5th International Conference on Corpus Linguistics (CILC2013)

The Use of Discourse Markers in Spanish Language Learners’ Written Compositions

An Vande Casteele*, Kim Collewaert

Vrije Universiteit Brussel, Faculteit Letteren & Wijsbegeerte, Pleinlaan 2, 1050 Brussel, Belgium

Abstract

The present paper reports on the use of discourse markers in short compositions written by Spanish language learners. The aim is to explore how frequently discourse markers appear in written productions of foreign language learners and to verify which kind of markers they mostly use. A learners’ corpus of approximately 26 500 words has been compiled for this study. By comparing the results of the language learners with the texts composed by a control group of Spanish native participants, we want to shed light on which types of discourse markers seem more or less evident to acquire and which problems learners seem to encounter in expressing discourse cohesion.

© 2013 The Authors. Published by Elsevier Ltd. Open access under CC BY-NC-ND license. Selection and peer-review under responsibility of CILC2013.

Keywords: corpus analysis; discourse cohesion; discourse markers; Spanish as a foreign language; written discourse

1. Introduction

The starting point for our research on discourse markers in Spanish as a foreign language is a study presented by Martí Sánchez (2008) and more precisely his statement that Spanish language learners seem to use discourse markers less frequently and in a less secure way than native speakers. Now, with our corpus-based analysis we attempt to verify this hypothesis. Therefore a contrastive investigation will be conducted on the basis of short narrative texts written by Spanish language learners. Their results will be compared with similar texts composed by a control group of Spanish native participants.

* Corresponding author. Tel.: +32-2-629-26-51; fax: +32-2-629-36-84.
E-mail address: An.Vande.Casteele@vub.ac.be
As discourse markers are said to be crucial resources in guiding the communication process, their acquisition will also play an important role in acquiring communicative competence. Communicative competence implies the ability to act and interact with language. As Widdowson (1996) argues, knowing a language requires more than understanding, speaking, reading and writing sentences, it also includes knowledge on how sentences are related and used to produce communicative effect. Hymes (1972: 277) stated already that communicative competence is knowing “when to speak, when not to, what to talk about and with whom, when, and in what manner to interact.” Therefore, in order to obtain effective communication, small words such as discourse markers cannot be neglected, Svartvik (1980: 171) highlights: “If a foreign language learner says “five sheep” or “he goed”, he can be corrected by practically every native speaker. If, on the other hand, he omits a well, the likely reaction will be that he is dogmatic, impolite, boring, awkward to talk to etc., but a native speaker cannot pinpoint an ‘error’.”

Finally, several studies have already been conducted on the use of discourse markers by non-native speakers. But still a lot has to be investigated. Most of these studies examine learners of English as a foreign language (e.g. Lahuerta Martínez, 2002, 2004 and Müller, 2005). Our study was conducted at the Vrije Universiteit Brussel in Belgium with Flemish learners of Spanish as a foreign language who approximately have a B1-B2 proficiency-level according to the Common European Framework. The mother tongue of our learners’ group is Dutch, but they all acquired other foreign languages as French and English. In the control group all participants were Spanish native speakers. Another interesting aspect in our study is the fact that the written assignment is the same both for the learners as for the control group and that the participants were not aware of our research topic.

2. Some initial thoughts on discourse markers

Discourse markers belong to a category of linguistic devices which seem to play an important role in interpreting texts. Despite their significant pragmatic role in discourse, they are hard to define (Fischer, 2006; Loureda Lamas and Acín Villa, 2010; Martín Zorraquino and Portolés Lázaro, 1999; Martín Zorraquino and Montolío Durán, 1998; Portolés Lázaro, 2007; Schiffrin, 1987).

From a semantic perspective, discourse markers are cohesive devices with a procedural role. They do not have an informational content, but they do contribute to it by establishing links between informational elements. According to Briz, Pons Bordería and Portolés (2008) four major semantic functions can be distinguished: discourse markers can be used as connective devices with an argumentative, reformulating or organizing function, they can operate as discourse modalizers, focalize an element in a discourse or serve to maintain contact and control in communication. To illustrate the high variety in discourse functions assumed by markers, Jucker and Ziv (1998: 1), for instance, enumerate several possible roles: “connectors, turn-takers, confirmation-seekers, intimacy signals, topic-switchers, hesitation markers, boundary markers, fillers, prompters, repair markers, attitude markers, and hedging devices.”

From a syntactic perspective, discourse markers are defined as extrapredicative elements. This means that they are syntactically independent segments in the sentence. Consequently they could be easily removed maintaining the utterance grammatically intact. This, however, does not imply that both versions, with or without the marker, express exactly the same meaning.

Furthermore, discourse markers tend to present a high degree of grammaticalization. This so-called high degree of grammaticalization is characterised by a desemantization and a pragmatic strengthening (Lehmann 1995 and Traugott 2003). So, the grammaticalization process implies a change whereby lexical items develop a more grammatical role and lose their semantic content. It is a discourse-oriented kind of change, as Günthner and Mutz (2004) have pointed out.

The key discourse role of markers is to “anchor the communication to the speaker’s attitudes towards aspects of the on-going interaction” (Östman 1981: 5). By doing so, they help making utterances as relevant as possible through the least amount of processing time (Andersen 2001).
3. Design of the corpus study

As stated before, the aim of our exploratory study on the use of discourse markers by language learners is to determine to what extent learners use similar markers as native speakers and in which aspects important differences can be observed. For the compilation of the corpus both native and non-native speakers were assigned a written composition, a story-telling task, titled *Nunca olvidaré el día en que...*, with an approximate length of 600 words. In order to avoid a focus on the use of discourse markers, participants were not informed of our research purposes. The Spanish language learners, for instance, wrote their assignments as part of a grammar course and thought it was an exercise on grammatical accuracy and more precisely on the use of verbal tenses. The fact that both groups received the same task without knowing our research goals, allows us to work with a homogeneous dataset. To illustrate this, here is a fragment of a text of one of our students:

1) “El verano pasado, he ido a Madrid con mi novio. Era mi primera vez en España y ¡era unas vacaciones maravillosas! Pero el fin de nuestro viaje no era tan bueno. Llegamos al aeropuerto demasiado temprano para estar seguro de no perder nuestro vuelo. Cuando estábamos en la gran sala de salidas, veíamos en los informaciones que nuestro vuelo era cancelado debido a una huelga en Francia. ¡Panicé! Fuimos al taquilla de Ryanair, esperando que sería posible darnos otro vuelo. Pero la gente de Ryanair nos decía que la primera opción era tres días más tarde y como ya habíamos sido en Madrid seis días, queríamos volver a casa. ¿Qué teníamos hacer entonces? Iban varios opciones: el tren, pero no eran places libres porque todo el mundo quería volver a casa. La otra opción era buscar otro vuelo. Habíamos buscado muchísimo y en fin hemos reservado el vuelo más barato que pudimos encontrar. Sin embargo, este vuelo aun era bastante costoso para nosotros, pobres estudiantes. Pero la mala suerte ya no había finido: nuestro nuevo vuelo tenía más de dos horas de retraso. Al fin y al cabo, hemos pasado más de once horas en el aeropuerto de Madrid. Cuando llegamos por fin en Amsterdam, y éramos buscado para el hermano de mi novio, llegamos a la casa a las una. ¡Cómo estaba contenta de ver mi cama!"

For the classification of the discourse markers our analysis was mainly based upon the proposal detailed by Martín Zorraquino and Portolés Lázaro (1999) in the Spanish reference grammar *Gramática Descriptiva de la Lengua Española* published by Bosque and Demonte. Some refinements were added by taking into account the analysis of Briz, Pons Bordería and Portolés (2008) and Martí Sánchez (2008). For instance, *por eso* and *por ello*, which are not incorporated in the classification scheme of Martín Zorraquino and Portolés Lázaro (1999), given the fact that they do not seem to fulfill entirely the grammaticalization parameter, are also considered in our study, based upon their appearance in the work of Martí Sánchez (2008) and in the online specialised dictionary on discourse particles *Diccionario de partículas discursivas del español* (Briz, Pons Bordería and Portolés, 2008). Furthermore, the conjunction *pero* was also included, since it can be considered as the prototype of argumentative connectors. *Pero* is not only the most frequently used connective device in counter-argumentation, it also presents a similar use as other argumentative discourse markers, such as *no obstante* or *sin embargo* (Fernández Silva, 2005), which clearly function as discourse markers. In the next section the corpus data and the results will be further discussed.

4. General results

First, some general results will be presented on how frequently discourse markers appear in the written narratives of the Spanish language learners and the texts of the native speakers. The database consists of 51 texts written by Spanish language learners (26 547 words), and, next to this, it also includes 10 texts (7511 words) written by native speakers (the control group). In the 51 texts the language learners produced, they use 458 discourse markers. Observe that the number of markers used per assignment varies from 0 to 25. This means an average of 8.98

---

1 Note that errors have not been corrected in the samples.
discourse markers per text. The native speakers, on the other hand, employ 94 discourse markers in 10 assignments, which leads to an average of 9.4 discourse markers per text. So, our results show that native speakers use more markers than language learners, but we have to admit that there is only a slight difference. Moreover, an important aspect we have not taken into account as far is the length of the productions. Some of the narratives were a lot shorter than others (around 200 words). If only the assignments which meet the required extension (approximately 600 words) are taken into account, the mean number of discourse markers used by the group of Spanish language learners increases to 10.25 per text. Consequently, in this case, Spanish language learners would use more discourse markers than the control group. In order to give a more correct insight on the data, the average number of discourse markers per 1000 words was calculated. Therefore the number of discourse markers in each text were counted and the total number of words. The results show now that the group of Spanish language learners on average use 17.25 discourse markers per 1000 words, whereas the native speakers use 12.5 discourse markers per 1000 words. So, the frequency of the selected discourse markers seems to be higher in the assignments written by non-native speakers.

At first, these figures are contrary to the general assumption that non-natives would use less markers than native speakers. Nevertheless a more detailed analysis on the kinds of discourse markers used should enable us to refine our findings.

Next to the necessity of more details on the discourse markers used, we also have to account for possible non-linguistic parameters which could influence the use of discourse markers. For instance, the context in which the task was produced can also have its impact. The more academic setting in which the non-natives produced their assignments, including a teacher correcting such assignments, could have caused that the Spanish language learners paid more attention to the use of discourse markers. The native speakers, on the contrary, did not produce their assignments in an institutionalized context.

5. Distribution of discourse markers in learner’s and native’s productions

The next section provides details on the distribution of the discourse markers used in the narrative texts produced by language learners versus the ones written by native speakers. First, we examined which categories of discourse markers appear more frequently in the texts. The analysis clearly shows that both the learners as the native speakers use the connecting devices most frequently. And within this category of connective discourse markers, the contrastive type prevails, thanks to the high occurrence of the marker *pero*, a rather simple marker well-known by language learners. So the high frequency of *pero* seems rather obvious. Against our expectations, more complex contrastive devices also appear in the assignments of the language learners. In their texts we also found *no obstante*, *en cambio* or *al contrario*. So there is variation in the choice of contrastive markers, even more than in the texts produced by the control group. Whereas natives only use *pero* and *sin embargo*, in the productions of the non-natives we also found *no obstante*, *en cambio* and *al contrario*. This high level of variation can also be explained by the academic setting in which the Spanish language learners produced their assignments. Nevertheless, our analysis shows also some problems in the use of markers by the language learners. They seem to hesitate more in choosing the right contrastive marker: often they use *pero* while *no obstante* or *sin embargo* would be more appropriate, or vice versa. And *pero* is clearly overused by the language learners, as illustrated by the following example and also by the example (1).

2) “Estabas fumando y yo te miraba. Una sonrisa iluminó mi rostro. De repente te tornaste y me viste. No tuve el reflejo de torcer la mirada, pero pienso que mi sonrisa desapareció. Luego miraste por otro lado, pero yo me quedé mirándote. No estaba intencionado de mi parte de ser sin delicadeza, pero de alguna manera esperé que volvieras a verme. Era lo que ocurrió algunos momentos después. Giraste la cabeza y nos mirábamos por segunda vez. Intenté sonreír, pero estaba muy tímida. La llegada del tren hizo terminar este momento.”

3) “Decidí que esperaría hasta que ya no lloviera antes de partir, no obstante en vez de dejar de llover, también comenzó a tronar.”
Furthermore, a closer look at other kinds of discourse markers, for instance the consequential markers, shows a difference in usage between the two groups of participants. Whereas in the assignments of the language learners the marker *entonces* is mostly used, the native group prefers *así que*. A remarkable aspect is that both groups choose consequential markers which are more typical of oral discourse. This choice could be argued by the type of assignment, the writing of a personal story, which can favour the use of discourse markers typical of the discourse. Next to this, in case of the Spanish language learners, another explanation could be added: one could suppose that they have not learned yet to choose the appropriate discourse markers in relation to the discourse register.

4) “Antes de empezar mis estudios de español, quería ir de vacaciones a España y seguir un curso de lengua, para saber si eso me gustaría. *Entonces*, durante las vacaciones de verano, fui al norte de España por dos semanas. Era un viaje de una organización, *entonces* estábamos en grupo.”

5) “No sabía muy bien cuál de estos momentos escoger para esta redacción *así que* pensando y pensando he decidido escribir sobre como nunca olvidaré el día en que nació mi hija.”

Within the category of ordering markers, opening, closing and continuative markers are often used, both by the language learners as by the native speakers. A slight remark still can be made: whereas the learners use markers of all three kinds of ordering devices, natives speaker only use the continuative type. Another remarkable aspect is the frequent usage of *entonces* in some texts written by the non-natives, as exemplified in (6) and (7):

6) “Por la tarde, tenía tiempo para descubrir la ciudad, porque Madrid tenía barres interesantes. Fui de compras en la Gran Vía donde había muchas bares, cines y tiendas. […] Había visto Atocha, la estación más antigua de Madrid, y la estación más grande de España. Dentro la estación había un jardín tropical que era hermosa. Es esta estación que conoció los ataques terroristas de 2004. *Entonces* en frente de la estación se encuentra el Centro de Arte Reina Sofía que yo había visitado. […] Luego, El Rastro era un mercado que haya cada domingo. […] Hubo visitado otro museo de pintores, que era famosa en España: el museo del Prado. En este museo hay cuadros de famosos pintores europeos. Había pintores españoles, franceses, flamencos, italianos, como Velázquez, El Greco, Goya, El Bosco, Rubens, Caravaggio. *Entonces*, la última cosa que había visto, es el Palacio Real, cerca de la estación de metro Opera, y de la Plaza de España.”

7) “El médico provoco las contracciones, pero mi angustia siguió, ya que siendo siempre con el monitorio, yo podía oír el corazón del bebe palpitar, y a cada contracción, el corazón retrasaba. Tenía miedo, hablaba a mi bebe, le decía coraje mi amor. ¡*Entonces* el médico me anunció que había demasiados riesgos… era necesario proceder a un cesárea! Me puse a llorar, mi marido hacía su posible para remontarme el moral, pero sabía que tenía también miedo.”

This particular inappropriate usage could be explained by language interference. *Entonces* as a consequential connective marker is usually translated into Dutch by *dan*. *Dan* is a frequent discourse marker in Dutch and is easily transferred into the interlanguage.

With respect to the reformulative markers, our comparison shows an interesting difference between the natives and the non-natives. Language learners used formal markers as *al fin y al cabo* and *a fin de cuentas* to summarize what was expressed previously, whereas the natives used *total* and *en fin*. These markers seem to belong more to informal speech. As an illustration:

8) “Sentado nuevamente pero apreciando todos los paisajes, las calles que se veían cuando salíamos de Madrid, los trenes que llegaban de quién sabía dónde, si *total*, conocíamos pocas ciudades en ese entonces de la tierra de los toros y el flamenco.”

9) “Nos volvimos más humanos, aprendimos de todos porque todos traíamos algo distinto dentro, *total*, veníamos de distintos lugares, era normal que todos hablemos diferente […]”
Finally, the typical Spanish metadiscoursive markers were more frequently used by the natives. Apparently, the acquisition of these devices seems to be more difficult for foreign language learners.

10) “Bien, pues era medio día cuando llegamos a la empresa, estuvimos esperando horas, sobre todo yo, pues me quedé esperando en un sofá que tenía una de las secretaria […]”

11) “Considero que yo soy una persona más expresiva pues el contacto con las personas me es muy grato,”

In the productions of the non-natives we found some cases, but often pues is used in a less natural way.

12) “A pesar de que helaba y pues la tierra estaba desbalizado, pudo evitar una viga.”

A final aspect is avoidance. According to the phenomenon “avoidance behaviour” introduced by Schachter (1974: 212), learners tend to rely on simple structures, avoiding expressions which are more difficult to produce: “The learner apparently constructs hypotheses about the target language based on knowledge he already has about his own language. If the constructions are similar in the learner’s mind, he will transfer his native language strategy to the target language. If they are radically different, he will either reject the new construction or use it only with extreme caution.” As shown in other studies on this topic, the avoidance-factor is hard to determine, but we still can assume that it plays a role in our results.

A related aspect is the one-to-one principle of interlanguage (Andersen, 1984) according to which acquisition of new forms seems to be facilitated when there is a unique correspondence between form and meaning, when there is an obvious form-function mapping. An interlanguage system “should be constructed in such a way that an intended underlying meaning is expressed with one clear invariant form (or construction)” (Andersen, 1984: 79). In that sense, unambiguous connective devices seem to be acquired more easily than highly multifunctional discourse markers. So, learners will be using more markers, when there is a clear correlation between form and function.

6. Final considerations

The corpus study on the use of discourse markers in Spanish language learners’ written compositions compared to the data of native speakers revealed that discourse markers were used almost as frequently by both groups. But the kind of discourse markers used in their productions was different. Language learners used extensively argumentative and ordering devices, mostly pero and entonces, sometimes in a less appropriate way. This cannot only be ascribed to the lack of knowledge on how to use discourse markers. We also assume a lack of instruction. Finally, avoidance behaviour and the fact that expressions with a one-to-one form-function mapping are more easily used, play also an important role in the kinds of discourse markers used in the production.

References


