($p = 8.56 \times 10^{-23}$), gender ($p = 4.15 \times 10^{-5}$), and length of residence ($p = 0.00345$) were all found significant in conditioning PM selection, with the results for these factor groups being identical for both runs. The results can be seen in Figure 3.

Input value: "DMs"
N=591
R^2 – 0.387

Factor Group	*N*	*%*	*Factor Weight*
Position (*Range – 0.482*)			
Medial	294	0.568	0.788
Initial	101	0.188	0.379
Final	196	0.112	0.306
Gender (*Range – 0.264*)			
Female	377	0.469	0.632
Male	214	0.145	0.368
Length of Residence (*Range – 0.224*)			
Birth/20+ yrs	433	0.434	0.612
<20 yrs	158	0.127	0.388

Figure 3. Multivariate results for DMs (*you know* and saber) vs. tags (*no* and *qué no*) in Arizona Spanish.

An analysis of the factor weights demonstrates that an utterance-medial position (0.788), self-identified female gender (0.632), and those who had been born/had lived in Tucson for 20+ years (0.612) favored the choice of a DM over a tag. This is interpreted here as an indication that these speakers rely on these sociolinguistic cues (utterance position,

self-reported gender, length of residence) in their selection of PM subtype. However, function was dropped for both models, indicating its insignificance as a predictor of PM choice in Arizona Spanish. The following paragraphs detail further distributional and other relevant information for each of the significant factor groups.

As seen in the multivariate analyses, utterance position was the only significant linguistic predictor conditioning speaker choice of tags over DMs. Specifically, factor weight (0.788) shows that an utterance-medial position favors the choice of a DM rather than a tag, while an initial (0.379) and final (0.306) position disfavors it. Furthermore, this would indicate that these latter positions are more favorable for the choice of a tag, a result verified by the distributional results presented in Table 3.

Table 3. Distribution of tags (*no* and *qué no*) vs DMs (*you know* and *saber*) by utterance position (note: the percentages in this and all following tables refer to the total n of tokens in that column).

	Tags	**DMs**	**Total**
Initial	82	19	**101**
	21%	*9%*	***17%***
Medial	127	167	**294**
	33%	*80%*	***50%***
Final	174	22	**196**
	45%	*11%*	***33%***
Total	**383**	**208**	**591**

PMs in general have the ability to occur in any syntactic position due to their supposed optionality to the syntax. However, each particular PM will typically have a positional preference (Martín Zorraquino and Portolés Lázaro 1999). Both of these points are reflected in the data here, in that while both sets of PMs can occur in any utterance position, there are clear preferences. DMs *you know* and *saber* overwhelmingly occur in an utterance-medial position, 80%, and are only found in initial and final positions in 9% and 11% of instances, respectively. Tags *no* and *qué no*, however, are more proportionally spread across positions, with a final position as the obvious preference at 45%, considered the canonical location for English and Spanish tags (Allerton 2009; Gómez González 2012, 2014). Furthermore, tags also occur in the medial and initial positions 33% and 21% of the time, respectively. Taking the results of the multivariate and distributional analyses together, and similar to Palacios Martínez's (2014) study of *innit*, we propose that, as tags in Arizona Spanish grammaticalize away from addressee-oriented functions, their syntactic ties are loosened *per se*, allowing for them to move to other structural positions, which would otherwise be occupied by PMs such as *you know* or *saber*. Furthermore, D'Arcy (2005) and Kern's (2019) findings that, as PMs grammaticalize they can move from clause boundaries to positions within the clause itself, might lend further insight into this phenomenon, and should be explored in future studies. The key finding here is that in Arizona Spanish, linguistically speaking, speakers rely mainly on position in deciding which PM to utilize in discourse.

What is more, as proposed by previous studies (Kluge 2011; Palacios Martínez 2014; Pichler 2013; Schleef and Mackay 2022), utterance position is posited to be closely correlated with discourse function. A cross-tabulation of these two factors verify this claim (Yates $p = 0$), as seen in Table 4.

Table 4. Cross-tabulation of broad discourse factor by utterance position for pragmatic markers in Arizona Spanish.

	Broad Discourse Function			
Position	***Addressee***	***Speaker***	***Exchange***	**Totals**
Initial	1	34	66	**101**
	2%	*31.5%*	*15%*	
Medial	2	42	250	**294**
	4%	*38%*	*58%*	
Final	45	35	116	**196**
	94%	*31.5%*	*27%*	
Totals	**48**	**111**	**432**	**591**

As seen above, addressee-oriented tokens occurred in a final position 94% of the time, whereas speaker-oriented tokens occurred more evenly split between the initial (31.5%), medial (38%), and final (31.5%) positions, and exchange-oriented tokens occurred more than half of the time in the medial position (58%), followed by the final (27%) and initial (15%) positions. Utilizing *no* as a case-in-point, we can see how a PM's position in reference to the utterance can display different discourse functions, as shown in examples (12), (13), and (14) below:

Addressee-oriented

(12) *Los mexicanos siempre usan vestido blanco, ¿**no**?* (CESA025)
'Mexicans always wear white dresses, right?'

Speaker-oriented

(13) *Y, **no**, ahora están muy mal los chamacos*. (CESA031)
'And, yeah, now the kids are really bad'.

Exchange-oriented

(14) *Siempre era ir con familia, ¿**no**?, a festejar.* (CESA044)
'It was always going with family, right, to celebrate'.

In reviewing the statistically significant crosstabulation of discourse function by position above, as well as the illustrative examples of how discourse function can shift based on the position in which a PM is used within the utterance, we take these findings as an indication that a final position will be correlated with functions which are conducive to a response, and a medial position with those which are not, thus confirming Pichler's (2013) hypothesis that position is correlated with function for tags and other PMs (Kluge 2011; Palacios Martínez 2014; Schleef and Mackay 2022).

Moving on, in the past, gender effects were either mixed for pragmatic variables (Childs 2021; Sankoff et al. 1997; Tagliamonte 2012), or not significant (Andersen 2022; Kern 2019; Kimps et al. 2014b; Said-Mohand 2007). For PM usage in Arizona Spanish, we do see an effect for self-reported gender, in that women favor usage of the DMs *you know* and *saber*, as seen in the factor weights presented in Figure 3 above, where DMs are favored for females (0.632) and disfavored for males (0.368). However, as seen in Table 5 below, results for tags are more evenly split between the genders.

Whereas females use approximately twice as many PMs as males, the percentage of tag usage between the two groups is approximately equal, with females using 52% of tags and males 48%. This is unsurprising given the established status of negative tags in Spanish. In comparison, females overwhelmingly use more DMs as compared to males, 85% vs 15%, respectively. While the results are skewed due to a larger number of females in the participant sample, it is worth bearing in mind that women oftentimes lead changes-in-progress, whether for an incoming standard variant or a less prestigious

one (Meyerhoff 2011). It is difficult to say with certainty which of these is the case here without more data, warranting further studies across multiple generations of speakers. It can be inferred, however, that as DMs grammaticalize from verbs-of-knowing to PMs, it is primarily females who are leading this change.

Table 5. Distribution of tags (*no* and *qué no*) vs DMs (*you know* and *saber*) by self-reported gender.

	Tags	DMs	Total
Female	200	177	**377**
	52%	*85%*	***64%***
Male	183	31	**214**
	48%	*15%*	***36%***
Total	**383**	**208**	**591**

The third factor found to be significant is the participants' self-reported length of residence. Due to the fact that most of our participants were millennials (born between 1980–1995), we decided to utilize length of residence as a proxy for age. More time living in one place is hypothesized as implying greater integration into the local speech community (Aaron and Hernández 2007; Esparza 2017; Sankoff et al. 1997; Torres and Potowski 2008). This is demonstrated in the multivariate analysis, in that those born/who had lived in Tucson for 20+ years were more likely to utilize DMs (0.612) than those who had lived in the city for less than 20 years (0.388). This is reflected in Table 6 below:

Table 6. Distribution of pragmatic marker Subtypes by length of residence in Tucson.

	Tags	DMs	Total
1–19 years	138	20	**158**
	36%	*10%*	***27%***
Birth/20+ years	245	188	**433**
	64%	*90%*	***73%***
Total	**383**	**208**	**591**

While those born in/have lived in Tucson for 20+ years use more PMs overall (73%), they also use significantly more DMs than their more recently arrived counterparts, 90% of all tokens. While we believe this indicates that PM usage may be constrained by a local standard, we also believe that it can be partially explained by specific patterning for *you know*. Utilizing a cross-tabulation of DM usage by length of residence, we see that whereas native Tucsonans/those who have lived there for 20+ years account for 80% of all instances of *saber*, this number rises to 95% for *you know*, indicating that use of this English variant is restricted nearly exclusively to those who are more integrated into this bilingual Borderlands community.

To conclude our discussion of which predictors condition PM selection in Arizona Spanish, the above results demonstrate that utterance position, self-reported gender, and length of residence were all significant in the choice between a tag (*no and qué no*) and a DM (*you know and saber*). Specifically, thus far, we have seen that a medial position is conducive to the use of a DM, while a final position is preferred for tags. Furthermore, while usage of tags is evenly split between males and females, DMs are overwhelmingly used by females in this community. Finally, while those born in/had lived in Tucson for 20+ years were more likely to use PMs overall, they were also found to be more likely to use DMs specifically, as compared to those who have more recently migrated to the area.

4.1. Discourse Function, Revisited

Returning now to discourse function, neither broad discourse function nor conduciveness to a response were found to be significant in the statistical analysis, a finding not unheard of (Childs 2021). Because both tags and DMs must undergo grammaticalization to be used as PMs, it is unsurprising that the cline of grammaticalization may lead them towards increasingly non-conducive functions (Carvalho and Kern 2019; Childs 2021; Croft 2022; Gómez González 2012, 2014; Molina 2005; Palacios Martínez 2014; Pichler 2013, 2021; Said-Mohand 2007; Travis 2005; Uclés Ramada 2020). This is perhaps compounded in this dataset by the finding that Spanish tends to exhibit more non-conducive functions than English overall (Gómez González 2014). That said, tags are assumed to be at a base-level addressee-oriented, even after grammaticalization (Gómez González 2012), some remnant of which can be seen in the breakdown by function in Table 7 below:

Table 7. Distribution of tags (*no* and *qué no*) vs DMs (*you know* and *saber*) by broad discourse function.

	Tags	**DMs**	**Total**
Addressee	44	4	**48**
	11%	*2%*	***8%***
Speaker	78	33	**111**
	20%	*16%*	***19%***
Exchange	261	171	**432**
	68%	*82%*	***73%***
Total	**383**	**208**	**591**

While both DMs and tags favor non-conducive (speaker-/exchange-oriented) functions, DMs do so overwhelmingly, at 98%, and are only addressee-oriented 2% of the time, while tags retain their original addressee-oriented functions 11% of the time. That said, the pragmatic system of Arizona Spanish seems to be grammaticalizing overall, in that 92% of all PMs are used with non-conducive functions, in line with past cross-linguistic findings. Furthermore, because the same pragmatic functions can be used to describe both PM subtypes, the fact that it was not found significant in the statistical analysis is telling, in that the system has grammaticalized to a point where function has ceased to be a deciding factor in PM selection for speakers. This has two possible implications: either tags and other types of PMs were never differentiated in the first place, the unlikely option, or grammaticalization eventually erases these differences entirely, driving tags towards increasingly non-conducive usages (Pichler 2013) and rendering them no different than other PMs (Palacios Martínez 2014).

As a final point on function, we identify here the distribution of narrow discourse functions, to see a breakdown of the purposes for which PMs are used in Southern Arizona. These results are presented in Table 8 and Figure 4.

Table 8. Distribution by *conduciveness* (Pichler 2013), broad and narrow discourse functions for pragmatic markers in Arizona Spanish.

	Conducive		***Non-Conducive***				
	Addressee		**Speaker**		**Exchange**		
	Information	**Action**	**Attitudinal**	**Challenging**	**Focusing**	**Phatic**	**Discursive**
N	47	1	67	44	243	39	150
%	8%	0%	11%	7%	41%	7%	25%

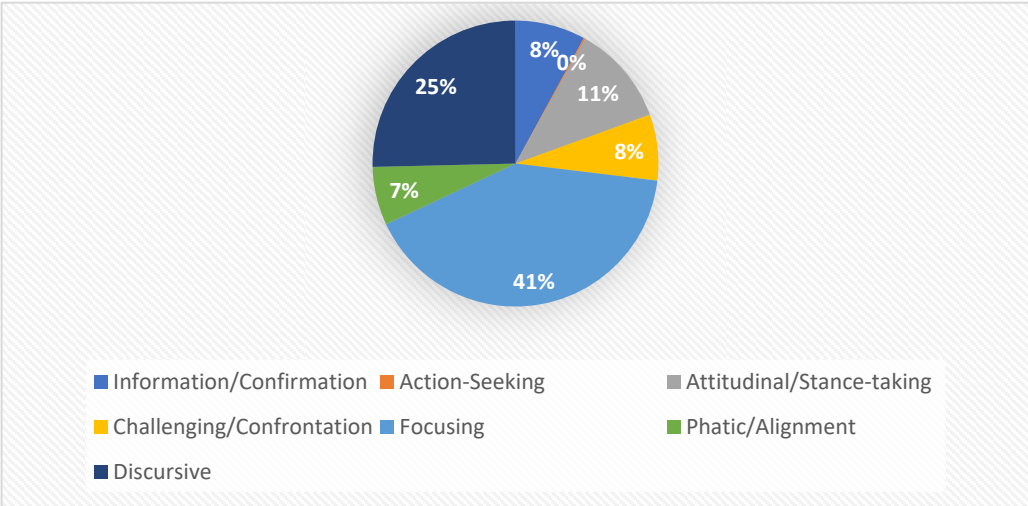

Figure 4. Pragmatic marker distribution by narrow discourse function in Arizona Spanish.

As can be seen above, focusing (41%) is the most frequent narrow discourse function for this dialect, followed by discursive (22%) and attitudinal/stance-taking (11%).

To conclude our discussion on function, we would like to emphasize two points. Due to the grammaticalization of the system overall, both tags and DMs as subtypes of PMs are moving towards functions which are non-conducive to a response; nevertheless, tags still retain some of their original addressee-oriented functionality (11%). Furthermore, because of this grammaticalization, function is found to be insignificant when the data is submitted to statistical analysis, indicative of the fact that when looking to utilize a PM in discourse, speakers of Arizona Spanish no longer rely on function in deciding which one to use. Instead, this choice is constrained by the PM's position within an utterance, as well as the gender and length of residence of the speaker in the local community.

4.2. You Know and Language Environment

Finally, we examine the incoming variant *you know* as it makes its way into Arizona Spanish. As in several past studies (Aaron 2004; Andersen 2022; Carvalho and Kern 2019; Sankoff et al. 1997), we take the use of *you know* in this variety as evidence for contact-induced change, presenting a discussion of how this variant influences codeswitching behavior, and hypothesizing as to its possible long-term impact on the system itself.

It is well-known that codeswitching triggered by a PM is dependent upon the marker's degree of integration into the system. For example, in their study of *so* and *entonces* in Chicago, Torres and Potowski (2008) found that *so* was so incorporated into the local dialect that it only triggered codeswitching behavior 4% of the time. In a similar vein, Aaron (2004), in her study of New Mexico, found that while *so* was relatively well-entrenched into the local variety, it still only triggered a codeswitch to English 16% of the time. Conversely, Kern (2019) found that, while the variants *like*, *como*, and *como que* in Arizona Spanish were semantically similar and used in parallel ways in both languages, there was very little crossover between languages, with *like* only being used 17 times for every 10,000 words in Spanish, and with no crossover to English. Therefore, given the fact that the present study also analyzes Arizona Spanish, it is reasonable to expect similar results for the PMs reviewed here.

This study considers four environments where *you know* was found in the dataset. The first is an environment where the variant is preceded and followed by English, as in (15) below. Next is a preceding Spanish environment with a following codeswitch to English, as in (16). The third is preceding English with a codeswitch to Spanish, as in (17). The last is a variant that is preceded and followed by Spanish (what Poplack 1980, might refer to as a tag switch), as in (18).

(15) *I was happy, I was content with myself,* ***you know****, just either playing or whatever.* (CESA076)
(16) *Pues la fiesta después,* ***you know****, like the after party.* (CESA049)
'Well the party after, you know, like the after party'.
(17) *Like somebody who only speaks Spanish is trying to find something or,* ***you know****, alguien que se ayuda, and nobody speaks Spanish.* (CESA076)
'Like somebody who only speaks Spanish is trying to find something or, you know, somebody that can help, and nobody speaks Spanish'.
(18) *Me gusta la cultura y soy,* ***you know****, rodeado con la cultura.* (CESA027)
'I like the culture and I'm, you know, surrounded with the culture'.

Given the fact that *you know* is an incoming loan from English into Spanish in Southern Arizona, by taking the above four language environments into consideration, we can see how exactly its incorporation into the system affects language choice and codeswitching behavior for this variety, as compared to native Spanish variants. A breakdown of these results can be seen in Table 9 below:

Table 9. Language environment and codeswitching behavior of Spanish-origin variants vs. *you know* (N = 84 tokens were excluded from this count due to their occurrence at a turn boundary, whether initial or final).

	Spanish Variants	***You Know***	**Total**
English→English	4	50	**54**
	1%	*35%*	***11%***
Spanish→English	9	32	**41**
	2%	*22%*	***8%***
English→Spanish	6	11	**17**
	2%	*8%*	***3%***
Spanish→Spanish	345	50	**395**
	95%	*35%*	***78%***
Total	**364**	**143**	**507**

Extending findings for *so* and *entonces,* as presented by Aaron (2004) and Torres and Potowski (2008), we can infer that an incoming loan's degree of integration into a recipient language can be demonstrated by proxy via the amount of time where it is used in a continuous (Spanish→Spanish) or codeswitched (English→Spanish) environment. Therefore, while only accounting for approximately 19% of the dataset here, *you know* is split fairly evenly for which language it conditions in the following environment. A following English-language clause is conditioned 57% of the time, while a following Spanish-language clause 43% of the time (compare the native Spanish variants, with a following English environment only 3% of the time, and a following Spanish environment 97%). In comparing this to Kern's (2019) finding that there is relatively little language mixing triggered by the incoming variant *like* in Spanish discourse, we assume that *you know* represents a further degree of integration into a monolingual Spanish mode for Arizona Spanish.

With that in mind, what are the possible effects that *you know* might have as it moves into the pragmatic system of this variety? It is well-known that pragmatic borrowing is common (Andersen 2022; Andersen et al. 2017). Indeed, pragmatics is a key reason why borrowing in contact situations occurs in the first place, in that loanwords might be filling a functional gap, bringing over functions from the donor language, or acquiring new functions in the recipient (Andersen 2022; Andersen et al. 2017; Bybee 2015). However, given the similarity in meaning and usage between *saber* and *you know* (Kluge 2011; Said-Mohand 2007; Sankoff et al. 1997), these possibilities do not seem to be the case here, in that both variants show similar broad functional distributions, being 97% and 99% non-

conducive, respectively. What is more likely is that, for the time being, *you know* seems to be co-existing with *saber*, as it moves into the system, sharing functions and distribution without attrition of the native variant (Aaron 2004; Andersen 2022; Sankoff et al. 1997; Torres and Potowski 2008). This can be seen in (19) below, where the speaker self-repairs from one variant to the other, further demonstrating their functional equivalence:

(19) ***You know***, ***sabes qué***, *tengo una sobrina, la hija de mi tía*... (CESA015)
'You know, you know, I have a niece, the daughter of my aunt...'

However, there are other long-term possibilities. For example, it is possible that contact situations can accelerate the grammaticalization process (Childs 2021; Palacios Martínez 2014; Pichler 2021), that the pragmatic repertoires of bilinguals have been shown to be more advanced in terms of grammaticalization (Carvalho and Kern 2019), and that, ultimately, as an incoming variant moves in, it may restructure the system, as it takes on new functions, forcing other variants to either adapt and move into their own functional niches or fade from use entirely (Aaron 2004; Andersen 2022; Childs 2021; Pichler 2021; Torres and Potowski 2008). Even if this does not occur, as *you know* grows in frequency, while sharing *saber*'s functional distribution, it is possible that lexical replacement may occur, as happened with *right* replacing *you know* in Canadian English (Denis and Tagliamonte 2016). In any case, the role that *you know* plays in the pragmatic system of Arizona Spanish, while incipient and in need of further longitudinal/apparent-time data, is critical in illustrating how pragmatic systems in contact can change and adapt.

5. Conclusions and Future Study

In conclusion, over the course of this study, we have demonstrated that utterance position, gender, and length of residence were all found to be significant in conditioning the choice between tags and DMs in Arizona Spanish. Importantly, regarding the research questions proposed in the introduction, the main takeaways from this work are threefold. First, the fact that discourse function was not significant provides us with incipient evidence indicating that functional differences between tags and DMs in this variety have been reduced to insignificance, as both categories move towards non-conducive functions. Because tags are hypothesized to be functionally different than other types of PMs, in that tags are supposed to be used for more addressee-oriented functions, this is important because it provides further evidence that grammaticalization can be a driving force not only in differentiating linguistic systems but also in levelling differences within said systems. Further study will be needed to verify this hypothesis. Second, the results demonstrate that, although function is not statistically significant in differentiating tags from DMs for this dialect, there is a correlation between utterance position and discourse function, which allows the speaker to better differentiate between the two in selecting a PM for use in discourse. And finally, for the English-origin loan *you know*, it is clear that this variant has slowly been incorporated into Arizona Spanish, as demonstrated by its distribution within codeswitched environments, where it is taking its place alongside *saber*, without attrition of this native variant.

Future research should compare the distribution of the negative-polarity tags *no* and *qué no* to the positive-polarity tags *sí* and *verdad*, as these PMs are also relatively frequent in CESA, and Pichler (2013) has previously noted that the polarity of the tag should correlate with different functions. In addition, it will be important to conduct an in-depth syntactic analysis of PM distribution, in the vein of D'Arcy (2005) and Kern (2019), to evaluate to what degree they have grammaticalized from markers to particles, available for use within clause boundaries. Finally, a comparative sociolinguistic (Tagliamonte 2003) analysis of this variety in comparison to the sister dialects in Sonora and New Mexico would provide critical insight into the local behavior of PMs as compared to regional trends.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data available upon request.

Acknowledgments: The author would like to thank Ana Carvalho and the members of the Sociolinguistics Research Lab (SoReLa) at the University of Arizona for valuable feedback throughout this process, as well as two anonymous reviewers for their suggestions on the manuscript. Previous iterations of the work were presented at two conferences, the *32nd Annual Graduate and Professional Symposium on Hispanic and Luso-Brazilian Literature, Language and Culture* (17–18 February 2023; Tucson, AZ, USA) and the *51st New Ways of Analyzing Variation* conference (13–15 October 2023; Queens, NY, USA). While all efforts at accuracy were made in the preparation of the study, it is not without fault; all errors remain my own.

Conflicts of Interest: The author declares no conflict of interest.

Notes

[1] All translations are my own.

[2] "[Marcadores de discurso] son unidades lingüísticas invariables, no ejercen una función sintáctica en el marco de la predicación oracional (o sea, están fuera de la sintaxis oracional) y poseen un propósito en el discurso: el de guiar las inferencias que se realizan en la comunicación. O sea, señalan las relaciones que existen entre unidades del discurso" (Portolés Lázaro 1998, pp. 25–26).

[3] We thank one reviewer for observing that whether we account for PM development under the framework of grammaticalization or not ultimately does not change the significance of our findings.

[4] While other factors were coded for, including intonation, anchor mood, and polarity, they are outside the scope of the current work, and will not be discussed further.

References

Aaron, Jessi. 2004. "So respetamos un tradición del uno al otro"—So and entonces in new mexican bilingual discourse. *Spanish in Context* 1: 161–79. [CrossRef]

Aaron, Jessi, and José Hernández. 2007. Quantitative evidence for contact-induced accommodation: Shifts in /s/ reduction patterns in salvadoran spanish in houston. In *Spanish in Contact: Policy, Social and Linguistic Inquiries*. Edited by Kim Potowski and Richard Cameron. Amsterdam: John Benjamins, pp. 329–43.

Allerton, D. J. 2009. Tag questions. In *One Language, Two Grammars: Differences between British and American English*. Edited by Günter Rohdenburg and Julia Schlüter. Cambridge: Cambridge University Press, pp. 306–23.

Andersen, Gisle. 2022. What governs speakers' choices of borrowed vs. domestic variants in discourse-pragmatic variables. In *Discourse-Pragmatic Variation and Change: Theory, Innovations, Contact*. Edited by Elizabeth Peterson, Turo Hiltunen and Joseph Kern. Cambridge: Cambridge University Press, pp. 251–71.

Andersen, Gisle, Cristiano Gino Furiassi, and Biljana Mišić Ilić. 2017. The pragmatic turn in studies of linguistic borrowing. *Journal of Pragmatics* 113: 71–76. [CrossRef]

Aschenberg, Heidi, and Óscar Loureda Lamas. 2011. *Marcadores del Discurso: De la Descripción a la Definición*. Madrid: Iberoamericana Editorial Vervuert.

Bybee, Joan. 2011. Usage-based theory and grammaticalization. In *The Oxford Handbook of Grammaticalization*. Edited by Bernd Heine and Heiko Narrog. Oxford: Oxford University Press, pp. 69–78.

Bybee, Joan. 2015. *Language Change*. Cambridge: Cambridge University Press.

Carvalho, Ana. 2012-. *Corpus del Español en el Sur de Arizona (CESA)*. Tucson: University of Arizona.

Carvalho, Ana, and Joseph Kern. 2019. The permeability of tag questions in a language contact situation: The case of spanish-portuguese bilinguals. *Pragmatics* 29: 463–92. [CrossRef]

Childs, Claire. 2021. Mechanisms of grammaticalization in the variation of negative question tags. *Journal of English Linguistics* 49: 419–58. [CrossRef]

Clayman, Steven, and Chase Raymond. 2021. You know as invoking alignment: A generic resource for emerging problems of understanding and affiliation. *Journal of Pragmatics* 182: 293–309. [CrossRef]

Crible, Ludivine, and Liesbeth Degand. 2019. Domains and Functions: A Two-Dimensional Account of Discourse Markers. *Discours* 24: 3–35. [CrossRef]

Croft, William. 2022. *Morphosyntax: Constructions of the World's Languages*. Cambridge: Cambridge University Press.

D'Arcy, Alexandra. 2005. Like: Syntax and Development. Ph.D. dissertation, University of Toronto, Toronto, ON, USA.

Degand, Liesbeth, Zoé Broisson, Ludivine Crible, and Karolina Grzech. 2022. Cross-linguistic variation in spoken discourse markers. In *Discourse-Pragmatic Variation and Change: Theory, Innovations, Contact*. Edited by Elizabeth Peterson, Turo Hiltunen and Joseph Kern. Cambridge: Cambridge University Press, pp. 83–103.

Denis, Derek, and Sali Tagliamonte. 2016. Innovation, right? change, you know? utterance-final tags in canadian english. In *Discourse-Pragmatic Variation and Change in English: New Methods and Insights*. Edited by Heike Pichler. Cambridge: Cambridge University Press, pp. 86–112.

Esparza, Sandra. 2017. /s/ Aspiration in Ciudad Juárez and Speech Accommodation. Master's thesis, University of Texas at El Paso, El Paso, TX, USA.

García Vizcaíno, María José. 2005. El uso de los apéndices modalizadores ¿no? y ¿eh? en español peninsular. In *Selected Proceedings of the Second Workshop on Spanish Sociolinguistics*. Edited by Lotfi Sayahi and Maurice Westmoreland. Somerville: Cascadilla Proceedings Project, pp. 89–101.

Gazdik, Anna. 2022. On the prosodic realization of spanish ¿no?—Tags from a pragmatic perspective. *Isogloss* 8: 1–19. [CrossRef]

Gómez González, María de los Ángeles. 2012. The question of tag questions in english and spanish. In *Encoding the Past, Decoding the Future: Corpora in the 21st Century*. Edited by Isabel Moskowich and Begoña Crespo. Cambridge: Cambridge Scholars Press, pp. 59–95.

Gómez González, María de los Ángeles. 2014. Canonical tag questions in english, spanish and portuguese. *Languages in Contrast* 14: 93–126. [CrossRef]

Johnson, Daniel. 2009. Getting off the goldvarb standard: Introducing rbrul for mixed-effects variable rule analysis. *Language and Linguistics Compass* 3: 359–83. [CrossRef]

Johnson, Daniel. 2023. *Rbrul* (version 4.2.2). Software. Available online: http://www.danielezrajohnson.com/rbrul.html (accessed on 18 December 2023).

Kern, Joseph. 2019. Like in english and como, como que, and like in spanish in the speech of southern arizona bilinguals. *International Journal of Bilingualism* 24: 184–207. [CrossRef]

Kimps, Ditte, Kristin Davidse, and Bert Cornillie. 2014a. A speech function analysis of tag questions in british english spontaneous dialogue. *Journal of Pragmatics* 66: 64–85. [CrossRef]

Kimps, Ditte, Kristin Davidse, and Bert Cornillie. 2014b. The speech functions of tag questions and their properties: A comparison of their distribution in colt and llc. In *Corpus Interrogation and Grammatical Patterns*. Edited by Kristin Davidse, Caroline Gentens, Lobke Ghesquière and Lieven Vandelanotte. Amsterdam: John Benjamins, pp. 321–50.

Kluge, Bettina. 2011. Camino de un marcador del discurso: Una comparación del español ¿sabes? Con el francés tu sais y el inglés you know. In *Marcadores del Discurso: De la Descripción a la Definición*. Edited by Heidi Aschenberg and Óscar Loureda Lamas. Madrid: Iberoamericana Editorial Vervuert, pp. 305–41.

Martín Zorraquino, María, and José Portolés Lázaro. 1999. Los marcadores del discurso. In *Gramática Descriptiva de la Lengua Española*. Edited by Ignacio Bosque and Violeta Demonte. Madrid: Real Academia Española, vol. 3, pp. 4051–213.

Meyerhoff, Miriam. 2011. *Introducing Sociolinguistics*. London: Routledge.

Molina, Isabel. 2005. La moda del ¿sabes? En madrid: Un análisis sociolingüístico. In *Filología y Lingüística: Estudios Ofrecidos a Antonio Quilis*. Madrid: Consejo Superior de Investigaciones Científicas—Universidad de Valladolid, vol. 1, pp. 1045–56.

Ocampo, Francisco. 2006. Movement towards discourse is not grammaticalization: The evolution of claro from adjective to discourse particle in spoken spanish. In *Selected Proceedings of the 9th Hispanic Linguistics Symposium*. Edited by Nuria Sagarra and Almeida Toribio. Somerville: Cascadilla Proceedings Press, pp. 308–19.

Onodera, Noriko. 2011. The grammaticalization of discourse markers. In *The Oxford Handbook of Grammaticalization*. Edited by Bernd Heine and Heiko Narog. Oxford: Oxford University Press, pp. 614–24.

Otheguy, Ricardo, Ofelia García, and Wallis Reid. 2015. Clarifying translanguaging and deconstructing named languages: A perspective from linguistics. *Applied Linguistics Review* 6: 281–307. [CrossRef]

Palacios Martínez, Ignacio. 2014. Variation, development and pragmatic uses of innit in the language of British adults and teenagers. *English Language and Linguistics* 19: 383–405. [CrossRef]

Pichler, Heike. 2013. *Structure of Discourse-Pragmatic Variation*. Amsterdam: John Benjamins.

Pichler, Heike. 2021. Grammaticalization and language contact in a discourse-pragmatic change in progress: The spread of innit in London English. *Language in Society* 50: 723–61. [CrossRef]

Poplack, Shana. 1980. Sometimes I'll start a sentence in spanish y termino en español: Towards a typology of code-switching. *Linguistics* 18: 581–618. [CrossRef]

Portolés Lázaro, José. 1998. *Marcadores del discurso*. Barcelona: Ariel.

Said-Mohand, Aixa. 2007. Aproximación sociolingüística del marcador del discurso tú sabes en el habla de jóvenes bilingües estadounidenses. *Southwest Journal of Linguistics* 26: 67–93.

Sankoff, Gillian, Pierrette Thibault, Naomi Nagy, Hélène Blondeau, Marie-Odile Fonollosa, and Lucie Gagnon. 1997. Variation in the use of discourse markers in a language contact situation. *Language Variation and Change* 9: 191–217. [CrossRef]

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

Schleef, Erik, and Bradley Mackay. 2022. Evaluation of pragmatic markers: The case of you know. In *Discourse-Pragmatic Variation and Change: Theory, Innovations, Contact*. Edited by Elizabeth Peterson, Turo Hiltunen and Joseph Kern. Cambridge: Cambridge University Press, pp. 40–60.

Silva-Corvalán, Carmen, and Andrés Enrique-Arias. 2017. *Sociolingüística y pragmática del español*. Washington: Georgetown University Press.

Tagliamonte, Sali. 2003. Comparative sociolinguistics. In *The Handbook of Language Variation and Change*. Edited by J. K. Chambers, Peter Trudgill and Natalie Schilling-Estes. Hoboken: Wiley, pp. 729–63.
Tagliamonte, Sali. 2006. *Analysing Sociolinguistic Variation*. Cambridge: Cambridge University Press.
Tagliamonte, Sali. 2012. *Variationist Sociolinguistics: Change, Observation, Interpretation*. Hoboken: Wiley-Blackwell.
Torres, Lourdes, and Kim Potowski. 2008. A comparative study of bilingual discourse markers in chicago mexican, puerto rican, and mexirican spanish. *International Journal of Bilingualism* 12: 263–79. [CrossRef]
Traugott, Elizabeth, and Richard Dasher. 2001. *Regularity in Semantic Change*. Cambridge: Cambridge University Press.
Travis, Catherine. 2005. *Discourse Markers in Colombian Spanish: A Study in Polysemy*. Berlin: Mouton de Gruyter.
Uclés Ramada, Gloria. 2020. Las funciones interactivas del marcador español '¿no?' las fronteras entre la atenuación y protección de la imagen. *Revista Signos: Estudios de Lingüística* 53: 790–814. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.