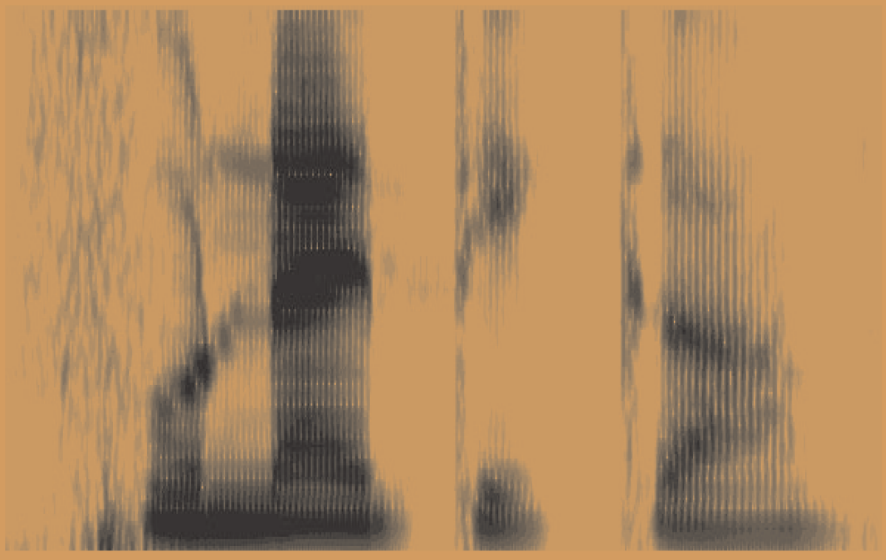
Estudios de Fonética Experimental

XX

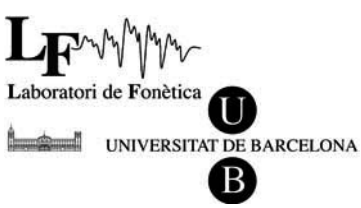


[f u ˈn ɛ t i k ə]

Laboratori de Fonètica
UNIVERSITAT DE BARCELONA

Barcelona, 2011

ESTUDIOS DE FONÉTICA EXPERIMENTAL XX



Barcelona, 2011

La revista *Estudios de Fonética Experimental* está recogida en:

Sumarios ISOC-Ciencias Sociales y Humanidades, producidos por el Centro de Información y Documentación Científica (CINDOC) del CSIC: http://bddoc.csic.es:8080/indez.jsp, http://resh.cindoc.csic.es/index.php

Latindex: http://www.latindex.unam.mx/
Criterios latindex cumplidos: 32

Dialnet de la Universidad de La Rioja: http://dialnet.unirioja.es/

CARHUS Plus+2010 de la Generalitat de Catalunya (AGAUR): Valoración A
https://www10.gencat.net/agaur_boga/AppJava/FlowControl?cmd=CercaRevistesRevCmd&view=VCercaRevistesRev&procedencia=CercaRevistes&Submit=Cercar&rev_titol=&rev_issn=1575-5533&rev_ambit=&rev_valoracio=T

MIAR Live 2010 (UB): MIAR (Difusió ICDS: 3.915)
http://miar.ub.es/consulta.php?issn=1575-5533

ULRICH'S Serials Analysis System:
http://www.ulrichsweb.com/ulrichsweb/Search/ViewSearchResults.asp?navPage=1&SortOrder=Asc&SortField=f_display_title&collection=SERIAL&QueryMode=Simple&ScoreThreshold=0&ResultCount=25&ResultTemplate=quickSearchResults.hts&QueryText=sn=1575%2D5533&

Todos los volúmenes de la revista en pdf están en las siguientes direcciones:
http://stel.ub.edu/labfon-ub/es/content/publicacion-estudios-de-fonetica-experimental
http://www.raco.cat/index.php/index/raco/cercaXR?searchJournal=399

Universitat de Barcelona
Dipòsit Legal: B-31.504-1984
ISSN: 1575-5533
Impressió: Gráficas Rey, S.L.
Tiratge: 350 exemplars

Estudios de Fonética Experimental

Dirección de «EFE» *Estudios de Fonética Experimental*
Laboratori de Fonètica
Universitat de Barcelona
Aulari Josep Carner, 5è
Gran Via de les Corts Catalanes, 585
08007 BARCELONA
T. 934035650
e-mail: labphon@ub.edu
http://stel.ub.edu/labfon-ub/es/publicacion-estudios-de-fonetica-experimental

Han actuado como revisores anónimos para uno o más artículos propuestos para este número los siguientes investigadores (por orden alfabético), tanto para los artículos aceptados como para los artículos rechazados:

Lourdes Aguilar Cuevas	(Universitat Autònoma de Barcelona-ESP)
María José Albalá	(Consejo Superior de Investigaciones Científicas-ESP)
Mercedes Cabrera	(Universidad de Las Palmas-ESP)
Micaela Carrera de la Red	(Universidad de Valladolid-ESP)
Ramon Cerdà Massó	(Universitat de Barcelona-ESP)
Maximiano Cortés Moreno	(Tainan University of Technology-TAIWAN)
Josefa Dorta Luis	(Universidad de La Laguna-ESP)
Eva Estebas Vilaplana	(Universidad Nacional de Educación a Distancia-ESP)
Pedro Gómez Vilda	(Universidad Politécnica de Madrid-ESP)
José Ignacio Hualde	(University of Illinois at Urbana-Champaign - USA)
Takuya Kimura	(Seisen University, Tokyo-JAPON)
Giovanna Marotta	(Università di Pisa-IT)
Victoria Marrero Aguiar	(Universidad Nacional de Educación a Distancia-ESP)
Philippe Martin	(Université Paris 7-FR)
Eugenio Martínez Celdrán	(Universitat de Barcelona-ESP)
José Matas Crespo	(Doctor en Lingüística por la Universitat de Barcelona-ESP)
Joan Carles Mora Bonilla	(Universitat de Barcelona-ESP)
Elsa Mora Gallardo	(Universidad de Los Andes, Mérida-Venezuela)
Brian Mott	(Universitat de Barcelona-ESP)
Rosa Miren Pagola Petrirena	(Universidad de Deusto-ESP)
Antonio Pàmies Bertrán	(Universidad de Granada-ESP)
Pilar Prieto Vives	(Universitat Pompeu Fabra-ESP)
Joaquín Romero Gallego	(Universitat Rovira i Virgili, Tarragona-ESP)
Paolo Roseano	(Universitat de Barcelona / Universitat Pompeu Fabra-ESP)
María Josep Solé	(Universitat Autònoma de Barcelona-ESP)
Dorota T. Szmidt	(Universitat de Barcelona-ESP)
Guillermo Toledo	(Consejo Nacional de Investigaciones Científicas y Técnicas -CONICET- ARG).
María del Mar Vanrell	(Universidad Autónoma de Madrid-ESP)

ÍNDICE

Artículos

Miscelánea

Notas y reseñas

«Estudios de Fonética Experimental» informa:

ARTÍCULOS

CORPUS BASED PROSODIC VARIATION IN BASQUE: Y/N QUESTIONS MARKED WITH THE PARTICLE *AL*

VARIACIÓN PROSÓDICA EN VASCO BASADA EN CORPUS: PREGUNTAS SÍ/NO MARCADAS CON LA PARTÍCULA *AL*

GOTZON AURREKOETXEA
Universidad del País Vasco
gotzon.aurrekoetxea@ehu.es

IÑAKI GAMINDE
Universidad del País Vasco
inaki.gaminde.@gmail.com

AITOR IGLESIAS
Universidad del País Vasco
txadonak@gmail.com

Artículo recibido el día: 8/11/2010
Artículo aceptado definitivamente el día: 13/07/2011
Estudios de Fonética Experimental, ISSN 1575-5533, XX, 2011, pp. 11-31

ABSTRACT

This paper describes the intonational variation between two generations in three different localities of the Basque Country, using data recorded and organized in the EDAK corpus (Dialectal Oral Corpus of the Basque Language) and analysing the usage of just one type of sentence, namely y/n questions. In the selected localities, there are two morphological ways of constructing this type of sentence: using the morphological marker *al* before the verb (or between the verb and the auxiliary) or not using it. First, we identify the phonological patterns that exist in these localities. Then, we analyse intergenerational variation, according to the five different phonological patterns found. Finally, we study the geo-prosodic variation which exists between older and younger people from these localities.

Keywords: *linguistic corpus, Basque language, prosody, socio-prosodic variation, geo-prosodic variation.*

RESUMEN

El artículo trata la variación entonacional entre dos generaciones en tres localidades distintas situadas en el centro del espacio lingüístico vasco en frases interro-gativas absolutas con datos del EDAK (Corpus dialectal oral del euskera). En dichas localidades hay dos posibilidades morfológicas para construir esta clase de oraciones: usando la marca interrogativa *al* delante del verbo (o entre el verbo y el auxiliar) o sin ella. En el análisis primeramente se han determinado los modelos fonológicos entonativos de las localidades en esta clase de oraciones. En segundo lugar se ha analizado la variación intergeneracional. Y finalmente, se ha analizado la variación geoprosódica entre dos generaciones de dichas localidades.

Palabras clave: *corpus lingüístico, vasco, prosodia, variación socio-prosódica, variación geo-prosódica.*

1. INTRODUCTION

This paper aims to show the prosodic variation that exists in several localities of the Basque Country, using one type of sentence: y/n questions. In order to do this, we use the Dialectal oral corpus of the Basque language or EDAK (Aurrekoetxea, Sánchez and Odriozola 2009). This research was carried out by the EUDIA research team, located at the University of the Basque Country (UPV-EHU). The current project has two aims: first, to obtain basic prosodic information from all varieties of the Basque language, and second, to complete the first study of sociolinguistic variation in Basque, using recorded data from two generations in each locality. To date, there are no studies on sociolinguistic variation in Basque other than the partial ones by Aurrekoetxea (2008 and 2010) and Ormaetxea (2008, 2011), and none on prosodic sociolinguistic variation. The EDAK corpus constitutes the first step to explore the basic principles of this kind of variation, and hence should be used as the basis for future work in this field.

This paper is organized into five main sections following the introduction: in section 2, we look at the strategies employed in Basque to form y/n questions; section 3 deals with the methodology used to gather information; in section 4 we provide a phonological analysis of the data; in sections 5 and 6 respectively, we analyse the socio-prosodic and the geo-prosodic variation in each of the dialects; and finally, we draw conclusions and make several proposals for future research.

2. Y/N QUESTIONS IN BASQUE

It is known that the languages of the world employ different strategies for constructing y/n interrogative sentences (Dryer 2005). These kinds of sentence have been one of the most often-analysed items of intonation research in neighbouring languages, including Catalan (Carrera et al. 2004; Fernández Planas et al. 2007; Martínez Celdrán et al. 2005; Pradilla and Prieto 2002; etc.), Galician (Fernández Rei et al. 2005) and Spanish (Dorta and Hernández 2005). The intonation of y/n questions has also been studied in Basque (Aurrekoetxea and Iglesias 2010; Gaminde 2004, 2006, 2010; etc.). A number of research teams and projects (GrEP[1], AMPER[2], etc.) have also taken them into account in their studies.

[1] http://prosodia.upf.edu/home/en/index.php

[2] Contini et al. 2002.

Apart from intonation contours, there are other strategies in Basque for building y/n questions, namely the use of interrogative markers. In some varieties, the particle *al* is used before the verb or between the verb and auxiliary in periphrastic constructions[3]. A different strategy is to add the suffix *-a* to the verb or to the auxiliary in periphrastic forms[4]. From a morphological point of view, there are thus three possible types of y/n questions: morphologically unmarked questions, questions marked with the particle *al* and questions marked with the suffix *-a.*

In this paper we test the hypothesis which states that the intonation curve is different in questions which use only intonation tools and in questions which use a morphological marker.

3. METHODOLOGY

3.1. Questionnaire, localities and informants

The questionnaire used for the EDAK project has 201 questions, of which 23 deal with questions of intonation, and out of these 23, 6 concern y/n questions. The questionnaire was created in line with the existing literature on Basque prosody. The aim of the corpus is to present material gathered from two generations in 100 localities of the Basque Country, selected after a thorough analysis of dialectal variation in Basque, and including the whole of the Basque-speaking territory on both sides of the Spain/France political border. This is the first step for analysing sociolinguistic variation in the Basque language. The research treats all of the Basque Country as one unit and gathers information from only one person from each generation in each locality. We are aware that this does not strictly satisfy the number of informants required by sociolinguistic research, but it allows us to isolate the first preliminary features, which must subsequently be developed in future research. We took into account the weakness of the corpus on this point, and took great care both in the selection of the informants and during the interviews themselves in order to gather appropriate data from each generation.

[3] In the sentences analysed in this paper, the particle *al* is always placed between the main verb and the auxiliary.

[4] There are phonological rules for the conbination of the suffix *-a* with the last vowel of the auxiliary verb: da + -a > dea, dü + -a >dia, etc.

To elicit the data, a member of the research team read each sentence in Spanish or French and asked the informant to translate it orally into his or her usual variety of Basque. This part of the questionnaire, related to intonation, was repeated three times, but at different moments during the interview, following the methodology of the AMPER project (Contini et al. 2002).

The research team tried to obtain an unbiased attitude from the informants when doing the questions. Although the interview is a formal task, informants were asked to speak as they would in a normal conversation with their friends. Informants were explicitly asked to employ their usual variety of Basque, without thinking about other ways of saying the same thing in other dialectal varieties or in the standard variety.

3.2. Data gathering, sound annotation and labelling, and storage

For data gathering and sound recording, laptops equipped with Audacity software[5] and USB microphones (PC Headset 960 USB)[6] were used.

The annotation task was carried out using the SFSWin software program[7]. Although this tool offers the option of automatic annotation, we opted for manual annotation because of the structure of the data gathered.

The output of the SFSWin program is a text file which connects the annotations to each question with their sound location in the recording. Once the sound had been annotated, we carried out the labelling task using the "txertatu_etiketak" ActivePerl script[8]. This script was adapted by Aholab, the Signal Processing Laboratory (http://aholab.ehu.es) of the University of the Basque Country (UPV-EHU), which we work with.

The data storage was carried out using two formats: TEI and MySQL. Each of these formats has tools for performing different analyses. The change from one format to the other is made automatically.

[5] http://audacity.sourceforge.net/

[6] www.logitech.com

[7] http://www.phon.ucl.ac.uk/resource/sfs/download.htm

[8] http://www.activestate.com/

The following step was the acoustic analysis of the data. For this task we used *Praat*[9]. The Segment-data script[10] is used to extract quantitative information automatically, so it can be used in quantitative analyses.

In this paper we examine the following 6 y/n questions:

1. Has he/she come [Etorri da]?
2. Has he/she bought it [Erosi dau]?
3. Has the friend arrived [Laguna sartu da]?
4. The friend, has he/she arrived [Laguna, sartu da]?
5. The friend has arrived [Laguna etorri da]?
6. Has he/she bought the bread [Erosi dau ogia]?

We examine data from three localities (see map of figure 1): Amezketa (which belongs to the Gipuzkoan area according to traditional dialect classification), Leitza and Sunbilla (these latter are included in the High Navarrese dialect area). The reasons for this choice of localities are connected with the aim of this paper, namely to study the differences in the intonation curve of y/n questions with and without a question cue. In all three selected localities, we found two ways of forming y/n questions: the use of the particle *al* (figures 2 and 3), and a change in the intonation contour in the absence of a particle (figures 4 and 5).

For this research, we use the Autosegmental/metrical (AM) model of intonational analysis to examine the recorded data (Hualde 2003; Prieto 2003 and 2006; Toledo 2007). This model is one of the most widely used for analysing intonation in different languages. It has also been applied to the Basque language by Elordieta (2003, 2007a and 2007b, etc.), among others.

Leaving aside the presence or absence of interrogative particles, and taking into account previous literature on this subject, we consider that the most relevant feature of y/n interrogative sentences in Basque is the pitch of the last part of the sentence, and more precisely, the part which follows the last accented syllable. For this reason we focus on the final syllables of the sentences to see whether the intonation curve is upward or downward.

[9] http://www.fon.hum.uva.nl/praat/

[10] Lyko, K. (2008): "Segment_data.praat" script. GNU General Public License.

With regard to the type of boundary tones, and following previous literature about Basque intonation (Elordieta 2000, 2003; Gaminde 2003, 2007, 2010), we consider four contours: H%, L%, HL% and LH%. These are the main border contours we have found in the Basque language in these kinds of sentence.

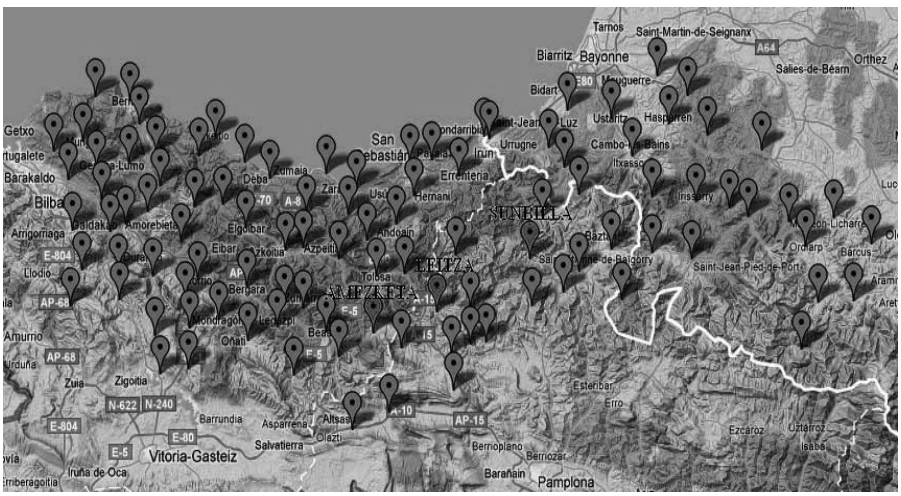


Figure 1. *Localities of the EDAK corpus and localities chosen for this study*[11].

4. PHONOLOGICAL ANALYSIS OF THE DATA

The phonological models found in the data from Amezketa, Leitza and Sunbilla are the following.

[11] Map made with Google Maps.

4.1. Y/n questions marked with the particle *al*

We observe two main types of intonation contour: curves with a final H% boundary tone, as in figure 2, and curves with a final L%, as in figure 3. All of the data gathered from these three localities follows one of these two patterns.

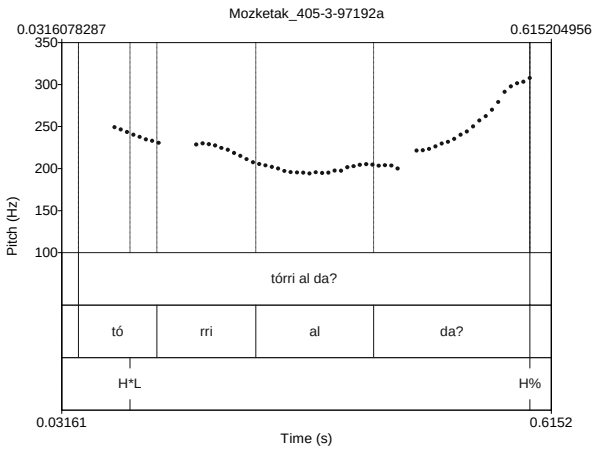


Figure 2. *Has he/she come? (Leitza).*

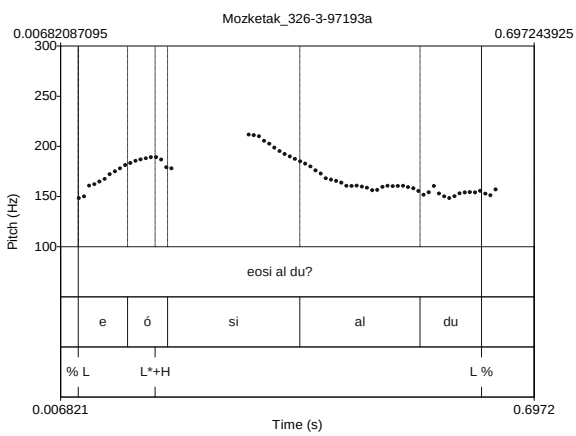


Figure 3. *Has he/she bought it? (Amezketa).*