A STATISTICAL APPROACH TO SPANISH AMERICAN PHONOLOGICAL UNITS

MICHELINA GUIRAO
MIGUELINA GUIRAO
Laboratorio de Investigaciones Sensoriales CONICET, Universidad de Buenos Aires
CC 83 1433 Buenos Aires, Argentina

ABSTRACT
In this paper we report an account of Spanish American phonological units. The sample consisted of 74460 syllables distributed in 43306 words. The frequency of occurrence of phonemes are presented each one labeled according to articulatory features. Dental have an incidence of almost four times more than velars. Palatals have a low frequency. Voiced phonemes gave higher figures than unvoiced ones. A table is presented with a ranking order of the first 50 syllables. Thirty one of these syllables are also words. Included is the percent incidence of each of the first 50 syllables in initial and final word position. Type CV is equally distributed in both positions and CVC tend to terminate words. An observation is made on the most frequent articulatory combinations encountered at both extremes of words.

INTRODUCTION
In this work we intent to address the problem of segmentation of morphemic units in continuous speech on the basis of statistical data. Other ares were to provide useful information for spectrogram reading and for cross language comparison. Our system consists of five vowels and seventeen consonants. All but four phonemes /j/ (spelled y and J), /w/ (ch), /r/ (/1/) and /l/ (6) can be represented with orthographic symbols.

In spoken Spanish the syllabic structure predominates over the morphemic one. It is also known that CV syllables are more than 50% of all syllabic types. The expansion of this pair produces CV+C, CV+V, CV+CV which add up to 90% of all syllabic types. CV constitutes more than 50% of all syllabic types. CV+V+C which adds up to almost four times more than velars. Palatales have a low frequency. Voiced phonemes gave higher figures than unvoiced ones. A table is presented with a ranking order of the first 50 syllables. Thirty one of these syllables are also words. Included is the percent incidence of each of the first 50 syllables in initial and final word position. Type CV is equally distributed in both positions and CVC tend to terminate words. An observation is made on the most frequent articulatory combinations encountered at both extremes of words.

DISTRIBUTION OF VOWELS AND CONSONANTS

In Figure 1 we present the five vowels ordered according to tongue height.

Voiced consonants added together make more than three fourths of all consonants. The most frequently pronounced consonants are dentals /d t s n l/ which summate 27%, labials /m p b/ 10% and velar /k/ 7%. These consonantal components represent 66.3% of the total sample but are only 3.3% while those of two are 34.8% of the total consonant occurrence. The most frequently pronounced consonants are dental /s/ /t/ and /l/ which summate 15%, central /e/ and back closed /o/ follow with relatively small differences. Closed vowels are less used. In Table 1 we present the first 50 items which represent 66.3% of the total sample but of five of the different syllables. Thirty seven are formed by a CV pair, five by CV, five by V and four by VC. The most frequently pronounced consonants are dental /s/ /t/ and /l/ which summate 27%, labials /m p b/ 10% and velar /k/ 7%. These consonantal components are mainly combined with the three called strong vowels /e a o/. Thirty seven are also word boundary. Syllables Percent

Table 1. The first fifty syllables (with asterisk are also words)

In Table 1 we present the first 50 items which represent 66.3% of the total sample but only 7% of the different syllables. Thirty seven are formed by a CV pair, five by CV, five by V and four by VC. The most frequently pronounced consonants are dental /s/ /t/ and /l/ which summate 27%, labials /m p b/ 10% and velar /k/ 7%. The consonantal components are mainly combined with the three called strong vowels /e a o/. Thirty one of the units listed in Table 1 are monosyllabic words. These units appear with a relative much higher incidence than multisyllabic Spanish pronunciation. Voiced is another apparent feature vowels and voiced consonants added together make more than three fourths of the total occurrences. Comparison with the Peninsular Spanish pronunciation gave minor deviations which are correlated with some differences in the phonological system /7/.

FREQUENTLY REPEATED SYLLABLES
It is noted that the high incidence of certain phonemes results from a group of frequently repeated syllables.

In Figure 2 we present the distribution of Spanish consonants by articulatory configurations.

Taking as a reference frontal half closed /e/ which is the most used sound, scoring 15%, central /a/ and back closed /o/ follow with relatively small differences. Closed vowels are less used. In Figure 2 consonants are distributed according with place, voicing and manner of articulation. The most frequently pronounced consonants are dental /s/ /t/ and /l/ which summate 27%, labials /m p b/ 10% and velar /k/ 7%. The consonantal components are mainly combined with the three called strong vowels /e a o/. Thirty one of the units listed in Table 1 are monosyllabic words. These units appear with a relative much higher incidence than multisyllabic Spanish pronunciation. Voiced is another apparent feature vowels and voiced consonants added together make more than three fourths of the total occurrences. Comparison with the Peninsular Spanish pronunciation gave minor deviations which are correlated with some differences in the phonological system /7/.

In general we observe that a large number of sounds are produced at the front of the mouth. Two frontal vowels plus dentals and labials summate over three thirds of all phonemes occurrences. Voicing is another apparent feature vowels and voiced consonants added together make more than three fourths of the total occurrences. Comparison with the Peninsular Spanish pronunciation gave minor deviations which are correlated with some differences in the phonological system /7/.
To count syllabic sounds located either at the onset or at the offset of words, we took again a sample formed by the first 50 syllables. Most of these items resulted from bisyllabic words. The range of scores is in the order of 1 to 10, 62 to 12.1 and velars 5%. Among the consonantal sounds listed in the first set we see that vowel alone or in VC pair are the most frequently produced sounds initiating syllables. Open vowel /a/ 9.8% is first. This initial vowels with exception of /a/ and /e/ were not registered. Fricative /s/ was also missing but another dental in /ra re/ and a palatal in /at at/ entered in Table 1. We also observed a clear predominance of dentals. This category covers about two thirds of the total percentage listed in the table. The first eleven sounds are strong vowels. The third category found are dental combined with strong vowels /b/ in /ba be bien bre/ follow with 4.7 and 3% respectively. Among the velars which are limited both sides being three times more frequent than at the end of words. Velar /k/ is similarly distributed at both extremes. Monosyllabic words that are part of larger lexical segments eleven words /es me e/ appeared more frequently at the beginning than when forming part of larger words. Non syllables which are part of larger lexical units are more frequently occurring in initial than in final position.

**SYLLABLES AT BOTH MARGINS OF WORDS**

To count syllabic sounds located either at the onset or at the offset of words, we took again a sample formed by the first 50 syllables. Most of these items resulted from bisyllabic words. The range of scores is in the order of 1 to 10, 62 to 12.1 and velars 5%. Among the consonantal sounds listed in the first set we see that vowel alone or in VC pair are the most frequently produced sounds initiating syllables. Open vowel /a/ 9.8% is first. This initial vowels with exception of /a/ and /e/ were not registered. Fricative /s/ was also missing but another dental in /ra re/ and a palatal in /at at/ entered in Table 1. We also observed a clear predominance of dentals. This category covers about two thirds of the total percentage listed in the table. The first eleven sounds are strong vowels. The third category found are dental combined with strong vowels /b/ in /ba be bien bre/ follow with 4.7 and 3% respectively. Among the velars which are limited both sides being three times more frequent than at the end of words. Velar /k/ is similarly distributed at both extremes. Monosyllabic words that are part of larger lexical segments eleven words /es me e/ appeared more frequently at the beginning than when forming part of larger words. Non syllables which are part of larger lexical units are more frequently occurring in initial than in final position.

**REFERENCES**


---

**SYLLABLES AT BOTH MARGINS OF WORDS**

To count syllabic sounds located either at the onset or at the offset of words, we took again a sample formed by the first 50 syllables. Most of these items resulted from bisyllabic words. The range of scores is in the order of 1 to 10, 62 to 12.1 and velars 5%. Among the consonantal sounds listed in the first set we see that vowel alone or in VC pair are the most frequently produced sounds initiating syllables. Open vowel /a/ 9.8% is first. This initial vowels with exception of /a/ and /e/ were not registered. Fricative /s/ was also missing but another dental in /ra re/ and a palatal in /at at/ entered in Table 1. We also observed a clear predominance of dentals. This category covers about two thirds of the total percentage listed in the table. The first eleven sounds are strong vowels. The third category found are dental combined with strong vowels /b/ in /ba be bien bre/ follow with 4.7 and 3% respectively. Among the velars which are limited both sides being three times more frequent than at the end of words. Velar /k/ is similarly distributed at both extremes. Monosyllabic words that are part of larger lexical segments eleven words /es me e/ appeared more frequently at the beginning than when forming part of larger words. Non syllables which are part of larger lexical units are more frequently occurring in initial than in final position.

**REFERENCES**

