

WORKSHOP ON INTERNATIONAL COOPERATION AND STANDARDIZATION OF
SPEECH DATABASES AND OF SPEECH I/O ASSESSMENT METHODS
CHIAVARI, SEPTEMBER 27-28, 1991

PHONETIC CRITERIA FOR THE DEVELOPMENT OF A SPEECH DATABASE IN SPANISH
(THE ALBAYZIN PROJECT)

Five speech research groups in Spain are developing a speech database that would be used in current and future research projects in the field of speech recognition and natural language processing in Spanish. A general overview of the project will be given in this Workshop by C. Nadeu (Universitat Politècnica de Catalunya) and more specific topics will be discussed by F. Casacuberta (Universitat Politècnica de València). In this contribution, the phonetic criteria for the development one of the databases included in the project will be presented.

The main problem in the design of a speech database is coping with the inherent variability of speech. Partial efforts have been made up to now to develop different phonetic corpora for Spanish, but they have not attained a large-scale coverage of the most important sources of variability found in speech. One of the databases being developed under the project Albayzin - called the 'phonetic database' - will try to approach such an exhaustive coverage. The following factors have been considered:

1. Speaker - dependent sources of variability

1.1. Sex of the speaker

Half of the speakers in the database will be male, and the other half female

1.2. Speaker's dialect

The most relevant geographical variations of Spanish will be taken into account, choosing speakers from six main dialectal areas (Madrid, Barcelona, Valencia, Granada, Valladolid and Galicia)

1.3. Speaking rate

This factor will most probably be controlled, but the possibility of having different speaking rates for the same corpus of utterances is not to be dismissed. Speaking rate variations will, of course, naturally appear in different speakers

1.4. Number of speakers

100 speakers will be used for the phonetic database, although only a few of them (6) will record the entire corpus.

2. Phonetic content

2.1. Number of items and phonetic balance

The corpus will consist of 200 sentences of ± 15 syllables each, divided in 20 groups of 10 sentences. Each group will be phonetically balanced.

2.2. Allophonic inventory and allophonic variability

Different phonetic textbooks give different allophonic inventories for Spanish. An exhaustive listing of all the allophones that appear in the literature has been made, and the causes that determine allophonic variation - also according to the literature in Spanish phonetics - have been classified as follows:

a.- adjacent allophone

a.1.- assimilation in place of articulation

dental stops, nasals, fricatives and laterals are sometimes assimilated to the place of articulation of the following allophone

a.2.- assimilation in voicing

a.2.1.- devoicing of voiced allophones

some voiced allophones are devoiced in syllable final position followed by a voiceless consonant

a.2.2.- voicing of voiceless allophones

voiceless fricatives are voiced before a voiced consonant

b.- syllabic structure

open and close vocalic allophones appear to be found in Spanish according to the syllabic structure of the word to which they belong; the approximant or obstruent character of stops depends on their position in the syllable; vowels and semivowels (glides) also alternate according to syllable structure

c.- position in the utterance

Position in the utterance is a main factor in the choice between fricative or affricate variants of the phoneme /j/ as well as in the choice between approximant and obstruent stops; vowel reduction also depends on this factor

d.- stress

Stress determines vowel reduction and also the choice between the fricative and the affricate variant of /j/

Taking all these factors into account, a total of 63 allophonic variants have been found. This inventory has been tentatively reduced to 33 units, which is close to the number of phonemes that is usually quoted for Spanish (24).

The following four factors of variability have been retained:

a. preceding and following allophone

b. position in the syllable

initial vs final position will be considered here c.

degree of stress

stressed vs unstressed is the main dichotomy to be considered, but different degrees of stress may also be taken into account. d. position in the word

In order to assess the validity of the choice of allophones and to study the influence of the different factors of variation considered, a pilot study is being undertaken on a corpus of spontaneous speech in Spanish. This study consists in:

- (1) transcription in conventional spelling of two 1 h conversations with 2 native speakers of Spanish
- (2) automatic phonetic transcription of the corpus, assigning also stress and pauses
- (3) statistical study of the corpus

The following data will be obtained for each allophone

(3.1.) frequency of occurrence of preceding and following allophones
this will allow us to know the most frequent and the most infrequent combinations in order to assess the relevance of assimilation and voicing/devoicing phenomena

(3.2.) frequency of occurrence in each of the three syllable positions: initial, final and medial
this data will allow to assess the need to include syllable position conditioned allophones in the final inventory

(3.3.) frequency of occurrence in stressed and in unstressed syllables
the need to include stressed and unstressed variants of a given allophone will be clarified in this study

It is hoped that this preliminary study will help us in gaining more insight into the weight of the different factors that are known to determine allophonic variation in Spanish, and also in determining the final allophonic content of the proposed database.

JOAQUIM LLISTERRI & DOLORS POCH

Departament de Filologia Espanola - Laboratori de Fonètica,
Facultat de Filosofia i Lletres, Universitat Autònoma de Barcelona