Adam Mickiewicz University – Poznań 1993
*Papers and Studies in Contrastive Linguistics vol. XXVII*

# AUTOMATIC PHONETIC TRANSCRIPTION OF SPANISH TEXT: NATIVE SPEAKER DIALECTS AND FOREIGN SPEAKER SIMULATION

JAMES L. WYATT
*Florida State University*

This paper discusses a personal computer program which transforms textual input in the Spanish language into a phonetic transcription of segmental phonemes.

The first part of this paper will discuss the program from a user's viewpoint. That is, what the program produces. Next will be a discussion of a general nature about the program itself, outlining the tasks performed to turn text into phonetic transcription. Finally, I shall describe some features of the Spanish language as related to this application of computer processing.

I shall not dwell on minute details of Spanish phonology, nor shall I discuss the computer program in any technical sense. My hope is that some will find this computer application of interest and that perhaps one or more may be inspired to put the computer to work in their particular field of interest, if they have not already done so.

## THIS PROGRAM AND THE USER

The program I have written will produce transcriptions of more or less typical varieties of Castillian Spanish and/or American Spanish.

In addition, the user may select certain optional Iberian and/or Spanish-American dialectal features, and may even request a phonetic transcription of a simulated native speaker of English who commits phonological errors in Spanish because of interference from English.

Simulations of native speakers of languages other than English could be built into the program.

The computer program, written in a version of the SNOBOL programming language for IBM and IBM-compatible personal computers, contains a sub-

stantial knowledge of the principal phonological rules for segmental phonemes in Spanish and a limited knowledge of conflicting phonological rules in English.

These latter rules are the basis for simulating predictable errors by a native speaker of English.

The individually selectable optional phonological rules for dialectal features of Iberian Spanish make it possible to produce a transcription generally typical of southern Spain. One optional rule allows transcriptions to include a variant allophone of the (s) phoneme typical in some areas of northwestern Spain. Additional optional rules may be incorporated in the program with little difficulty.

The program utilizes a preselected set of rules to transform a transcription of Castillian Spanish into a transcription typical of at least some areas in the Americas.

The user may select on a one-by-one basis from a set of American dialectal rules to create transcriptions including several dialectal features found in American Spanish.

Although this program was developed for experimental programming purposes rather than for any particular practical application, several possible uses based on modifications of this program come to mind.

Perhaps the most obvious application of a program of this nature, modified to meet requirements, would be in the teaching of Spanish phonology to learners of the language.

Related to that application is training to change from one dialect to another, for whatever reason. Native or non-native speakers might want to change dialectal features on a more or less permanent basis for social or other reasons. Or speakers of Spanish might need training to gain facility in switching back and forth from one dialect to another, as in acting for the stage, cinema, or television.

Still another possible application lies in the area of text to speech processing. This program's output could be used as input by a program capable of generating the audible allophones of the Spanish language.

Although this program does not at present output pitch patterns, the information needed to create pitch patterns is derivable from the transcription itself if the input does not contain exclamations or yes-no questions. It would be a simple matter to accommodate these exceptions.

We have discussed what the program does in terms of input and output. Now, let us turn to what the program does to make output possible.

## THE TEXT TO TRANSCRIPTION PROGRAM

At the beginning of execution, the program requests the user's options from the screen. The user responds on the keyboard with "Y" or "N" (for "yes", or "no") to each option offered.



The computer program reads data contained in an input data file and writes transcriptions to an output file, or with a simple change the program will accept input from the keyboard and write transcriptions to the screen.

Input may be single sentences or entire paragraphs.

After requesting the user's options, the program reads a record of data (from a disk file or the keyboard) and proceeds to the following:

1. Convert all input to lowercase and assign stress to lexical items. (Most Spanish words with stress are unmarked in natural text; written stress marks occur only when stress is not predictable).

2. Remove all spaces and punctuation from the text, inserting "#" for sentence final pause, and "/" for internal pauses assumed for commas, colons, and semicolons. (Of course, readers of the natural text might insert pauses where punctuation does not occur, and they might ignore punctuation and not pause).

3. Substitute phonetic symbols for alphabetic characters, make required allophonic adjustments because of phonetic environment and progressive and regressive assimilation processes, and break the string of phonetic symbols into syllables. At the end of this step, the program's memory contains two strings representing Castillian Spanish. One string is a limited transcription consisting only of characters found in the standard IBM character set. The other string provides for special phonetic characters designed for this purpose.

4. Transform the two strings produced for Castillian Spanish into Iberian dialectal transcriptions, according to user-selected options.

5. Transform the two Castillian Spanish strings into a variety of American Spanish, if selected by the user.

6. Transform the two American Spanish, strings according to individually selectable dialectal features.

7. Produce, if selected, a transcription of simulated pronunciation of Spanish by a speaker of American English who transfers features of English phonology to Spanish. The input for this section is the original input record marked for stress, the same as for the program section that produces Castillian transcriptions.

Output may occur in any of the steps producing transcriptions and is according to the options selected by the user.

Among the user options are those to produce a limited transcription using only the IBM character set, and/or a transcription in a phonetic alphabet. Options allow the output to be labelled according to type of transcription.

The limited transcription makes use of lowercase and some uppercase roman alphabet characters and some of the IBM upper ASCII character set # 2. Some of these latter characters are: Ã, Ú, Ó, #, Õ, Š, and ›. The first three characters represent voiced fricative allophones of the /bdg/ stop phonemes, respectively. The sign # represents the sole allophone of the voiced

palatal lateral phoneme, Õ and Š present the voiceless and voiced allophones, respectively, of the interdental fricative phoneme /Õ/, and ⟩ represents the voiceless palatal affricate phoneme.

The full phonetic transcription uses key codes for most lowercase roman characters and for a number of special phonetic characters. The full phonetic transcription specifications have not been implemented for the computer screen, but are downloaded to a dot matrix printer before printing the program's output.

Most of the special phonetic characters are based on characters used by one or another American author of works on phonology and do not adhere to the International Phonetic Alphabet. Several characters are somewhat unique because of the limitations imposed by a matrix eight dots high and 11 dots across.

Three characters, those for the fricative allophones of the voiced stop phonemes /bdg/, are slightly different from the usual characters for the sake of greater legibility.

Four characters are unlike any ever seen by the author. These consist of the roman letters *s* and *z* with diacritics above, graphically suggesting the position of the tongue.

Two characters' diacritic marks rise from the top of the *s* and *z* and make a left turn. These represent the voiceless and voiced allophones of the predorsal /s/, widely used in southern Spain and most of the Americas. The "left turn" is intended to show the position of the blade or predorsum of the tongue which contacts the alveolar region.

The diacritic over two other /s/ allophones curls up and to the right, forming a small hook. This is intended to represent a strongly retroflexed apex in the production of the /s/ phoneme's allophones typical of some areas of northwestern Spain, the allophones of the so-called "fat" or "thick" /s/.

These unusual phonetic signs were inspired by the diacritic used by at least one author to represent the typically Castillian apico-alveolar /s/. That diacritic is a straight vertical mark on top of the symbols for that phoneme's allophones. The author of this paper interpreted the vertical mark as the apex of the tongue at a right angle to the alveolar ridge.

The phonetic character definitions for the printer were created by a program the author wrote in the BASIC programming language. Before printing the output of the phonetic transcription program, the BASIC program is run with the printer on line to receive the character specifications.

Also, before printing the phonetic transcriptions produced by the SNOBOL program, a second very short program in SNOBOL processes the disk transcription files to set margins and line length. This program assures line breaks at the end of a syllable.



Samples of transcription output are contained in figures 1 and 2. Figure 1 is the multiple output of one sentence, taking advantage of all possible options offered by the program. Figure 2 is the output of one paragraph using special phonetic characters for Castillian, American, and dialectal American Spanish.

Figures 3a and 3b are charts of symbols used in full phonetic transcriptions, and figures 4a and 4b are charts of symbols used in limited IBM character set transcriptions.

The SNOBOL phonetic transcription program presently consists of slightly more than 700 statements. The program, with comment lines, amounts to approximately 30.000 characters, and was written in version 2.18 of SNOBOL4+ by Catspaw, Inc. (Salida, Colorado 81201, U.S.A.).

Using a 16 MHz 80386 IBM-compatible AT machine, the program's compilation time is less than three seconds, and the execution time to process the paragraph in the three versions in figure 2 was about 95 seconds, including the time necessary to respond from the keyboard to 30 screen prompts. Slightly more than 4,000 statements were executed to produce this output.

Before leaving this discussion of the transcription program itself, I should like to recommend SNOBOL4+ as ideal for this type of application. The SNOBOL language is known for its pattern matching and string manipulation features, and that is what this text-to-transcription program is about. SNOBOL4+ is a large model version of SNOBOL for the personal computer and has a number of useful features not found in standard SNOBOL.

## THE SPANISH LANGUAGE AND THIS APPLICATION

The Spanish language lends itself well to the automatic production of phonetic transcription.

This is especially true as it applies to assigning stress to words of text and dividing the stream of speech symbols into syllables, the mechanics of which we shall now discuss.

Spanish has only two levels of stress. Orthography marks the great majority of exceptions to these two simple rules concerning stress:

1. If a written word ends in a vowel or *n* or *s*, and if stress is on the next to last syllable, a written stress mark is not used, and

2. If a written word ends in any consonant except *n* or *s*, and if stress is on the last syllable, a written stress mark is not used.

If the phonetic facts are exceptions to these rules, then orthography requires a written acute accent mark over the stressed vowel.

The computer program described here makes use of these rules and deals with words not bearing written stress on stressed vowels in next-to-last and last syllables.

Some words to which the rules apply in citation form, however, do not normally bear stress in an utterance. These words fall into relatively small sets of words according to syntactic function.

The program contains these normally weakly stressed words, and it checks each word of input to see if it should be subjected to the stress assignment routine.

The syntactic categories of weakly stressed words include: definite and indefinite articles, relative pronouns, direct and indirect object pronouns, prenominal possessives, simple prepositions and conjunctions, and obligatory contractions of preposition plus definite article. The group of unstressed words numbers 62 items.

A very slight problem exists, however, in that several unstressed words look like words that are stressed. The spelling "como" may be either a conjugated form of a verb, and thus bear stress on the next to last syllable (according to the rules), or the spelling may be that of an unstressed conjunction or an adverb. And the spellings "un" and "una" may represent stressed cardinal numbers or unstressed indefinite articles.

There are several other possibilities of erroneously failing to mark a stressed syllable, but the functional load of the several ambiguous items is greatly tilted in favor of accepting the group of unstressed words at face value.

Two alternatives to rule out errors would be to include in the program a query to the user at run time to determine whether one of the assumed unstressed words is, in fact, stressed. The other alternative would be to parse each sentence containing one of the few possibilities of error to determine whether stress should be assigned. This would be a truly major undertaking, probably far beyond the capabilities of a single individual during a reasonable period of time.

After stress has been assigned to all words in the input record, the computer program removes all blanks between words, inserts phonetic characters for alphabetic characters, makes allophonic adjustments, and proceeds to break the string of phonetic symbols into syllables, making use of a rule concerning a single consonant between vowels and several rules concerning consonant clusters.

Any single consonant between vowels begins a syllable. There are 12 consonant clusters of two that may begin syllables. Clusters of three consonants may end in a cluster of two that begins a syllable, with the first of the group terminating the previous syllable; if not, the last consonant of the cluster of three begins a syllable, and the first and second consonants end the preceding syllable. A group of four consonants always splits in the middle, the first two ending a syllable, the last two (always one of the 12 two-consonant clusters previously mentioned) beginning a syllable.



Also, as is true at the word level, vowels and semivowels in the string of phonetic symbols combine to form diphthongs and triphthongs, and other vowel groups may or may not lie within a single syllable. And, as at the word level, a combination of one stressed and one unstressed identical vowel, generally collapses into a single stressed vowel, while two stressed identical vowels do not collapse.

Many word boundaries are obliterated in the process of syllable segmentation. Some syllables do coincide with word beginnings or endings, but many do not.

Let us now consider conversion from one Spanish dialect to another and selecting dialectal features.

The Castillian variety of Spanish includes two phonemes not typical of southern Spain and Spanish America, the interdental fricative phoneme /Ö/ with voiceless and voiced allophones, and the voiced palatal lateral phoneme /#/ (from the ASCII symbols in figure 4a).

In addition, the point of articulation of the /s/ phoneme's allophones is substantially different in Castillian Spanish and the language in southern Spain and most of Spanish America, as was noted above when phonetic characters were discussed.

To convert from Castillian Spanish to a representation of the language typical of southern Spain, the program simply substitutes appropriate /s/ allophones for those of /Õ/, the /y/ allophone for that of /#/, and the variant /s/ allophones for the Castillian allophones of /s/.

In order to produce a transcription of Spanish considered to be typical of several general regions of the Americas, the program utilizes a set of rules for American Spanish based on the optional rules for the Spanish of southern Spain.

Some may wonder why the output of the Spanish of southern Spain was not used to produce American Spanish. The reason is that the dialectal features for Iberian Spanish are selected by the user. Those options may be chosen or rejected on an individual basis. The rules for American Spanish are a set of rules used by the program.

Dialectal features available to alter typically American Spanish include the following (some symbols are from the ASCII symbol chart in figure 4a):

1. Word final velar allophone of /n/ in syllable initial position and before pause.
2. Aspiration of /s/ in syllable final position.
3. Affricate allophone of /y/.
4. Use of [v] for /b/ in words spelled with letter *v*.
5. Fricative allophone for trilled alveolar /R/ and alveolar /r/ in word initial and final position.
6. Tap allophone for /r/ in word initial position.

7. Open allophone of /e/ for closed allophone.
8. Substitution of /h/ for /x/.
9. Substitution of word final [o] for [Uo].

The final section of the computer program produces a transcription of a simulated native speaker of English without a good knowledge of Spanish phonology. The bad rendition is based on preservation of word boundaries, a lack of use of the fricative allophones of /b/, /d/, /g/, aspiration after the syllable initial voiceless stops /ptk/, syllabification according to English consonant cluster groupings, treating semivowels as full vowels, the tap /r/ phoneme for the trilled /R/ phoneme, and using schwa for word final unstressed /a/ in syllables following a syllable with a stressed vowel.

These faults are readily observed in the Spanish of many Americans who understand Spanish quite well but who are unaware of their phonological problems with the language.

This computer application dealing with Spanish has been challenging. At some time in the future I plan to optimize the code, implement phonetic characters for the screen, and perhaps add more dialectal features. I have wanted to share the results at this stage with those with similar interests.

## SAMPLE OF RUN WITH ALL OPTIONS

Input:
Los hispanistas se reunieron en Dinamarca.
Input with marked stress:
los hispanístas se reuniéron en dinamárca #
Castillian: ASCII Font
lo - sis - pa - nís - tas - se - Reu - nié - ro - nen - Di - na - már - ka #
Castillian: Phonetic Font
lo - śiś - pa - niş - taś - śɛ - r̄ɛu̯ - nié - ro - nẽn̨ - di - nã - már - ka #
Dialectal Iberian: ASCII Font
lo - sis - pa - nis - tas - se - Reu - nié - ro - nen - Di - na - már - ka #
Dialectal Iberian: Phonetic Font
lo - śiś - pa - niş - taś - śɛ - r̄ɛu̯ - nié - ro - nẽn̨ - di - nã - már - ka #
General American: ASCII Font
lo -sis - pa - nís - tas - se - Reu - nié - ro - nen - Di - na - már - ka #
General American: Phonetic Font
lo - śiś - pa - níş - taś - śɛ - r̄ɛu̯ - nié - ro - nẽn̨ - di - nã - már - ka #



Dialectal American: ASCII Font
lo - sih - pa - níh - tah - se - Reu - nié - ro - Nen - Di - na - már - ka #
Dialectal American: Phonetic Font
lo - ših - pa - níh - tah - śɛ - r̄ɛu̯ - nié - ro - ŋen̨ - di - nã - már - ka #
"Gringo" Spanish: ASCII Font
los i - spa - ní - stas se re - u - ni - é - ron en di - na - már - k'a #
"Gringo" Spanish: Phonetic Font:
los i - spa - ní - stas se re - u - ni - é - ron en di - na - már - k'a #

Figure 1

Sample Run with Restricted Options

Castillian: Phonetic Font

e - nũ - nag - θé - śo - de - kom - ʃián - θa / dé - śoś - ke - pro - ƀó - ka - la - ʃa - mi - lia - ri - đá - di - kom - bi - bén - θia - de - loź - bal - ne - a - rióś / la - eɱ - ʃér - ma - dɛl - ko - ra - θun - mɛ - ʃi - rió - śu - mál / koɳ - tó - doź - loź - de - śo - fo - ka - θió - neś / bio - lén̨ - taś - pal - pi - ta - θió - neś / bér - ti - goś / śiŋ - ko - peś / ko - láp - śoś / eŋ - ke - śe - bé - ɬe - gár - laú̯ - ti - ma - ó - ra # mién - tra - śa - bla - ba / la - mi - rá - ba - yó - a - teɳ - ta - mẽɳ - te # é - rau̯ - nã - mu - xér - ko - mo - de - tréiɳ - tai - θiŋ - ko - a - tréiɳ - tai - śéi - śá - ñoś / eş - tro - pe - á - da - po - rɛl - pa - de - θi - mién̨ - to / al - mẽ - noştál - krei / au̯ŋ - ke - pro - lon - gá - do - e - lɛg - śá - mẽn / em - pe - θé - a - śu - po - nér - kɛu̯ - bié - śe - ál - go - má - śa - ɬá - de - lo - ʃi - śi - ko - en - śu - ʃu̯i - na #

General American: Phonetic Font

e - nũ - nag - śé - śo - de - koɱ - ʃián - śa / dé - śoś - ke - pro - bó - ka - la - ʃa - mi - lia - ri - dá - di - kom - bi - bén - śia - de - loż - bal - ne - a - rióś / la - eɱ - ʃér - ma - dɛl - ko - ra - śón - mɛ - ʃɛ - fi - rió - śu - mál / koɳ - tó - doż - loż - de - tá - yeż - de - śo - ʃo - ka - śió - neá / bio - lén̨ - taś - pal - pi - ta - śió - neś / bér - ti - goś / / śiŋ - ko - peś / ko - láp - śoś / eŋ - ke - śe - bé - ye - gár - laú̯ - ti - ma - ó - ra # mién - tra - śa - blá - ba / la - mi - rá - ba - yó - a - teɳ - ta - mẽɳ - te # é - rau̯ - nã - mu - xér - ko - mo - de - tréin - tai - sin - ko - a - tré - in - tai - śéi - śá - ñoś / eş - tro - pe - á - da - po - rɛl - pa - de - śi - mién̨ - to / al - mẽ - noş - tái - krei / au̯ŋ - ke - pro - loŋ - gá - do - e - lɛg - śá - mẽn / em - pe - śé - a - śu - po - kɛu̯ - bié - śe - ál - go - má - śa - yá - de - l o - fi - śi - ko - en - śu - ʃu̯i - na #

Dialectal American: Phonetic Font

ɛ - ŋũ - ŋag - śé - śo - dɛ - koɱ - ʃián - śa / dé - śoh - kɛ - pro - vó - ka - la - ʃa - mi - lia - ri - dá - di - kon - vi - vén - śia - dɛ - loh - bal - nɛ - a - mál / la - eɱ - ʃér - ma - dɛl - ko - ra - són - mɛ - ře - fi - rió - śu - mál / koɳ - tó - doh - loh - dɛ - tá - ĵeh - dɛ - śo - ʃo - ka - śio - nɛh / vio - lén̨ - tah - pal - pi - ta - śió - nɛh / vé - ti - goh / sin - ko - pɛh / ko - láp - śoh / ɛŋ - kɛ - śɛ - vé - ĵe - gár - laú̯ - ti - ma - ó - ra # mién̨ - tra - śa - blá - ba / la - mi - rá - ba - ĵo - a - tɛɳ - ta - mén̨ - tɛ # é - rau̯ - nã - mu - hér - ko - mo - dɛ - tréiɳ - tai - śiŋ - ko - a - tréiɳ - - tai - śéi - śá - ñoh / ɛh - tro - pɛ - tro - pɛ - á - da - da - po - rɛl - pa - dɛ - śi - mién̨ - to / al - mẽ - noh - rál - - krɛi / au̯ŋ - mẽ - noh - tál - krei / au̯ŋ - kɛ - pro - loŋ - gáo - ɛ - lɛg - áá - mẽŋ / ɛm - pɛ - śé - a - śu - po - nér - kɛu̯ - bié - śɛ - ál - go - má - śa - ĵá - dɛ - lo - fi - śi - ko - ɛn - śu - řu̯i - na #

Figure 2

Consonant Allophones: Phonetic Character Set

| MANNER OF ARTICULATION | | POINT OF ARTICULATION | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | BILABIAL | LABIO DENTAL | INTER DENTAL | DENTAL | ALVEOLAR | PALATAL | VELAR | BILABIO VELAR | GLOTTAL |
| STOPS | VL | p | | | t | | | k | | |
| | VD | b | | | d | | | g | | |
| SLIT FRICATIVES | VL | | f | θ | | | | x | | h |
| | VD | b | v | δ | ɖ | ř | y | g | w | |
| GROOVE FRICATIVES | VL | | | | ʂ | ʃ ʂ ʃ | š | | | |
| | VD | | | | ʐ | ʒ ʐ ʒ | ž | | | |
| AFFRICATES | VL | \ | | | | | ĉ | | | |
| | VD | | | | | | ĵ | | | |
| NASALS | VD | m | ɱ | | ɳ | n | ñ | ŋ | | |
| LATERALS | VD | | | | ɭ | l | ɫ | | | |
| TAPS | VL | | | | | | | | | |
| | VD | | | | | r | | | | |
| TRILLS | VL | | | | | | | | | |
| | VD | | | | | r̄ | | | | |

Figure 3a

Vowel Allophones: Phonetic Character Set

| | | FRONT | | CENTRAL | | BACK | |
|---|---|---|---|---|---|---|---|
| | | ORAL | NASAL | ORAL | NASAL | ORAL | NASAL |
| HIGH | | i | ĩ | | | u | ũ |
| MID | CLOSED | e | ẽ | | | o | õ |
| | OPEN | ɛ | ɛ̃ | | | ɔ | ɔ̃ |
| LOW | | | | a | ã | | |

Figure 3b



Consonant Allophones: Phonetic Character Set

| MANNER OF ARTICULATION | | POINT OF ARTICULATION | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | BILABIAL | LABIO DENTAL | INTER DENTAL | DENTAL | ALVEOLAR | PALATAL | VELAR | BILABIO VELAR | GLOTTAL |
| STOPS | VL | p | | | t | | | k | | |
| | VD | B | | | D | | | G | | |
| SLIT FRICATIVES | VL | | f | θ | | | | x | | h |
| | VD | ß | v | δ | Ø | | y | r | w | |
| GROOVE FRICATIVES | VL | | | | | s | | | | |
| | VD | | | | | z | | | | |
| AFFRICATES | VL | | | | | | ¢ | | | |
| | VD | | | | | | j | | | |
| NASALS | VD | m | | | | n | ñ | N | | |
| LATERALS | VD | | | | | l | £ | | | |
| TAPS | VL | | | | | | | | | |
| | VD | | | | | r | | | | |
| TRILLS | VL | | | | | | | | | |
| | VD | | | | | R | | | | |

Figure 4a

Vowel Allophones: ASCII Character Set

| | FRONT | CENTRAL | BACK |
|---|---|---|---|
| HIGH | i | | u |
| MID | e | | o |
| LOW | | a | |

Figure 4b