ISSN 2385-4138 (digital)
https://doi.org/10.5565/rev/isogloss.212

Isogloss 2022, 8(1)/11
1-23

The influence of first and second language on the acquisition of pragmatic markers in Spanish: Evidence from an experimental study

Gijs Mulder
Radboud University
gijs.mulder@ru.nl

Gert-Jan Schoenmakers
Radboud University
gert-jan.schoenmakers@ru.nl

Helen de Hoop
Radboud University
helen.dehoop@ru.nl

Received: 28-03-2022
Accepted: 06-07-2022
Published: 07-07-2022

How to cite: Mulder, Gijs, Schoenmakers, Gert-Jan & Helen de Hoop. 2022. The influence of first and second language on the acquisition of pragmatic markers in Spanish: evidence from an experimental study. *Isogloss. Open Journal of Romance Linguistics* 8(1)/11, 1-23.
DOI: https://doi.org/10.5565/rev/isogloss.212

Abstract

This paper reports on an experimental study of the use of two Spanish markers of epistemic modality and evidentiality, *creo que* 'I believe that' and *pienso que* 'I think that', by native speakers, and by Dutch and German learners of Spanish. We found a clear preference for *creo que* among the native speakers of Spanish, but with differences between the main varieties of the language. For Dutch and German learners

the preference for *creo que* was significantly weaker, and for beginning learners of Spanish it was significantly weaker than for advanced learners. While for the Dutch learners of Spanish this pattern reflects the preference for *I think* over *I believe* in their L1, the preference that the German-speaking learners have for *creo que* can be seen as evidence for a general tendency noticed in the literature, namely the interference of an L2 (in this case English) instead of the L1 (in this case German) in L3 performance (in this case Spanish).

Keywords: experiment, epistemic/evidential expression, Dutch, German, English.

1. Introduction

Understanding and using pragmatic markers in a second language can be difficult, even for proficient speakers. The reason is that pragmatic markers are particularly common in spontaneous speech and usually serve multiple functions depending on the context in which they are used. Pragmatic markers do not contribute to the propositional content, but speakers use them to guide the hearer in the interpretation process: they give structure and coherence to the discourse or mark the speaker's stance towards the conversation or the hearer (Fraser 1999, Foolen 2011, Aijmer & Simon-Vandenbergen 2003, Hogeweg et al. 2016). Their presence and appropriate usage are thus strongly associated with native speaker competence (Haselow 2021). A general problem for non-native speakers is that while many pragmatic markers are etymological and/or semantic cognates, they often correspond only partially or not at all in meaning (Aijmer & Simon-Vandenbergen 2006). Languages also exhibit differences in the frequencies and pragmatic functions of those markers that are each other's correlates (Romero-Trillo 2019).

Research on the L2-acquisition of pragmatic markers reveals varying patterns. Numerous corpus-based studies have shown that L2-speakers tend to use pragmatic markers less frequently or with less variation than L1-speakers in similar contexts (Müller 2015, Staples & Fernández 2019). However, specific markers are used and developed by learners in different ways. Depending on their function and the stage of the learning process, L2-speakers use them excessively or not at all (Polat 2011, Ament et al. 2020). The existence of L1-correlates might influence the L2-acquisition of pragmatic markers, even if there are no exact form-function relations between the markers of the two languages (Hogeweg et al. 2016).

In this paper we focus on two cognition verbs in first person, *I believe* and *I think*, that are used as markers of indirect evidentiality and epistemic modality in different ways in different languages. *I believe* and *I think* are markers of evidentiality in that they "show the kind of justification for a factual claim which is available to the person making that claim", Anderson's (1986: 274) first criterion for evidential expressions, and "are not themselves the main predication of the clause", Anderson's (1986: 274) second criterion. Notoriously, indirect evidentiality overlaps with epistemic modality, the domain pertaining to the speaker's subjective evaluation of the degree of certainty of some propositional content. Especially in the case of inferential evidentiality, the two domains overlap, since they are related to speakers' cognitive states (Wiemer 2018). Evidence indicated by verbs of cognition such as *believe* and *think* is clearly subjective, *i.e.*, it is not necessarily accessible to hearers. There are different ways in which speakers may have obtained such evidence; in Dutch, for

example, the present tense expression *denk ik* 'I think' is used for inferred evidentiality, whereas the past tense variant *dacht ik* 'I thought' is used for assumed evidentiality (Griffioen, de Hoop & Mulder 2018).

Empirical studies of various (unrelated) languages show that evidential expressions, including markers such as *I think* and *I believe*, often have unreliability or uncertainty as part of their meaning (cf. Simon-Vandenbergen 2000, San Roque & Bergqvist 2015, Spronck 2015, Foolen et al. 2018). The addition of *I think* or *I believe* to a main proposition thus signals a degree of epistemic (un)certainty, while the speaker has reasons to believe, based on certain evidence, that something is, was, or will be the case (cf. Anderson's first criterion, quoted above).

Diachronic studies confirm the close relationship between indirect evidentiality and epistemic modality. Epistemic parentheticals like *I think* and *I believe* underwent a route through a process of increasing subjectification (Traugott 1995: 38), from the act of cognition to evidentiality and epistemic modality (Brinton 1996: 163). English, Dutch, German, and Spanish are among the many languages that have developed such evidential/epistemic usage of the expressions *I think* and *I believe* (see for example Thompson & Mulac 1991, on English; de Hoop et al. 2018, on Dutch; Schoonjans 2012, on German; and Mulder 2017, 2018, on Spanish).

In Spanish, the phrase *(yo) creo* 'I believe' usually combines with a complement clause introduced by the complementizer *que* 'that', but it can also occur in medial and final position as a parenthetical construction, signaling a process of grammaticalization (Mulder 2018). Likewise, de Hoop et al. (2018) assume that the Dutch evidential parentheticals *geloof ik* 'believe I' and *denk ik* 'think I', which are attached to the end of a sentence or inserted in the middle, have developed out of the evidential/epistemic meanings of the main clauses *ik geloof dat* 'I believe that' and *ik denk dat* 'I think that'. This process of grammaticalization can go even further. In Afrikaans, the expression *ek glo* 'I believe' has developed into a fully grammaticalized discourse marker *glo* '(I) believe', expressing indirect evidentiality (Boye & Harder 2009: 19). Similarly, Schoonjans (2012) argues that German developed a discourse marker *glaub(e)* '(I) believe' from the parenthetical use of *ich glaub(e)* 'I believe'. For English, Aijmer (1997) has argued that *I think* has developed into a pure discourse marker.

Thus, there is a lot of cross-linguistic evidence for a grammaticalization cline of expressions like *I think* and *I believe*. It has also been noted that variants of *I think* and *I believe*, such as with or without a complementizer, can have various functions irrespective of their degree of grammaticalization. For example, Aijmer (1997) argues that the presence or absence of the complementizer in *I think (that)* does not determine its grammatical status as either a main clause proposition or an epistemic modifier or discourse marker. Kearns (2007) convincingly argues that, although the complementizer is usually absent in highly frequent epistemic markers such as *I think (that)*, its absence is in no way indicative of its grammatical status. Instead, Kearns (2007: 503) concludes that the absence of the complementizer "correlates with a shift in informational prominence, where the embedded clause becomes more prominent than the matrix", but this is irrespective of its particular interpretation or syntactic status in a given context. In other words, there is no one-to-one mapping from forms to meanings, nor vice versa (cf. Dehé & Wichmann 2010). Schoonjans (2012) describes the different uses of German *glaub(e)* 'believe' as a multidimensional continuous space, rather than as a simple continuous scale along the lines of

grammaticalization. We therefore assume, following Dehé & Wichmann (2010) among others, that there is no one-to-one mapping between the form and interpretation of expressions such as *I think (that)* and *I believe (that)*.

Given the above considerations, the use of expressions like *I think (that)* and *I believe (that)* might be relevant for the study of pragmatic transfer (Kasper 1992; 2010). Félix-Brasdefer and Lavin (2009) found that, among English learners of Spanish, these expressions belong to the preferred forms for the expression of epistemic modality and opinions. Since these verbs do not fully correspond structurally, functionally, and distributionally to their counterparts in L2, their use by learners might give rise to negative pragmatic transfer (Kasper 2010: 147; Pearson & Hasler-Barker 2021: 425).

Although Spanish *creo* 'I believe' is still used in its original, lexical meaning 'have faith or confidence (in someone or something)' or 'hold (something) as true', it more frequently functions as a marker of inferred or assumed evidentiality, epistemic modality, or personal belief (Mulder 2018). There is a limited set of syntactic structures for *creo* in these functions. In the 20th century oral section of the *Corpus del español* (Davies 2002), *creo* – with or without personal pronoun *yo* 'I'– is followed by a *que*-complement clause in about 85% of its occurrences. Mulder (2018) found that in Peninsular Spanish *creo* virtually never has its lexical meaning in combination with a complementizer *que* 'that'.

The present experimental study focuses on two pragmatic markers, *creo que* 'I believe that' and *pienso que* 'I think that', *i.e.*, variants with a null subject and a complement clause introduced by the complementizer *que* 'that'. We will compare native speakers' preferences for either of these two expressions to the preferences of native Dutch-speaking and native German-speaking learners of Spanish. In order to keep the experiment as simple and straightforward as possible, we abstract away from any different interpretations these two forms may have in the domains of evidentiality, epistemic modality, and (inter)subjectivity. In all experimental contexts we provide, another proposition than *creo* 'I believe' or *pienso* 'I think' carries the primary information, while *creo* or *pienso* functions as a marker of secondary information, in accordance with its evidential/epistemic use.

Below are two of the items used in the experiment, which illustrate the evidential/epistemic readings of the marker *creo que* 'I believe that'. All sentences in our study were initially taken from Twitter, as will be explained in Section 2. Note that (1) illustrates a case of inferential evidentiality: the speaker infers from seeing the dark clouds in the sky that it is going to rain. In (2), *creo que* is used as a marker of epistemic modality, presumably in combination with hearsay evidentiality.

(1) *Creo que va a llover, porque hay muchas nubes negras en el cielo.*
'I believe it's going to rain, because there are dark clouds in the sky.'

(2) *¿Juega Messi en el partido de hoy? - Ni idea, creo que está lesionado.*
'Is Messi playing?' 'I don't know, I believe he's injured.'

In Spanish, the phrase *pienso* is far less common than *creo* (Mulder 2018). Mulder (2018) compares the occurrences of three evidential/epistemic expressions, namely *creo (que)* 'I believe (that)', *pienso (que)* 'I think (that)', and *me parece (que)* 'it seems to me (that)' in two large on-line Spanish corpora. He finds that *creo* is by far the most

frequent of the three (75% in both corpora), while *pienso* is only used in 10%. Also, *pienso* less frequently combines with the complementizer *que* 'that' (65%) than *creo* (85%). Finally, although *pienso que* is frequently used as an evidential/epistemic expression, Mulder concludes that it also often refers to the actual act of thinking or reasoning, thus expressing primary information. An example is given in (3) (Mulder 2018: 106):

(3) *A veces pienso que sería mejor si ni siquiera hubiera nacido.*
'Sometimes I think that it would have been better if I had never been born.'

To sum up, *creo (que)* 'I believe (that)' is used more frequently than *pienso (que)* 'I think (that)' in Spanish, and in addition it is used relatively more frequently as a marker of secondary information, *i.e.*, to denote evidentiality, epistemic modality, or personal belief.

Like in Spanish, Schoonjans (2012) finds in his German dataset that *(ich) glaub(e)* 'I believe' is by far the most frequent expression of this type (2392 occurrences), followed at a large distance by *(ich) denk(e)* 'I think' (96 occurrences). Wilton's (2019) corpus of 57 interviews in German with German professional football players also contained more instances of *ich glaub(e)/glaub(e) ich* 'I believe' (69 instances) than of *ich denk(e)/denk(e) ich* 'I think' (48 instances).

The relative frequencies of *I believe* and *I think* seem to be the other way around in Dutch, as suggested by de Hoop et al. (2018: 93) and confirmed by our corpus study presented in Section 4 below. English *I think* is more frequent than *I believe* as well (Baumgarten & House 2010). Baumgarten & House (2010) find that *I think* is the single most frequent *I* + verb combination in corpora of both spoken American and British English (see also Thompson & Mulac 1991, Simon-Vandenbergen 2000).

Thus, whereas Spanish and German prefer 'I believe' as a marker of evidentiality and epistemic modality, in Dutch and English 'I think' is the preferred expression. This raises the question where these different preferences across languages come from, since believing and thinking are clearly different cognitive experiences. The Spanish philosopher Ortega y Gasset established a distinction between *ideas* 'ideas' (the act and effect of thinking) and *creencias* 'beliefs' and famously wrote: *Las ideas se tienen, en las creencias se está* 'Ideas are ours, in beliefs we are' (Ortega y Gasset 1940: 3). 'Ideas' are the thoughts that arise about a particular aspect of our environment or our inner life; we produce them or we adopt them from others. 'Beliefs', on the other hand, are always there, they regulate our existence; we count on them, but we are not explicitly aware of them. For example, in order to go out on the street, we do not need to reflect on the fact that the street exists. Assuming this fundamental difference between thinking and believing, it is difficult to see a rationale behind the preferences languages have for one or the other verb for their pragmatic expressions. Is it because the Dutch and English are more focused on the results of a reflection than the Spanish and German? Or are the Spanish and German more interested in shared beliefs? We will not speculate about possible explanations for these differences among the four languages. Instead, we will focus on the proportions of *creo que* and *pienso que* in native Spanish speakers, compared to Dutch and German learners of Spanish, hypothesizing that the latter two groups are influenced by transfer from their L1.

The aim of the experimental study is twofold. First, while Mulder's (2018) study was limited to Peninsular Spanish, we will investigate whether the preference for *creo que* 'I believe that' over *pienso que* 'I think that' holds for native speakers from the varieties of Spanish in Latin-America as well. Second, we put forward the hypothesis that German-speaking learners of Spanish will closely resemble the native Spanish speakers in preferring *creo* over *pienso* (Mulder 2018), since this preference resembles the one in their mother tongue (Schoonjans 2012, Wilton 2019), while Dutch learners of Spanish will use *pienso* relatively more frequently than the native Spanish speakers, under the influence of their L1 (de Hoop et al. 2018).

2. The experiment

We conducted an on-line questionnaire to investigate the preference for *creo que* 'I believe that' or *pienso que* 'I think that' by three groups of Spanish speakers: native speakers, Dutch L2-speakers, and German L2-speakers of Spanish. The stimuli and data are available in a repository.[1]

2.1. Participants

405 participants completed the questionnaire. Data from 94 participants were excluded on the basis of six criteria defined prior to statistical analysis: i) their native language was not Spanish, Dutch, or German; ii) they indicated to have both Spanish and Dutch or German as their native language; iii) their native language was Spanish, Dutch, or German, but they lived in countries where this was not the common language; iv) they indicated that they had moved to another country before the age of 11 (in which case we cannot know which language influenced them the most); v) Spanish was not listed as one of the languages they speak; or vi) they chose other options than *creo* or *pienso* in at least three out of twelve target sentences. Data from 311 participants (80 male, 224 female, 7 other/prefer not to say; age range: 16-74, mean age: 34.2, SD: 13.0) were entered into statistical analysis. The participants came from sixteen different countries.

The L1 Spanish participants came from Spain and eleven Spanish speaking countries of Latin-America. 79 participants came from Spain (20 male, 59 female, 3 other/prefer not to say; age range: 21 - 68, mean age: 38.7, SD: 12.8). The number of participants from five countries was too low to ensure reliable results (three or less participants), so we combined their responses with those from neighboring countries (the groupings reflect the traditional dialect zones of Latin-America, as will be explained in Section 3.1). The Latin-American participant groups included data from 43 Andean Spanish speakers (15 male, 28 female; age range: 19 - 52, mean age: 31.3, SD: 7.3), 11 Caribbean Spanish speakers (1 male, 10 female; age range: 20 - 74, mean age: 40.4, SD: 19.6), 24 Chilean Spanish speakers (7 male, 16 female, 1 other/prefer not to say; age range: 23 - 73, mean age: 43.3, SD: 14.0), 17 River Plate Spanish speakers (2 male, 14 female, 1 other/prefer not to say; age range: 16 - 54, mean age: 33.8, SD: 10.5), and 14 Mexico/El Salvador Spanish speakers (7 male, 7 female; age range: 21 - 63, mean age: 31.3, SD: 12.1). Note that the number of participants was not identical for each participant group; the participant groups from the River Plate

[1] https://doi.org/10.34973/5tq7-p857

area (17), Mexico/El Salvador (14), and the Caribbean (11) were considerably smaller than the other participant groups.

69 participants were Dutch L1 speakers, of which 37 participants reported to speak Spanish at elementary or limited level (9 male, 28 female; age range 17 - 71, mean age: 27.9, SD: 14.2) and 32 were fluent speakers of Spanish (12 male, 20 female; age range 20 - 60, mean age: 35.4, SD: 12.8). 54 participants were native speakers of German, of which 27 participants reported to speak Spanish at elementary or limited level (5 male, 21 female, 1 other/prefer not to say; age range 18 - 56, mean age: 30.5, SD: 10.5) and 27 were fluent speakers of Spanish (4 male, 23 female; age range 18 - 66, mean age: 27.9, SD: 10.8).

2.2 Material & design

The main part of the questionnaire consisted of 48 'fill-in-the-blank' sentences, 12 of which were target sentences. All sentences, target and filler, were based on examples found on Twitter, because we wanted them to be as natural and diverse as possible. Twitter typically contains speech acts that are relevant for the topic at hand (statements, claims, opinions), and represents diverse styles of communication in any regional variation. For the sake of clarity, some of the sentences were simplified. Half the target sentences originally contained *creo que* 'I believe that', the other half *pienso que* 'I think that', but both options were acceptable, as judged by three native speakers from different Spanish-speaking countries (Spain, Mexico, and Chile). In the selection of the target sentences, we took into account Mulder's (2018) finding that many examples cannot be unambiguously classified under the heading of one of three categories, *i.e.*, inferred evidentiality, assumed evidentiality, or epistemic modality. Especially the latter category is rather heterogeneous due to the different qualifications of the epistemic commitment marked by *creo que*, varying from probability to uncertainty (cf. Mulder 2018: 110-112). The target sentences include instances of each of the three categories, and can be found in the Appendix. Each sentence was presented with a drop-down menu at the position of the blank, which contained four possible options to select an answer from. In the target sentences, the options were *creo*, *pienso*, and two other options (the complementizer *que* was already given in each sentence). These were *puedo* 'I can', *quiero* 'I want', *necesito* 'I need', or *me gusta* 'I like'. Each option occurred equally often throughout the experiment; the order of the options was randomized per experimental trial.

Fillers were added with the aim of disguising the target sentences. They were also based on examples from Twitter and all contained common first-person verbs. Some of these verbs were grammatically interchangeable, just like the sentences with *creo* and *pienso*. Twelve filler sentences were compound sentences with frequently used first-person cognition verbs: six with *entiendo que* 'I understand that' and six with *sé que* 'I know that'. The remaining 24 fillers were simple clauses, so both compound and simple sentences were equally represented.

The questionnaire was conducted in Qualtrics, where we divided the questions into three blocks of four target and twelve filler sentences. The order of sentences within each block was randomized, as was the order in which the blocks were presented.

2.3 Procedure

Participants were given the option to read the instructions of the questionnaire in English, Spanish, Dutch, or German. After selecting their preferred language, they received information about the questionnaire, and gave their consent by continuing with the questionnaire. They were then asked to fill in their age and gender before continuing to the main part of the questionnaire after a brief description of the experiment. They could select their preferred answer from a drop-down menu, which was initially set to a non-selectable empty option. Before concluding the questionnaire, participants answered questions about which countries they had lived in at various times in their lives, where they lived at the time, and which languages they spoke. The participants were asked to rank their own proficiency on one of three levels: basic or elementary proficiency (only having basic knowledge), conversational or limited proficiency (being able to have a simple conversation), and full proficiency (a fluent speaker). This information was used to categorize participants according to proficiency in Spanish.

2.4 Analysis

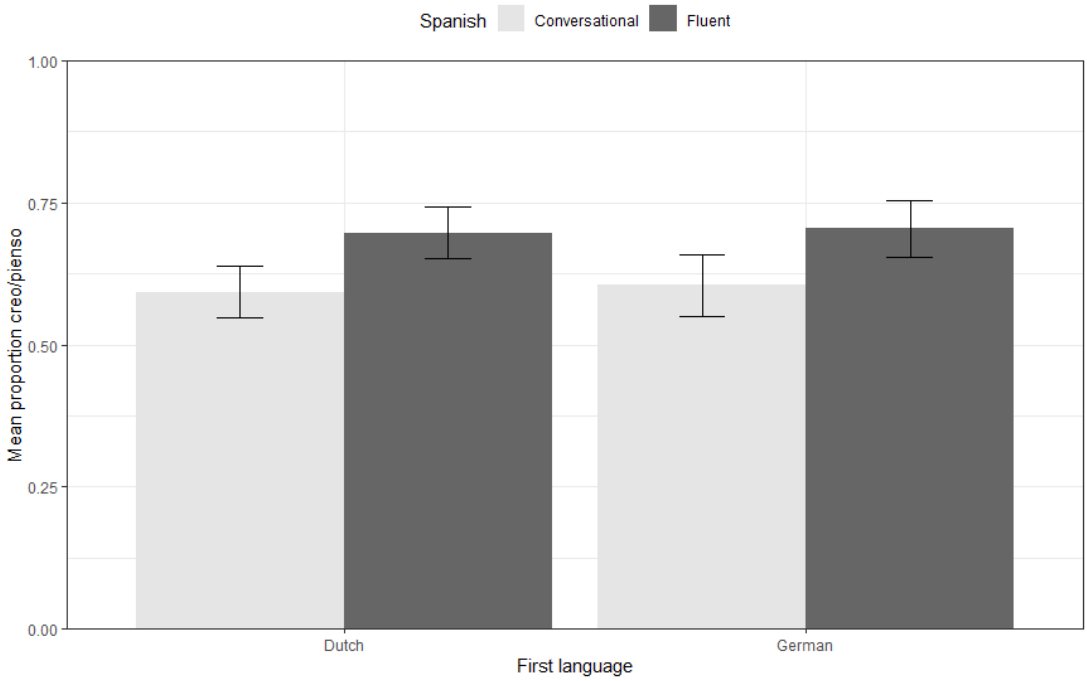
Since only twelve participants judged their Spanish proficiency level as 'basic', their data were combined with those from participants who judged their Spanish proficiency level as 'conversational' (*viz*., being able to have a simple conversation) prior to statistical analysis. We removed responses with answers other than *creo* or *pienso* in order to calculate the *creo que/pienso que* proportions for each participant group.

We first conducted a binomial test to confirm that there is a distinct preference for *creo que* among native speakers of (European) Spanish (cf. Mulder 2018), *i.e.*, whether the observed number of *creo* responses is significantly higher than what could be expected on the basis of chance. This prediction is borne out; the proportion of *creo* responses (775/946 responses, 81.9%) was significantly higher than chance level (50%), $p < .001$ (one-sided).

Next, we ran two statistical models to test whether there are differences in the *creo que/pienso que* ratios of a) Dutch and German learners of Spanish, depending on their proficiency in Spanish, and of b) Spanish speakers from different geographical regions (including Dutch-speaking and German-speaking learners of Spanish).

The mean *creo que/pienso que* proportions by the Dutch and German learners of Spanish are visually presented per participant group in Figure 1. This figure shows that the more advanced learners used *creo que* more often than the less advanced students, apparently regardless of their native language, but even the learners who speak Spanish at best conversationally have a preference for *creo que*.

Figure 1. Mean *creo que/pienso que* proportions of Dutch and German learners of Spanish, per proficiency level, where *pienso que* = 0 and *creo que* = 1.



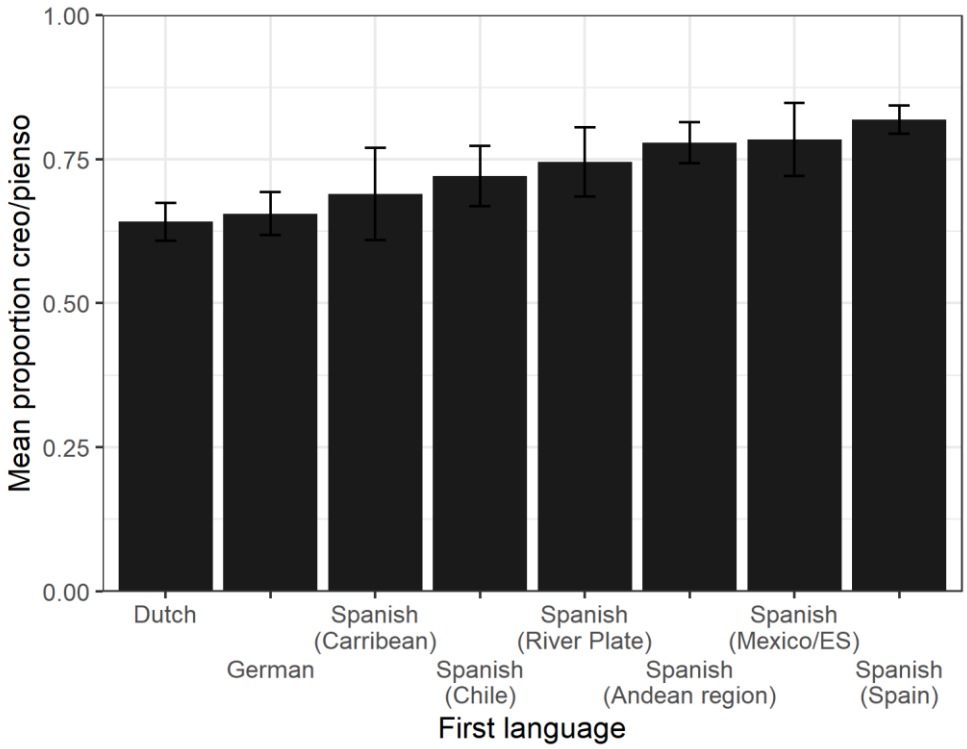
To test the statistical validity of these effects, we performed a generalized linear mixed effects model on the data, using the software R (version 4.0.5) and the *lme4* package (Bates, Mächler, Bolker & Walker 2015), with the *creo que/pienso que* ratio entered as the binary dependent variable (with *pienso que* as the reference category), and *Native language* (Dutch, German) and *Spanish proficiency* (conversational, fluent) as the independent variables. The independent variables were encoded using deviation contrasts (-.5, +.5). The model included a by-item random intercept and a by-item random slope for the effect of *Native language*; a more complex random structure led to singularity issues.

The model yielded a significant main effect of *Spanish proficiency* ($\beta = 0.46$, $SE = 0.11$, $z = 4.06$, $p < .001$), which confirms that the more advanced learners of Spanish chose *creo* over *pienso* more often than the less advanced learners of Spanish, regardless of their mother language. The effect of *Native language* was not significant ($\beta = 0.03$, $SE = 0.15$, $z = 0.18$, $p = .857$), nor was the interaction effect between *Native language* and *Spanish proficiency* ($\beta = -0.02$, $SE = 0.23$, $z = -0.11$, $p = .917$). Thus, the effect of proficiency was not modulated by the participants' native language, nor did we find evidence for a difference between the *creo que/pienso que* proportions of Dutch and German learners of Spanish at large. This finding is unexpected, since there exists a preference for *ik denk* 'I think' in Dutch (de Hoop et al. 2018) and a preference for *(ich) glaub(e)* 'I believe' in German (Schoonjans 2012, Wilton 2019).

We then compared the *creo que/pienso que* ratios of participant groups based on their native language. This comparison includes the data from Dutch-speaking and German-speaking learners of Spanish, as well as the data from native speakers of Spanish, whom we categorized by their country of residence. The *creo que/pienso que* ratios for each group are shown in Figure 2, which shows that the speakers of European

Spanish have the strongest preference for *creo que* in our experiment, and the Dutch and German learners of Spanish the lowest.

Figure 2. Mean *creo que/pienso que* proportions of different groups of Spanish speakers, based on their mother tongue, where *pienso que* = 0 and *creo que* = 1.



We performed a second generalized linear mixed effects model on these data, which predicts the *creo que/pienso que* proportion (with *pienso que* = 0 and *creo que* = 1) as a function of *Participant group* (with European Spanish set as the reference category). This way, we can test whether the mean *creo que/pienso que* proportions per participant group in our experiment are significantly different from that of speakers of European Spanish. The model included a by-item random intercept; a more complex random structure led to singularity issues.

The model indicates that the mean *creo que/pienso que* proportions of native speakers of European Spanish are significantly higher than those of all other participant groups, except for the participant groups from Mexico/El Salvador (β = -0.23, SE = 0.21, z = -1.08, p = .281) and the Andean region (β = -0.26, SE = 0.14, z = -1.89, p = .059). More specifically, we find significant differences between the *creo que/pienso que* proportions of European Spanish and Dutch learners of Spanish (β = -0.96, SE = 0.11, z = -8.48, p < .001), German learners of Spanish (β = -0.90, SE = 0.12, z = -7.43, p < .001), Caribbean Spanish speakers (β = -0.74, SE = 0.21, z = -3.51, p < .001), Chilean Spanish speakers (β = -0.58, SE = 0.16, z = -3.64, p < .001), and River Plate Spanish speakers (β = -0.45, SE = 0.18, z = -2.45, p = .014). Thus, the preference for *creo que* over *pienso que* is not only stronger in speakers of (European) Spanish than in Dutch and German learners of the language, it is also stronger than in speakers of most of the Latin-American varieties of Spanish.

3. Discussion

The aim of our study was to test three hypotheses, namely that (i) native speakers of (European) Spanish prefer the evidential/epistemic marker *creo que* 'I believe that' over *pienso que* 'I think that', (ii) German-speaking learners of Spanish show the same pattern as Spanish natives, and (iii) Dutch-speaking learners of Spanish use relatively more *pienso que* than *creo que*. In addition to the first hypothesis, we wanted to check whether the preference for *creo que* also holds for speakers of the varieties of Spanish in Latin-America.

The results of the experiment confirm the first and third of our hypotheses, but quite surprisingly, the second hypothesis is falsified. We also found a correlation between the level of proficiency and the strength of the preference for *creo que*. Whereas the preference for *creo que* over *pienso que* was the strongest among Spanish speakers, it was the weakest for the beginning Dutch and German-speaking learners of Spanish, with the Dutch and German fluent speakers of Spanish in between. Before we turn to the unexpected result that falsified the second hypothesis in Section 3.3, we will first elaborate on the results that verified the first and third hypotheses in Sections 3.1 and 3.2.

3.1 The distribution of creo que *and* pienso que *in the Spanish-speaking world*

This section elaborates on the result that native speakers of Spanish prefer *creo que* 'I believe that' over *pienso que* 'I think that', but that we do find some dialectical variation. The question we addressed was whether native speakers from other dialect zones than Spain have a similar preference for the evidentiality/epistemic modality marker *creo que* over *pienso que* as Mulder (2018) observed for Peninsular Spanish. While Mulder's study was limited to Peninsular Spanish, he tentatively observed dialectical variation (2018: 105). Therefore, we investigated whether the preference for *creo que* over *pienso que* holds for native speakers from other Spanish dialect zones as well.

Peninsular Spanish indeed has a clear preference for *creo que* 'I believe that', confirming our first hypothesis. What our data also show is that the preference for *creo que* is significantly weaker in almost all Latin American Spanish varieties than in European Spanish (except the Andean region and Mexico/El Salvador).

Since dialectical features in Spanish America rarely coincide with political boundaries (Penny 2000: 136), over the last 150 years varying proposals have been made for divisions into dialect zones, based on different criteria (Lipski 1994, 2013, Moreno Fernández & Ueda 2018: 723). While Lipski has a fine-grained classification of ten zones, based on a combination of different criteria, for the purposes of this paper it is sufficient to follow the basic classification into five regions of Moreno-Fernández & Otero Roth (2008). It can fairly well be applied to the grouping in Figure 2: the Caribbean (of which we have data from Venezuela, Cuba, and the Dominican Republic), Mexico and Central America (Mexico and El Salvador), the Andean region (Ecuador, Peru, and Colombia), River Plate (Argentina and Uruguay), and Chile.[2] It

[2] One caveat is in order here: it is a simplification to present Colombia and Venezuela as uniform dialectal zones. The Spanish of the Caribbean coast of both countries is usually classified as Caribbean Spanish, while the interior and highland areas are part of the Andean region. However, in our experiment most Colombian data came from interior cities like Bogotá and Medellín.

should be pointed out that in these classifications pragmatic data – which are essential for the topic of this paper – are not taken into account. Only recently and incidentally has the socio-pragmatic level been studied in dialect studies of Spanish (e.g., Placencia 2018, Albelda Marco & Briz Gómez 2010). However, the results for the distribution of *creo que* 'I believe that' and *pienso que* 'I think that' are, roughly, in line with the traditionally observed dialectal differences. The zones with the highest preference for *creo que* (Mexico, Andes) are traditionally considered more conservative varieties of American Spanish, as they deviate the least from the standard features of Northern-Central Peninsular Spanish (Moreno Fernández & Otero Roth 2008: 33); on the other hand, we find the highest above-average preference for *pienso que* in dialects of the Caribbean and the Southern Cone (River Plate, Chili), which are considered more innovative varieties.

The less pronounced preference for *creo que* in Caribbean Spanish could be due to the presence and influence of English in this area (Moreno Fernández & Otero Roth 2008: 38). Recall that *I think* is by far the most common complement taking predicate in English (Baumgarten & House 2010), which could explain a possibly higher rate of *pienso que* in the Caribbean.

3.2 Dutch-speaking learners of Spanish

The acquisition of expressions such as *creo que* 'I believe that' and *pienso que* 'I think that' in a foreign language requires the acquisition of their pragmatic functions (Kasper & Rose 2002). Hogeweg et al. (2016) show that the acquisition and use of discourse particles in German is highly influenced by the existence of a similar set of particles in the learner's native language Dutch, even if the exact form-function relations within the sets of particles in the two languages differ. For Dutch learners of Spanish the largely interchangeable expressions *creo que* and *pienso que* may seem identical to the equally interchangeable Dutch expressions *ik geloof dat* 'I believe that' and *ik denk dat* 'I think that', while they may not be aware of the fact that their frequency distributions differ significantly. According to Zhang and Sabet (2016: 334), such frequency differences between L1 and L2 need "not necessarily be labelled as overuse or underuse", as they may also reflect different focuses and preferences among different groups of speakers.

We assumed that Dutch-speaking learners of Spanish would use relatively more *pienso que* 'I think that' than native Spanish speakers. Indeed this turned out to be the case. The preference for *ik denk* 'I think' over *ik geloof* 'I believe' in Dutch could thus explain the 'overuse' of *pienso que* by Dutch learners of Spanish compared to native Spanish speakers in this experiment. Such a pattern is reminiscent of the one found by Liu (2013), where Chinese learners of English use more *I think* in sentence medial or final position to express their opinion than English natives. Liu's (2013) explanation of this pattern is that it is triggered by the Chinese expression *wo juede* 'I think', which has that particular function in those positions.

Since traditionally little attention is paid to teaching pragmatic functions in foreign language education (cf. Aijmer 2011), this may explain why Dutch learners of Spanish choose *pienso que* 'I think that' significantly more often than native Spanish speakers. While any teacher of Spanish in the Netherlands is probably aware of the abundant use of *pienso* by beginning learners with a Dutch background, our results show that even fluent speakers of Spanish use it significantly more often than native

speakers of Spanish. These outcomes may illustrate the transfer of the L1 Dutch preference for *ik denk* 'I think' over *ik geloof* 'I believe' to their Spanish.

3.3 German-speaking learners of Spanish

Section 3.2 above discussed the confirmed hypothesis that Dutch-speaking learners of Spanish, even the advanced learners who fluently speak Spanish, chose *pienso que* 'I think that' over *creo que* 'I believe that' significantly more often than the Spanish natives. On the basis of this outcome we would expect German-speaking learners of Spanish to show the opposite pattern, *i.e.*, to pattern with the native speakers of European Spanish. However, the results of our experiment indicate that the German-speaking learners of Spanish pattern with the Dutch-speaking learners instead, and use *creo que* significantly less often than native speakers of European Spanish. In fact, we did not find evidence for a difference between the Dutch and German learners in the experiment. This surprising result cannot be accounted for as a case of L1 transfer, given the preference for *ich glaube* 'I believe' over *ich denke* 'I think' in German (Schoonjans 2012, Wilton 2019). In Section 4 we will explain this finding as a consequence of interference from the L2 of the German participants, *i.e.*, English, in which *I think* is the preferred phrase, like in Dutch.

3.4 Acquisition of pragmatic markers

The results of our study indicate that even fluent L2-speakers do not use pragmatic markers like *I think* and *I believe* in the same way as native speakers. This is a remarkable result, but the question is whether our participants correctly estimated their own language proficiency level. Unfortunately, in our experiment we did not collect any independent information about the participants' language proficiency level, nor about their learning trajectory and experience with the language. However, Hogeweg et al. (2016), in a similar experiment to ours, had an extra task in addition to the question about the participants' own estimation of language proficiency level, with which they could independently test this. In the experimental task the participants had to choose from four pragmatic discourse markers, while in the filler items, they had to choose from four prepositions. This showed that the participants' own estimation of their level of proficiency matched their performance on choosing the correct preposition, but that the performance on choosing the correct pragmatic marker differed for speakers from different language backgrounds. Compared to L2-learners of German with the same proficiency level but a different native language, Dutch speakers of L2 German performed better on the use of the pragmatic marker in German. Thus, the analysis showed that language background served as a predictor for performance on the pragmatic marker independent of subjects' judgments about their own level of German and their performance on the preposition task. Native speakers of Dutch were better at using *doch* in German, even though their overall proficiency level was no better than that of learners of German with another native language.

Our study raises questions about the acquisition and instruction of pragmatic markers in a foreign language. It would be interesting to see the effects of early acquisition and the role of instruction. In our experiment we did not investigate the age of acquisition of L2, nor the number of years of instruction, but we know from the literature that pragmatic markers are not or hardly dealt with explicitly in L2 acquisition (e.g., Foolen 2010). Hernández (2011) found that in the acquisition of pragmatic markers English learners of Spanish especially benefit from so-called input-

flooding: language forms are offered in functional, communicative activities so that attention is drawn to it in the most natural way possible.

Teaching pragmatics in a foreign language is of great importance, in particular to prevent negative pragmatic transfer, discussed above. As noted by Savvidou and Economidou-Kogetsidis (2019), pragmatic deficiencies can result in misunderstandings that result from the hearer misinterpreting the speaker's intentions and recognizing the force of the speaker's utterance differently than the speaker intended. This makes pragmatic errors in an L2 more serious than grammatical errors, in the sense that such errors are more difficult for hearers to recognize, leading L2 speakers to be judged more harshly when they make these pragmatic errors. That explicit pedagogical interventions can be effective in teaching pragmatics in an L2 is now recognized by many scholars and teachers, but research on it is still relatively small, compared to research on learning vocabulary, phonology, morphology, syntax, and even semantics. Cohen (2020) notes that not only teaching itself, but also the measurement and assessment of pragmatic performance still needs to be developed.

4. English L2 interference in Spanish L3 acquisition

This section explores the idea that German-speaking learners of Spanish are influenced by their L2 English, in which *I think* is the preferred expression, rather than by their L1 in which *I believe* is preferred. However, we first need to back up our claims concerning the differences in distribution between the two phrases in Spanish, Dutch, German, and English. If there is a preference for *ich denke* 'I think' in German, then the explanation for the German data might just be the same as for the Dutch data.

In order to find further evidence for our assumptions on the frequency distributions between *I think that* and *I believe that* in the four languages, we conducted a Twitter corpus study. The corpus contains tweets from April 24, 2021, which was a randomly chosen date. We collected all tweets that contained the expressions *creo que* 'I believe that', *pienso que* 'I think that', and their translations in German, English, and Dutch. In German, we only collected *ich glaube dass* 'I believe that' and *ich denke dass* 'I think that' (with and without a comma between the verb and the complementizer), but not other spelling variants, of which German has many (Schoonjans 2012, Wilton 2019). This might explain why we found less tweets in German than in Dutch within the same time interval of 24 hours. We did not examine the interpretations of the two expressions *I think* and *I believe* in the data, as we were only interested in the raw frequency patterns. Due to the large numbers of tweets we obtained with the target expressions in English and Spanish, we restricted the time interval for these two languages, as indicated in Table 1 below.

Table 1: Numbers and percentages of *creo que* and *pienso que* and its counterparts in German, Dutch, and English in a Twitter corpus, with the highest percentage per language in bold.

	Spanish	German	Dutch	English
time interval	24-04-2021 23:14-23:59	24-04-2021 24 hours	24-04-2021 24 hours	24-04-2021 21:49-23:59
'I believe that' N	3463	189	76	310
%	**60.5%**	**53.1%**	5.6%	14.9%
'I think that' N	2265	167	1290	1776
%	39.5%	46.9%	**94.4%**	**85.1%**

The results are more or less in line with what we expected, although German shows a weaker preference for 'I believe that' than we had predicted based on the literature (Schoonjans 2012, Wilton 2019). Yet, the difference in frequency distribution between German and Dutch is enormous, given the percentage 'I believe that' is 53.1% in German, while it is only 5.6% in Dutch. Based on this, we would have expected that the German learners of Spanish in the experiment would choose *creo que* more often than the Dutch learners. However, the German learners did not differ significantly from the Dutch learners, while they did differ significantly from the Spanish speakers in the experiment.

Like the Dutch-speaking learners of Spanish, the German-speaking learners used *creo que* 'I believe that' significantly less often than Spanish native speakers. The fact that the results did not reveal statistically relevant differences between the German and Dutch-speaking learners cannot be accounted for as transfer from their L1s, since the distribution pattern in Dutch is completely different from that in German (see Table 1). Therefore, we put forward the explanation that at least the German performance (but perhaps also the Dutch) in Spanish must be due to the influence of their L2, which is English.

English speakers are known to use *I think* extensively (Simon-Vandenbergen 2000, Baumgarten & House 2010). The results of our Twitter corpus study presented above also show about 85% *I think that* versus 15% *I believe that* in English. Moreover, speakers of English as an L2 or lingua franca have been shown to use *I think* frequently as well. Zhang and Sabet (2016) study the use of *I think* and *I believe* as well as some other related phrases by L1 and L2 speakers of English in a classroom setting in three different countries (USA, China, and Iran). For all three groups, *I think* is by far the most used (79% for the native speakers of English, 99% and 73% for the Chinese and Persian speakers of English), and *I believe* the least (4%, 0%, 3%, respectively). The data also show that L2 speakers use more *I think* than the L1 speakers: the Chinese-speaking learners of English use it 4.5 times more, and the Persian-speaking learners 3.5 times more than the native English speakers.

Wilton (2019) investigates two data sets, one set containing interviews in German with German professional soccer players. The other set contains interviews with soccer players of different L1s, including German. These interviews are held in English as a lingua franca. Whereas *ich glaub(e)* 'I believe' is used more frequently than *ich denk(e)* 'I think' in the German dataset, in the English dataset *I believe* is not used at all, whereas *I think* makes up 92% of all occurrences of a first person pronoun plus a verb of cognition (Willet 2019: 225). We can conclude from this study that in identical circumstances, namely a professional interview situation, German soccer

players use *I think* but not *I believe* when they speak English, although they prefer *ich glaub(e)* over *ich denk(e)* when they speak German.

The influence of English as an L2 on the acquisition of an L3 is well-established (see e.g., Stadt et al. 2018, Westergaard 2021). Hammarberg (2001) points out that there is in fact a general tendency to activate an earlier L2 rather than the L1 in L3 performance. This is also suggested in Bohnacker (2006), who finds that word order in L2 English hinders L1 Swedish learners from acquiring V2 word order in L3 German, even though Swedish has the same V2 word order as German.

That English as L2 for German-speaking learners can interfere with Spanish L3, has been put forward by Eibensteiner (2019), who shows that aspectual knowledge of L2 English affects the acquisition of past tenses in L3 Spanish. He argues that German-speaking learners of Spanish L3 activate their L2 English as a default transfer source in this domain.

The present study adds further evidence to the assumption that English as a second language interferes with the acquisition of Spanish by native speakers of German. In our study it is not the acquisition of morphology, syntax, or semantics, but of pragmatics where we see this effect. Because English is also the L2 for Dutch-speaking learners of Spanish (Stadt et al. 2018), we cannot exclude the possibility that the Dutch data are also the result of L2 influence rather than L1 influence, or both (Bohnacker 2021). Stadt et al. (2018) argue that the influence of English L2 on the acquisition of L3 is dependent on the degree of exposure to the L2, rather than on L2 proficiency. We assume that the influence of English on performance in Spanish is in any case no less for Dutch-speaking than for German-speaking learners, because of the high exposure to English in the daily life of the Dutch-speaking population.

5. Conclusion

We experimentally investigated the use of the two evidential/epistemic expressions *creo que* ‘I believe that’ and *pienso que* ‘I think that’ among native speakers of various varieties of Spanish as well as Dutch and German-speaking learners of Spanish. In accordance with earlier findings, *creo que* was found to be preferred over *pienso que* by native speakers of Spanish. Our data also showed that the preference for *creo que* is the strongest in European Spanish, and weaker in almost all varieties of Spanish in Latin-America.

Both Dutch and German-speaking learners of Spanish were found to use relatively more *pienso que* ‘I think that’ than Spanish speakers. We found a correlation between the level of proficiency and the strength of the preference for *creo que* ‘I believe that’. Whereas the preference for *creo que* over *pienso que* was the strongest among Spanish speakers, it was the weakest for the least advanced Dutch and German learners of Spanish, with the Dutch and German speakers who were fluent in Spanish in between.

For the Dutch-speaking learners the relatively more frequent use of *pienso que* ‘I think that’ can straightforwardly be explained by their first language, which has a clear preference for *ik denk* ‘I think’ over *ik geloof* ‘I believe’. For the German-speaking learners, we have argued that it is not their mother tongue, but English as their second language, that causes the decreased use of *creo que* compared to the Spanish natives. We conclude that the interference of the second language English can

explain the performance of the German-speaking learners of Spanish as a third language, while for the performance of the Dutch-speaking learners we cannot tell whether it is the influence of their first or second language, or both.

Acknowledgments

We are most grateful to Danai Psathas for designing the experiment as part of her Honours project within the Radboud Honours Academy, which is also gratefully acknowledged. Furthermore, our thanks go to Maria van de Groep and Michelle Suijkerbuijk for collecting the experimental data reported in Section 2, and to Wessel Stoop for collecting the Twitter data presented in Section 4. We are grateful to the reviewers and editors for their constructive comments on an earlier version of this article.

References

Aijmer, Karin. 1997. *I think* – an English modal particle. In T. Swan, & O. Jansen Westvik (eds.), *Modality in Germanic languages. Historical and comparative perspectives*, 1–47. Berlin/New York: Mouton de Gruyter. https://doi.org/10.1515/9783110889932.1

Aijmer, Karin. 2011. *Well I'm not sure I think…* The use of *well* by non-native speakers. *International Journal of Corpus Linguistics* 16(2): 231–254. https://doi.org/10.1075/ijcl.16.2.04aij

Aijmer, Karin, & Anne-Marie Simon-Vandenbergen. 2003. The discourse particle *well* and its equivalents in Swedish and Dutch. *Linguistics* 41(6): 1123–1161. https://doi.org/10.1515/ling.2003.036

Aijmer, Karin, & Anne-Marie Simon-Vandenbergen. 2006. *Pragmatic markers in contrast*. Amsterdam: Elsevier. https://brill.com/view/title/23247

Albelda Marco, Marta, & Antonio Briz Gómez. 2010. Aspectos pragmáticos: cortesía y atenuantes verbales en las dos orillas a través de muestras orales. In A. Izquierdo, & J. M. Enguita Utrilla (eds.), *La lengua española en América: Normas y usos actuales*, 237–260. Universidad de Valencia.

Ament, Jennifer, Júlia Barón Páres, & Carmen Pérez-Vidal. 2020. A study on the functional uses of textual pragmatic markers by native speakers and English-medium instruction learners. *Journal of Pragmatics* 156: 41–53. https://doi.org/10.1016/j.pragma.2019.09.009

Anderson, Lloyd B. 1986. Evidentials, paths of change, and mental maps: Typologically regular asymmetries. In W. Chafe, & J. Nichols (eds.), *Evidentiality: The linguistic coding of epistemology,* 273–312. Norwood, NJ: Ablex.

Bates, Douglas, Martin Mächler, Ben Bolker, & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1): 1–48. https://doi.org/10.18637/jss.v067.i01

Baumgarten, Nicole, & Juliane House. 2010. *I think* and *I don't know* in English as a lingua franca and native English discourse. *Journal of Pragmatics* 42(5): 1184–1200. https://doi.org/10.1016/j.pragma.2009.09.018

Bohnacker, Ute. 2006. When Swedes begin to learn German: From V2 to V2. *Second Language Research* 22(4): 443–486. https://doi.org/10.1191/0267658306sr275oa

Bohnacker, Ute. 2021. On transfer and third language acquisition: A commentary on Schwartz & Sprouse. *Linguistic Approaches to Bilingualism* 11(1): 37–44. https://doi.org/10.1075/lab.20101.boh

Boye, Kasper, & Peter Harder. 2009. Evidentiality: Linguistic categories and grammaticalization. *Functions of Language* 16(1): 9–43. https://doi.org/10.1075/fol.16.1.03boy

Brinton, Laurel. 1996. *Pragmatic markers in English: Grammaticalization and discourse functions*. Berlin/New York: Mouton de Gruyter. https://dx.doi.org/10.1515/9783110907582

Cohen, Andrew D. 2020. Considerations in assessing pragmatic appropriateness in spoken language. *Language Teaching* 53: 183–202. https://doi.org/10.1017/S0261444819000156

Davies, Mark. 2002. *Corpus del español: 100 million words, 1200s-1900s*. http://www.corpusdelespanol.org

Dehé, Nicole, & Anne Wichmann. 2010. Sentence-initial *I think (that)* and *I believe (that)*: Prosodic evidence for use as main clause, comment clause and discourse marker. *Studies in Language* 34(1): 36–74. https://doi.org/10.1075/sl.34.1.02deh

Eibensteiner, Lukas. 2019. Transfer in L3 acquisition. How does L2 aspectual knowledge in English influence the acquisition of perfective and imperfective aspect in L3 Spanish among German-speaking learners? *Dutch Journal of Applied Linguistics* 8(1): 67–83. https://doi.org/10.1075/dujal.19003.eib

Félix-Brasdefer, J. César & Erin Lavin. 2009. Grammar, prosody, and turn expansion in second language conversations. In J. Collentine, M. García, B. Lafford, & F. Marcos Marín (eds.), *Selected proceedings of the 11th Hispanic Linguistics Symposium*, 53–67. Somerville, MA: Cascadilla Proceedings Project.

Fraser, Bruce. 1999. What are discourse markers? *Journal of Pragmatics* 31(7): 931–952. https://doi.org/10.1016/S0378-2166(98)00101-5

Foolen, Ad. 2010. Partikels volgens de schijf van vijf [Particles according to the 'disc of five]. *Internationale Neerlandistiek* 48(2): 41–51.

Foolen, Ad. 2011. Pragmatic markers in a sociopragmatic perspective. In G. Andersen, & K. Aijmer (eds.), *Pragmatics of society*, 217–242. Berlin/New York: Mouton de Gruyter. https://doi.org/10.1515/9783110214420.217

Foolen, Ad, Helen de Hoop, & Gijs Mulder. 2018. Evidentiality: How do you know? In A. Foolen, H. de Hoop, & G. Mulder (eds.), *Empirical evidence for evidentiality*, 1–16. Amsterdam/Philadelphia: John Benjamins.
https://doi.org/10.1075/hcp.61.01foo

Griffioen, Laura, Helen de Hoop, & Gijs Mulder. 2018. Over het verschil in evidentialiteit tussen *denk ik* en *dacht ik* [About the difference in evidentiality between *I think* and *I thought*]. *Internationale Neerlandistiek* 56(2): 121–140.
https://doi.org/10.5117/IN2018.2.002.GRIF

Hammarberg, Björn. 2001. Roles of L1 and L2 in L3 production and acquisition. In J. Cenoz, B. Hufeisen, & U. Jessner (eds.), *Cross-linguistic influence in third language acquisition: Psycholinguistic perspectives*, 21–41. Clevedon: Multilingual Matters. https://doi.org/10.21832/9781853595509-003

Haselow, Alexander. 2021. The acquisition of pragmatic markers in the foreign language classroom: An experimental study on the effects of implicit and explicit learning. *Journal of Pragmatics*, 186: 73–86.
https://doi.org/10.1016/j.pragma.2021.09.017

Hernández, Todd A. 2011. Re-examining the role of explicit instruction and input flood on the acquisition of Spanish discourse markers. *Language Teaching Research* 15(2): 159–182. https://doi.org/10.1177/1362168810388694

Hogeweg, Lotte, Helen de Hoop, Stefanie Ramachers, Frans van der Slik, & Verena Wottrich. 2016. The L2-acquisition of the German particle *doch*. *International Review of Applied Linguistics in Language Teaching* 54(3): 201–227.
https://doi.org/10.1515/iral-2016-9981

de Hoop, Helen, Ad Foolen, Gijs Mulder, & Vera van Mulken. 2018. *I think* and *I believe*: Evidential expressions in Dutch. In A. Foolen, H. de Hoop, & G. Mulder (eds.), *Empirical evidence for evidentiality*, 78–97. Amsterdam/Philadelphia: John Benjamins. https://doi.org/10.1075/hcp.61.04hoo

Kasper, Gabriele. 1992. Pragmatic transfer. *Second Language Research* 8(3): 203–231. https://doi.org/10.1177/026765839200800303

Kasper, Gabriele. 2010. Interlanguage pragmatics. In M. Fried, J.-O. Östman and J. Verschueren (eds.), *Variation and Change: Pragmatic perspectives*, 141–154. https://doi.org/10.1075/hoph.6.11kas.

Kasper, Gabriele, & Kenneth R. Rose. 2002. Theories of second language pragmatic development. *Language Learning* 52(1): 13–62. https://doi.org/10.1111/j.1467-1770.2002.tb00023.x

Kearns, Kate. 2007. Epistemic verbs and zero complementizer. *English Language and Linguistics* 11(3): 475–505. https://doi.org/10.1017/S1360674307002353

Lipski, John M. 2012. Geographical and social varieties of Spanish: An overview. In J.I. Hualde, A. Olarrea, & E. O'Rourke (eds.), *The handbook of hispanic linguistics*, 1–26. Chichester: Wiley-Blackwell. https://doi.org/10.1002/9781118228098.ch1

Lipski, John M. 1994. *Latin American Spanish*. London/New York: Longman.

Liu, Binmei. 2013. Effect of first language on the use of English discourse markers by L1 Chinese speakers of English. *Journal of Pragmatics* 45(1): 149–172. https://doi.org/10.1016/j.pragma.2012.11.002

Moreno Fernández, Francisco, & Jaime Otero Roth. 2008. *Atlas de la lengua española en el mundo*. Fundación Telefónica and Ariel.

Mulder, Gijs. 2017. When feeling is thinking: A lexical-semantic analysis of evidential and epistemic predicates in Spanish. In J. I. Marín Arrese, G. Haßler, & M. Carretero (eds.), *Evidentiality revisited: Cognitive grammar, functional and discourse-pragmatic perspectives*, 105–121. Amsterdam/Philadelphia: John Benjamins. https://doi.org/10.1075/pbns.271.05mul

Mulder, Gijs. 2018. *(Yo) creo que* as a marker of evidentiality and epistemic modality: Evidence from Twitter. In A. Foolen, H. de Hoop, & G. Mulder (eds.), *Empirical evidence for evidentiality*, 100–120. Amsterdam/Philadelphia: John Benjamins. https://doi.org/10.1075/hcp.61.05mul

Müller, Simone. 2005. *Discourse markers in native and non-native English discourse*. Amsterdam/Philadelphia: John Benjamins. https://doi.org/10.1075/pbns.138

Ortega y Gasset, José. 1940. *Ideas y creencias*. Buenos Aires: Espasa-Calpe.

Pearson, Lynn, & Maria Hasler-Barker. 2021. Second language acquisition of Spanish pragmatics. In D. A. Koike, & J. C. Félix-Brasdefer (eds.), *The Routledge handbook of Spanish pragmatics*, 423–439. New York: Routledge. https://doi.org/10.4324/9780429455643-31

Penny, Ralph. 2000. *Variation and change in Spanish.* Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139164566

Placencia, María E. 2018. Variación pragmática. In J. Gutiérrez-Rexach (ed.), *Enciclopedia de lingüística hispánica*, 797–809. New York: Routledge. https://doi.org/10.4324/9781315713441

Polat, Britanny. 2011. Investigating acquisition of discourse markers through a developmental learner corpus. *Journal of Pragmatics* 43(15): 3745–3756. https://doi.org/10.1016/j.pragma.2011.09.009

Romero-Trillo, Jesús. 2019. Prosodic pragmatics and feedback in intercultural communication. *Journal of Pragmatics* 151: 91–102. https://doi.org/10.1016/j.pragma.2019.02.018

San Roque, Lila, & Hendrik Bergvist. 2015. Epistemic marking in typological perspective. Special issue of *STUF - Language typology and universals* 68(2): 123–232. https://doi.org/10.1515/stuf-2015-0007

Savvidou, Christine, & Maria Economidou-Kogetsidis. Teaching pragmatics: Nonnative-speaker teachers' knowledge, beliefs and reported practices. *Intercultural Communication Education* 2(1): 39–58. https://dx.doi.org/10.29140/ice.v2n1.124

Schoonjans, Steven. 2012. The particulization of German complement-taking mental predicates. *Journal of Pragmatics* 44(6–7): 776–797. https://doi.org/10.1016/j.pragma.2012.02.011

Simon-Vandenbergen, Anne-Marie. 2000. The functions of *I think* in political discourse. *International Journal of Applied Linguistics* 10(1): 41–63. https://doi.org/10.1111/j.1473-4192.2000.tb00139.x

Spronck, Stef. 2015. Stance as participant structure: A Jakobsonian approach to the pragmatics and semantics of evidentiality. *Belgian Journal of Linguistics* 29(1): 193–216. https://doi.org/10.1075/bjl.29.09spr

Stadt, Rosalinde, Aafke Hulk, & Petra Sleeman. 2018. The influence of L2 English on L3 French acquisition in bilingual education. *Dutch Journal of Applied Linguistics* 7(2): 227–245. https://doi.org/10.1075/dujal.18006.sta

Staples, Shelley, & Julieta Fernández. 2019. Corpus linguistics approaches to L2 pragmatics research. In N. Taguchi (ed.), *The Routledge handbook of second language acquisition and pragmatics*, 241–254. London/New York: Routledge. https://doi.org/10.4324/9781351164085-16

Thompson, Sandra A. & Anthony Mulac. 1991. A quantitative perspective on the grammaticization of epistemic parentheticals in English. In E. Closs Traugott, & B. Heine (eds.), *Approaches to grammaticalization Vol. 2*, 313–329. Amsterdam/ Philadelphia: John Benjamins. https://doi.org/10.1075/tsl.19.2.16tho

Traugott, Elizabeth. 1995. Subjectification in grammaticalisation. In D. Stein, & S. Wright (eds.), *Subjectivity and subjectivisation: Linguistic perspectives*, 31–54. Cambridge: Cambridge University Press. https://doi.org/10.1017/cbo9780511554469.003

Westergaard, Marit. 2021. Microvariation in multilingual situations: The importance of property-by-property acquisition. *Second Language Research* 37(3): 379–407. https://doi.org/10.1177/0267658319884116

Wiemer, Björn. 2018. Evidentials and epistemic modality. In A. Y. Aikhenvald (ed.), *The Oxford handbook of evidentiality*, 85–108. Oxford: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198759515.013.4

Wilton, Antje. 2019. Epistemic stance markers in German and English as a lingua franca media sports interviews. *Alicante Journal of English Studies* 32: 217–243. https://doi.org/10.14198/raei.2019.32.09

Zhang, Grace Q. & Peyman G.P. Sabet. 2016. Elastic 'I think': Stretching over L1 and L2. *Applied Linguistics* 37(3): 334–353. https://doi.org/10.1093/applin/amu020

Appendix

The twelve target-sentences used for the experiment are listed below. Between square brackets the forms that were used in the original examples are given.

- *[Creo que] va a llover, porque hay muchas nubes negras en el cielo.*
  'I believe it's going to rain, because there are dark clouds in the sky.'
- *Estoy de viaje y [creo que] el internet está muy lento porque veo noticias de hace 18 años.*
  'I'm travelling and I think the internet is very slow because I see news from 18 years ago.'
- *Antes estaba gorda, [pienso que] he adelgazado, pero no es verdad.*
  'I used to be fat, I think I've lost weight, but it's not true.'
- *Tengo dolor de cabeza y estoy cansada, [pienso que] puedo tener la gripe.*
  'I have a headache and I'm tired. I think I may have the flu.'
- *Quería ir pero no hay más tickets. [Pienso que] el festival es muy popular porque es único en los Estados Unidos.*
  'I wanted to go but there are no tickets anymore. I believe that the festival is so popular because it is unique in the United States.'
- *[Pienso que] ya estoy lista para casarme. Aprendí a hervir agua.*
  'I think I'm ready now to get married. I learned how to boil water.'
- *Soy la persona más fría y, al mismo tiempo, la persona más sensible del mundo. [Creo que] estoy demente*.
  'I am the coldest person and, at the same time, the most sensitive person in the world. I think I'm insane.'
- *[Creo que] estoy haciendo las cosas mal porque nadie siente cariño por mí. No sé si reír o llorar.*
  'I think I'm doing things wrong because no one cares about me. I don't know whether to laugh or cry.'
- *¿Juega Messi en el partido de hoy? - Ni idea, [creo que] está lesionado.*
  'Is Messi playing?' 'I don't know, I believe he's injured.'
- *Sobre Eurovision, [pienso que] va a ganar Francia*

'About Eurovision, I think that France will win'

- *[Creo que] a veces soy demasiado directa con algunas personas y normalmente no me hace ningún bien.*
  'I think that sometimes I'm too direct with some people and usually it doesn't do me any good.'
- *Le envié un mensaje anónimo a mi crush pero [pienso que] ella sabrá quién fue.*
  'I sent an anonymous message to my crush but I think she'll know who it was.'