

The Way to Introduce Spanish Varieties with Corpus in Classrooms*

Yan He

Guangdong University of Foreign Studies South China Business College
Guangzhou, China

Abstract—Based on the theoretical framework of Data-Driven Learning, this paper use the Atlas interactivo de la entonación del español and Corpus PRESEEA as resources for the teaching of Spanish varieties, whose objective is not that the learners get to use one or the other dialect, but that they learn to recognize them and appreciate them because they are part of the Pan-Hispanic culture. In addition, it is proven that the use of corpus can be an alternative to be exerted by teachers in language classrooms.

Keywords: *Data-Driven Learning, corpus, Spanish varieties, language teaching*

I. INTRODUCTION

The immense language community of Spanish speakers brings together a multitude of speaking communities with their own distinctive profiles. When the teaching of Spanish takes place in a non-Spanish-speaking context - China, in this case -, it becomes necessary to emphasize the Spanish variety to develop multicultural and multilingual competence on the part of the students. Fortunately, with the advent of technology, a new view of language learning and teaching has emerged: attempts to integrate computers as tools in language classrooms.

This paper suggests language corpora can enhance the quality of Spanish varieties teaching and learning in foreign language classrooms. Language teachers can play an important role in motivating the students to use corpora for their learning

II. DATA-DRIVEN LEARNING (DDL), CORPORA AND SECOND LANGUAGE ACQUISITION

DDL which was first pointed out by Johns (1991) is a method in which learners can discover grammatical patterns, word meanings or other aspects of language through searching linguistic data and investigating large amounts of authentic language. Johns expressed that DDL entails a shift in the role of teachers and students. As Chambers (2010: 355) said, teacher works as a “facilitator of the learning process, helping the learners to interpret the data, and giving them

advice on how best to search the corpus and analyse their search results.”

The term *corpus* is used to refer to a large collection of authentic texts, both written and spoken, in electronic format designed to be representative of a language variety. The use of corpora is no longer an activity interesting only to a small group of linguists. There is every reason to believe that corpus linguistics is going to impact every aspect of the way languages are taught, learnt and researched: authenticity, objectivity, verifiability, exposure to large amounts of language, new insights into language studies and enhance learner motivation. However, the introduction of corpora in the classroom might mean a tough job of teachers and learners. Both teachers and students are more used to traditional methods and may sometimes find using corpora in the classroom quite challenging: on the one hand teachers and students should be familiarized with consulting methods, otherwise they will find corpora boring and difficult; on the other there is a skeptical attitude among teachers, because of the lack of training about the exploitation of these tools and with little confidence in their effectiveness for teaching.

Boulton & Tyne (2013: 107) emphasize that “the DDL may rather be viewed as an extension of ordinary practice insofar as it builds on many aspects of existing language teaching and learning methodology.” Basically there are two ways to introduce corpora in the classroom: 1) the teacher uses the corpus as sources of resources to extract phenomenon of the language in which he wants the students catch on; 2) the learners directly consult the corpus to solve any doubt or linguistic curiosity.

In this present paper, we intend to prove that the corpora should be used as an important instrument for language teachers in their classes. Language teachers can play an important role in motivating the students to use corpora for their learning.

III. THE CORPUS AS RESOURCES FOR TEACHING SPANISH DIALECTS

Just as there are difference between American English, there are difference between the Spanish spoken in Spain (as known peninsular Spanish) and the Spanish spoken in Latin America. When starting to learn Spanish, students quickly find out that there are several ways to speak the language. A

*Fund: This study was supported by grants GDJG201801 from the Department of Education of Guangdong Province (China).

lot of Spanish dialects coexist throughout the world. For example, a child is called “niño” in peninsular Spanish, whereas those who speak Latin American Spanish use the word “nene” or “guagua”. Our students are being trained to perform in different professional activities not only related to Spain, but also to Latin America, so that we should think about the following question: what dialects and varieties of Spanish they come into contact as they travel throughout Spanish speaking countries? In class, it is necessary to expose them to as many varieties of Spanish as possible, not so much so that they learn to use them, but rather so that they become aware of value of the diversity that this entails.

Although Internet and the exchange of cultural products provide access to a multitude of samples of Spanish dialects, in fact neither publishers nor teachers usually have samples of this type specially designed for teaching. The corpus is not the panacea, but it allows taking to the classroom samples of Spanish as the concrete manifestation of speakers from different origin. They are resources adaptable to different teaching/learning contexts and to the needs of the learners, but we must be aware that the corpus is, first and foremost, structured tool designed for linguistic research and not for teaching. Although there are currently many types of corpus in Spanish, only some of them offer the facility to extract data related to dialects. In this present paper, we are going to focus on two of them due to their didactic potential: the *Atlas interactivo de la entonación del español* and the *Corpus PRESEEA*.

The *Atlas interactivo de la entonación del español* is a collaboration among various universities and research groups in Spain and America. The goal of this project is “to systematically present a series of audio and video materials for the study of prosody and intonation in the various dialects of Spanish”¹. Using the interactive map, the users can easily access the following materials: first, examples of different types of intonation contours which are used to encode various semantic nuance; second, a fragment of speech obtained by means of *Map Task* technique; third, a video interview containing a short clip of spontaneous speech.

The Project for the Sociolinguistic Study of Spanish and America² is a spoken corpus which aims to “coordinate sociolinguistic researchers from Spain and the Hispanic America in order to make possible comparisons between different studies and materials, as well as a basic information exchange. 40 research groups in different Spanish-speaking areas are involved in the project. The data are classified according to the origin, the sex, the age and the level of education of the informants. The corpus consists of semi-directed interviews, lasting between 45 minutes and an hour and a half, whose themes focus on various aspects of daily life (greetings, friendships, customs, etc.)

IV. SUGGESTIONS FOR INCORPORATING CORPORA INTO CLASSROOMS

In this section we present some proposals for teaching/learning of the diatopic varieties of Spanish from the corpus mentioned above. They can be easily adapted to any method, in different learning contexts, taking into account the educational level of learners. The basic idea is that in classrooms teachers can work with dialects to gradually sensitize learners about the cultural and linguistic richness of the language they are learning. These tasks can be considered as an approach to the linguistic variation of the Spanish-speaking world and extensions of the daily classroom teaching.

The *Atlas interactivo de la entonación del español* can be used with students of intermediate and advanced levels for activities focused on discrimination of allophones and prosodic features in different Spanish-speaking areas. A first activity consists in selecting different types of sentences from the “Sentences” menu (Statements, Yes-no questions, Wh-questions, Echo questions, Imperatives, and Vocatives) and the dialects (American Spanish: Mexican, Central American, Caribbean, Amazonian, Andean, South American Cone and other dialects; European Spanish: Northern Castillian, Central-Southern Castillian, Andalusian and Canarian). For example, you can extract, among the imperative statements, the request in all dialects and select one or two for each of the five varieties (Central-Southern Castillian, Andalusian, Mexican, Andean, South American Cone) collected in the corpus. In the record of each statement, there are the situation in which it is issued, the audio file, the translation, the fundamental frequency contour and the description of the melodic curve (“Fig. 1”):

¹ Website of the *Atlas interactivo de la entonación del español*: <http://prosodia.upf.edu/atlasentonacion/>

² Website of the Corpus PRESEEA: <http://preseea.linguas.net/Corpus.aspx>

Situation: You want to go to the movies with a friend. S/he tells you that s/he has work that s/he needs to do, but you know that s/he can leave it for later. What do you say to convince him/her to accompany you?

Semantic interpretation: Request.

Speaker: MOT

Utterance: Va, vente al cine, hombre.

Translation: Come on, man, come (with us) to the cinema.

▶ -0:00

0.5 1.5 2

300
260
220
180
140
100

Fundamental frequency (Hz)

Va, vente al cine, hombre

4 1 0 4 4

L* HL% L+H* L% L* HL%

Fig. 1. Melodic curve of interactive Altas of Spanish intonation.

Teachers can copy melodic curves in disorder on a sheet, ask students to associate them with the corresponding statement by the listening and then show them to which area they correspond. Another possible activity is to invite students to work directly with the corpus: group the transcripts of the sentences and observe the ways in which they are constructed to see grammatical, pragmatic and lexical differences, for example:

Activity. In the following “Table I” you will find the transcript of the sentences you just heard. Taking into account that all conversations are informal, what do you observe? Now look at the vocabulary: what similarities and differences do you notice?

TABLE I. INTERACTIVE ALTAS OF SPANISH INTONATION

Spanish-speaking areas	Imperatives: request
Madrid	Va, vente al cine, hombre.
Ja n	Va, vente!
Mxico DF	Ay, ya! Vamos al cine, no seas payaso.
Lima	Ah, ven conmigo, pues!
Buenos Aires	Dale! Ven  Ponete buena onda!

The Corpus *PRESEEA* allows to select the texts with the orthographic transcription, taking into account the variables

of gender, age and educational degree. The data was collected according to questions organized by topics, which can be of great help when selecting material for the ELE classroom. Following a thematic criterion you can select two or three fragments of interviews with informants of different origin who speak of the same topic and ask the students to observe the characteristics of one or another dialect. Each interview begins with a brief exchange of the use of formal and informal treatments in the speaking community, which helps to elaborate tasks focused on both content and form. The questions related to the treatment are shown as follows: 1) Cmo quiere que le trate, de to de usted? Cmo tratas a tus amigos? Y si son personas mayores? Si es alguien joven que no conoces? Y si es una persona mayor, hombre o mujer, al que le preguntas por una calle? Y a tu mdico? A los vecinos con los que no tienes mucho contacto?

A possible activity based on the Corpus *PRESEEA* is shown below:

Activity: you will hear three conversations that a Spanish, a Uruguayan and Peruvian talk about treatments that are commonly used in their country in formal and informal relationships. Listen to the texts several times and try to complete the following “Table II”:

TABLE II. ACTIVITY. ¿WHICH IS THE FAVORITE TREATMENT?

Questions	Spain	Uruguay	Peru
<i>¿Which is the favorite treatment: tú, vos or usted?</i>			
<i>When you use formal treatment and when informal?</i>			
<i>In which of the three varieties of Spanish is the use of "tú" more frequent?</i>			
<i>What kind of relationship express the use of "tú" in these three varieties of Spanish?</i>			

Activity: look at the transcription of some fragments of the conversations of two Uruguayan speakers and see when change the use of "vos".

E.: [...] Bueno / gracias por dedicarme este tiempo para poder hacer la entrevista.

I.: de nada.

E.: nosotras / por supuesto que nos conocemos de hace tiempo / nos tratamos de vos / as íno más / contame cuando vos te tenés que dirigir a alguien ¿cómo lo tratás?

I.: depende en qué ámbito.

E.: por ejemplo

I.: en los trabajos cuando no hay mucha confianza / de usted / bueno y cuando nos conocemos / tenemos confianza / nos tuteamos.

E.: ... ¿y sit e encontrás con alguien por la calle / y tenés que preguntarle algo? ¿te dirigís?

I.: de usted.

E.: de usted ¿siempre? / ¿y si es de tu edad?

I.: no / si es de mi edad nos tuteamos.

(MONV_M12_020 (Corpus PRESEEA))

E.: [...] bueno / vamos a ver / para empezar / nosotros dos ¿nos tratamos siempre de vos / ¿no?

I.: sí

E.: y vamos a seguir tratándonos de vos / en esta entrevista ¿o alguna vez nos trataos de túo de usted?

I.: no

E.: no

I.: che ¿cómo andás? / che esto / siempre as í

E.: che y vos ¿no? ¿túno?

I.: no / túno

E.: tú no / usted no // vos ¿a quién? ¿cómo tratás vos a alguien? ¿en general a la gente conocida? ¿a tus amigos por ejemplo?

I.: che / ven íac á/ o ¿cómo andás? / ven í/ siempre as í

E.: ¿con vos?

I.: sí/ che vos

E. ¿usted / no?

I.: rara vez // bueno sí/ en el ámbito de trabajo / viste / sí / cuando me llaman del directorio / ya / lo trato de usted / ah í sí / hay otra formalidad

E.: [...] suponete siguiendo este tema / una mujer o un hombre mayor / al cual / le querés hacer una pregunta en la calle / en fin / ¿cómo le decís? ¿túo usted?

I.: si es una persona mayor

E.: mayor

I.: y no lo conozco le digo usted

E.: ¿y a los vecinos tuyos / por ejemplo?

I.: ¡ah no! che

E.: che / todos todos che

I.: sí

E.: [...] ¿y a tus padres también los trataste siempre de vos?

I.: vos y che vieja

(MONV_H22_004 (Corpus PRESEEA))

Then, according to the above transcripts, you should try to summarize the conjugation rules related to "vos".

V. CONCLUSION

In this paper we have tried to show the potential of some corpus as resources to bring ELE students the varieties of Spanish. The objective of the proposed activities is not that students learn to use one or the other dialect but that they become aware of the linguistic richness of the immense Spanish-speaking community. It is not about promoting attention to regionalisms, but rather of bringing to the classroom some basic principles of panhispanism, which recognizes and values diversity of the unity of the language. We want our students to know how to interact respectfully with the different Spanish-speaking cultures, it is important that they learn to appreciate their dialect. We defend that the corpus aren't the panacea, but they are a valuable resource, another tool to help introduce and teach the varieties of Spanish in the classrooms.

REFERENCES

- [1] Bernardidi, Silvia (2004): "Corpora in the classroom. An overview and some reflections on future development". In Sinclair, John M.: How to use corpora in Language Teaching. Amsterdam: John Benjamin Publishing, pp. 15-36.
- [2] Boulton, Alex & Henri Tyne (2013): "Corpus linguistics and data-driven learning: a critical overview". In Bulletin VALS-ALSA 97, pp. 97-118.
- [3] Chambers, Angela (2010): "What is data-driven learning?". In O'Keefe, Anne; McCarthy, Michael (eds.): The Routledge Handbook of Corpus Linguistics. Routledge: London/New York, pp.345-358.
- [4] Cruz Piñol, Mar (2012): Lingüística de corpus y enseñanza del español como 2/L. Madrid: Arco Libros.

- [5] Gilquin Gaetanelle & Sylviane Granger (2010): "How can data-driven learning be used in language teaching?" In O'Keeffe, Anne; McCarthy, Michael (eds.): *The Routledge Handbook of Corpus Linguistics*. Routledge: London/New York, pp.359-370.
- [6] Johns, Tim (1991): "Should you be persuaded: two examples of data-driven learning". In *ELR Journal*, 4, pp.1-16.
- [7] Moreno Fernández, Francisco (2000): "Corpus para el estudio del español en su variación geográfica y social: el corpus PRESEEA". In *Oralia: Análisis del discurso oral*, 8, pp.123-140.
- [8] Giménez Folqués, David (2017): "La enseñanza de la gramática panhispánica en clase de español como lengua extranjera (ELE)". In CAUCE. *Revista Internacional de Filología, Comunicación y sus Didácticas*, n.º 40, pp.89-95.