JEL UERJ **VI**

Jornadas de Estudos da Linguagem

02 a 04 de DEZEMBRO de 2010

Research, Development and Application of a Learning Resource for Enhancing Listening and Spoken Skills in Spanish.

Elena Paz Vizcaya.

Dublin Institute of Technology

Digital Media Centre, Aungier St.

Dublin 2, Ireland.

mailto:elena.pazvizcaya@dit.ie

_____

Native speech is directed towards native listeners, not designed for comprehension and analysis by language learners. Speed of delivery - or economy of effort - produce a speech signal to which the native listener can assign the correct words — there are no discrete words in the speech signal itself. Experience of using timescaling with recorded English has highlighted the benefit of making slowED speech available to the language learner or researcher, as opposed to slow speech – i.e delivered slowly. The main contribution to knowledge of this project is to generate a unique research and analysis corpus (audio resource) of informal NS-NS Spanish speech recorded at a high level of audio quality and which is amenable to linguistic analysis, and which can therefore act as a resource to test the research hypotheses.The work carried our to date has focused on the analysis of the acoustic parameters related to segmental features of the speech units.

Keywords: Corpus linguistics & Language pedagogy, discourse analysis, speech prosody, connected speech, dialogic fluency.

_____

1- Introduction

When we speak, besides linking the sounds with which we form words, we change the pitch and the intensity, lengthen some elements and modify the quality of the voice. Thus, our words are perceived as emphatic, assertive, rushed, dubious or suggestive among many other nuances. In order to interpret messages, speakers not only focus their attention on the phonological contrasts between segments; they are sensitive to other elements of the signal that provide information regarding aspects like lexical stress, sentence modality or the speaker's intention. The aim of this research is, on the basis of analysis of the nature of Spanish spoken language, to help learners of Spanish as L2 derive benefit from the digital-media language learning technologies developed by Dublin Institute of Technology (DIT) research group. It is hoped that these Technology-Enhanced Language Learning (TELL) tools will resolve the problems encountered by

learners of Spanish as L2 in the acquisition and improvement of their aural skills. In this paper, the process and analysis carried out to identify suitable segments to apply the time-scaling tool and test the research hypotheses – segments which will later populate the nascent Dynamic Speech Corpus – is explained.

2- Methodology

"The variety of expressions and the functional diversity of intonation are manifested in their fullness in the non-scripted speech". (Hidalgo 2008). 'Duologues' are natural, relaxed dialogues between two people who know each other and recorded in such a manner that each interlocutor's performance can be studied in isolation, thus avoiding problems normally caused by cross-talk and back-channelling. To this end the isolation booths in the Digital Media Centre (DMC) have been made available to record the samples which ultimately will allow the construction of a corpus *'a principled collection of spoken or written language stored on computer and available for quantitative and qualitative analysis'* (O'Keeffe 2006) of natural spoken Spanish for analysis and provide a basis for effectively applying the time-scaling tool. The recordings are designed and recorded using the facilities of the CSAL (Cognition, Speech and Audio Laboratory) in the DMC where real, natural, emotional spontaneous speech can be recorded with a high degree of naturalness. A database for effective linguistic and emotional analysis must contain *'real emotional assets'*, and '*not obtained from broadcast sources and not using actors'* (Cullen et al 2006). After the recording phase, listening to the WAV file is the next step on the way to linguistic analysis. Careful listening to these recordings reveals several salient linguistic features, for example, 'sinalefas' and their speed of delivery. After listening to the WAV file, transcription of the acoustic signals into their orthographic, idealised forms – called the *orthotext*– is the next step. The transcriptions used in the corpus, therefore, are not phonemic or phonetic transcriptions, but represent the idealised or citation form, which can then be used in a search string to locate reduced phonetic features in the speech signal. Segmenting, labelling and tagging the sequences are the next stages in the creation of the speech corpus. The 'flow-unit' is set as the standard database unit by the DMC linguistic group. This is *'a segment which has its own flow and pragmatic integrity. It is speaker determined and can be timed, and is bounded by pauses – no matter how brief – or marked by a sudden change of tack/thought'* (Campbell et al 2006). The purpose of segmentation is to facilitate tagging the duration of each spoken sequence to discover the features of natural speech for inclusion in the corpus. Tagging is defined as the process to time the WAV file so as to calculate the delivery speed of each sequence. The speed here, based on Laver (1994: 158), refers to *'articulation rate'* as opposed to *'speaking rate'*. The phonetic variables considered for the statistical analysis of the flow units are:

1. The average duration of the flow unit (excluding empty and full pauses).
2. The number of phonetic syllables in each flow unit (as opposed to phonological syllables –*orthosyllables*– transcribed as orthotext).

3- Results

A small corpus of four 'duologues' (four speakers – one male and three females–), amounting to one hour and ten minutes recorded, has been analyzed.

CORPUS		Nº segments	Empty pauses	Full pauses	Flow units analyzed
1st Duologue Jesus-Patricia-1	Speaker 1	403	2	73	**328**
	Speaker 2	677	6	264	**407**
2nd Duologue Jesus-Patricia-2	Speaker 1	732	1	134	**597**
	Speaker 2	303	0	51	**252**
3rd Duologue Susana-María-1	Speaker 3	773	0	140	**633**
	Speaker 4	548	2	68	**478**
4th Duologue Susana-María-2	Speaker 3	760	1	115	**644**
	Speaker 4	728	3	83	**642**
TOTAL		4924	15	928	**3981**

Results show that in the 4924 segments, the 3981 flow units analyzed may have between 1 and 23 syllables. The most frequent in spontaneous dialogic speech, are those between 1 and 4 – *this is orthosyllables*– The number of empty pauses (15) is significantly small compared to the number of full pauses (928) which are frequent in natural communicative situations (Schriberg, 1999)
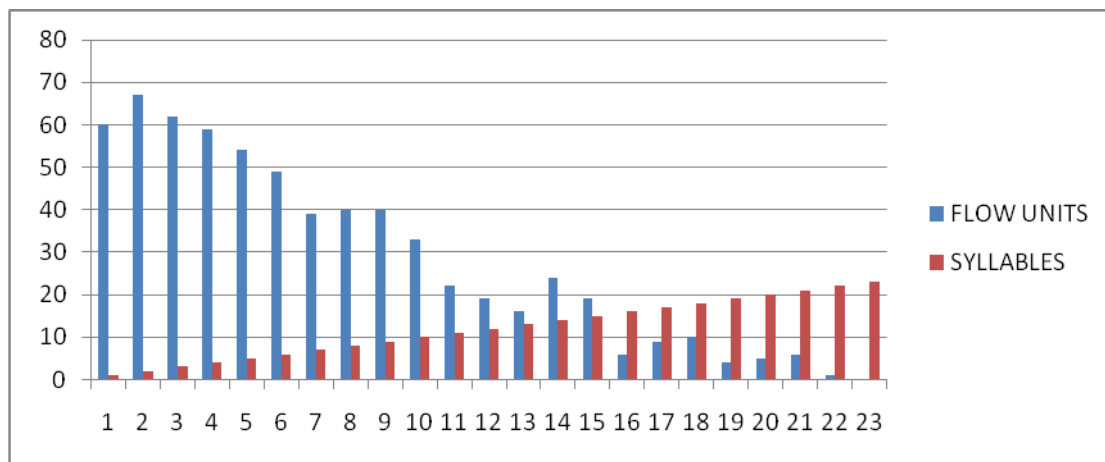
This analysis also shows that the reduction in the number of syllables due to **coarticulation**, "*mutual influence between adjacent sounds*" (Llisteri, 2003:11) or **sinalefa**, *"the pronunciation in one syllable of groups of vowels resulting of the joining together of words in the speech chain"* (Quilis, 1996: 150) is more evident in those flow units that have between 5 and 15 syllables. The average drop is between 2 and 3 syllables. Less frequent segments of between 15 to 25 syllables per flow unit show the same amount of reduction.

The average duration of the flow units studied (from 5 to 18 syllables) is of 1.6 seconds. It is important to highlight the fact that smaller flow units (1 to 4 syllables) and bigger ones (19 – 23) syllables have been discarded and obviously show significantly different values.

4- Conclusions

Although frequent, flow units of 1, 2 or 3 syllables are not suitable segments to test the research hypotheses since we can barely find the examples of reductionist techniques that may pose difficulties for learners of Spanish. The same would apply to pauses; there is not data of interest for this study in them. Those speech units that have between 5 and 15 syllables have proven to show many of the features that may cause that 'blur' of informal Spanish speech flow to which learners are expected to attune themselves to.

"Language pedagogy – following academic phonology – has avoided dealing with 'the extremely messy' aspects of everyday speech" (Cauldwell, 1998; 2000).

Therefore, those segments that show a significant reduction are suitable to apply the time-scaling tool and make them available to the language learner in order to identify which reductionisnt techniques, proper of connected speech in Spanish – *asimilations, elisions, syllabic contraction, etc. –* prove to be more difficult to overcome.

_____

REFERENCES

Hidalgo Navarro, A. 2008. *Algo más sobre la función demarcativo-integradora de la entonación: el caso de las series enumerativas.* Actas del XXXVII Simposio Internacional de la Sociedad Española de Lingüística (SEL), editadas por Inés Olza Moreno, Manuel Casado Velarde y Ramón González Ruiz, Departamento de Lingüística hispánica y Lenguas modernas. Pamplona, Servicio de Publicaciones de la Universidad de Navarra, 2008. ISBN: 84-8081-053-X

O'Keeffe, A. 2006. *Manufacturing pseudo-intimacy in media discourse*. Addressed at IVACS Conference, 2006.

Cullen, C., Vaughan, B., Kousidis, S., Wang, Y., MacDonnail, C. and Campbell, D. 2006. Generation of High Quality Audio Natural Emotional Speech Corpus using Task Based Mood Induction. Addressed at InSciT2006 Conference.

Campbell, D., Meinardi, M., Richardson, B. and Wang, Y. 2006. Natural English in Speaking and Listening Activities. Addressed at MATSDA/ACELS Conference, 2006.

Laver, J. 1994. *Principles of Phonetics*. Cambridge: Cambridge University Press.

Schriberg, **E.** 1999. *Phonetic consequences of speech disfluency".* Proceedings of the international Congress Of Phonetic Sciences ((ICPhS-99), Vol.I, pp. 619-622 San Francisco.

Llisterrí. J. 2003 "*Lingüística y tecnologías del lenguaje*", Lynx. Panorámica de Estudios Lingüísticos (Departament de Teoria dels Llenguatges, Universitat de València) 2: 9-71.

Quilis, A. (1993). *Tratado de fonología y fonética españolas*. Madrid: Gredos.

Cauldwell, R. 2002. *"Phonology for Listening: Relishing the messy".* richard@speechinaction.com