The Hablarte Text-to-Speech System for Spanish

Helen E. Karn

Office of the Dean, Georgetown College, Box 571003, Georgetown University, Washington, DC 20057-1003
Email: kmnh@gunet.georgetown.edu

Abstract: This paper presents an overview of Hablarte, a prototype text-to-speech (TTS) system for unrestricted Spanish texts. The focus of this paper is on the first five text processing modules. Each module performs a separate text processing function whose input is the output of the previous module. This allows each module to be evaluated independently. The order of these modules is phonological phrase boundary placement, grapheme-to-phoneme conversion, primary stress assignment, syllabification, resyllabification, phoneme-to-allophone conversion, intonation contour assignment, and rhythm assignment. The output is a set of phonetic symbols and prosodic notations which are the input to the speech synthesizer. Work is currently underway to recode these text algorithms in the Cold Fusion Markup Language as the front end for a Spanish diphone synthesizer for TTS synthesis in World Wide Web applications.

TEXT DIVISION AND PHONOLOGICAL PHRASE BOUNDARY PLACEMENT

The input to the first module is an electronic text in standard Spanish orthography, which is then divided into sentences (based on punctuation) and words. Three small lexicons (function words, auxiliaries, and verbs) are used to assign a grammatical category to each input word. If an input word is not found in one of the dictionaries, it is assigned the category of content word. A compound word lexicon and a lexicon of prefixed words are used to flag input words for special processing in the stress assignment and syllabification modules. Content words are also checked for the adverbial suffix -mente; if present, the word is divided into its base and suffix, so that both undergo primary stress assignment. Preliminary phrase boundaries are then placed by a modification of an f-group parser for English (1) and readjusted for shorter phrases and phrases not ending in punctuation (2).

GRAPHEME-TO-PHONEME CONVERSION

The second module converts the graphemes of the input to phonemes. The number of rules in this module is relatively few. The grapheme <x> is converted to the consonant cluster /ks/ in all forms of words derived from México (e.g. mexicano, mexicanas) and to the phoneme /x/ elsewhere (e.g. Ximénez). The digraphs <ch, ll, rr> are converted to the phonemes /ʃ, l, r/, respectively. The grapheme <c> becomes the fricative /s/ before front vowels and /k/ elsewhere. Similarly, the grapheme <g> is converted to the fricative /x/ before front vowels (e.g. general) and /g/ elsewhere. (Hablarte generates /s/ for all instances of <c> and <z>, but a rule to produce /θ/ before front vowels can easily be added for Castilian Spanish speakers whose dialects show that distinction.) Acute accents remain to facilitate primary stress assignment, but <i> is converted to /wi/ if it follows /g/ and precedes a front vowel. The graphemes <i, u> become the glides /j, w/, respectively, when they precede or follow a vowel which does not contain a graphemic accent.

PRIMARY STRESS ASSIGNMENT

The third module assigns primary stress to one vowel per word; secondary stress is not assigned. Each input word is first scanned for an acute accent, as in /löχiko/. If the input word does not contain an acute accent, Hablarte assigns primary stress to the vowel of the penultimate syllable by converting the vowel to its accented counterpart if the word ends in a vowel, /n/ or /s/, as in /ágwa/ and /kásas/. Otherwise, the vowel of the ultimate syllable is assigned primary stress, as in /ablár/. To find the penultimate/ultimate vowel, Hablarte reverses the word and replaces the desired vowel (first or second) before reversing the word (i.e. /kimono/ →/onomik/ →/onómik/ →/kimóno/).
SYLLABIFICATION

The fourth module divides each input word into syllables, using a correction of the Hualde syllabification algorithm (3). In the revised algorithm (4), the order of the syllabification rules are: Node projection, Complex nucleus, CV rule, Complex onset, Consonantization, Coda rule, Complex coda, /s/-Adjunction, and Coda simplification. For each word in the input text, the vowel is taken as the starting point. Glides and consonants preceding and following the vowel are then attached to the vowel to form the syllable, first from the left (forming the onset) and then to right (forming the coda). The phonotactics of Spanish are used to constrain onset and coda formation. Syllabification occurs from left to right, beginning with the first vowel in each word and ending when all vowels have been processed or when the end of the input word is reached.

RESYLLABIFICATION

The fifth module readjusts the syllable boundaries between the words within a phonological phrase. Resyllabification does not occur across phonological phrases. Morphemic boundaries, such as {sub-}, are maintained as syllable boundaries because Hablarte generates Andante (careful) speech. If no morphemic boundaries are present, the consonant of a coda, if present, is moved to the onset of the following syllable if that syllable does not have an onset. If the following syllable does have an onset but is identical to the preceding coda, then the consonant in the coda of the first syllable is deleted. If the consonant is a continuant, then the consonant of the second syllable is marked for longer duration. (Most, if not all, codas are continuants because of the strict constraints on Spanish codas.) Similarly, if a syllable has no coda, the following syllable contains no onset, and the vowel nuclei are the same, the two syllables are merged, with the vowel marked for longer duration. As with syllabification, processing occurs from left to right, beginning with the last coda of the first word and ending when all word-final codas have been processed or the end of the input phrase is reached.

CURRENT AND FUTURE WORK

At present, Hablarte’s text processing algorithms are being recoded in the Cold Fusion Markup Language so that it may be accessible through the World Wide Web. In addition, Hablarte’s original formant-based speech synthesizer is being replaced by the Spanish male diphone database of the Mbrola project (5). For updates to the work presented in this paper, please access the following URL: http://www.georgetown.edu/users/karnh/home

REFERENCES


