FOCUS AND ACCENT IN A DUTCH TEXT-TO-SPEECH SYSTEM

Joan L.G. Baart
Phonetics Laboratory, Department of General Linguistics
Cleveringaplaats 1, P.O.Box 9515
2300 RA Leiden, The Netherlands

Abstract

In this paper we discuss an algorithm for the assignment of pitch accent positions in text-to-speech conversion. The algorithm is closely modeled on current linguistic accounts of accent placement, and assumes a surface syntactic analysis of the input. It comprises a small number of heuristic rules for determining which phrases of a sentence are to be focussed upon; the exact location of a pitch accent within a focussed phrase is determined mainly on the basis of the syntactic relations holding between the elements of the phrase. A perceptual evaluation experiment showed that the algorithm proposed here leads to improved subjective speech quality as compared to a naive algorithm which accents all and only content words.

1. Introduction

This paper deals with the prosodic component of a text-to-speech system for Dutch, more in particular with the rules for assigning pitch accents (sentence accents) to words in an input sentence. Whereas other work on accent rules for Dutch speech synthesis (Kager & Quené, 1987) did not assume a syntactically analysed input, I will here work from the assumption that the text-to-speech system has a large dictionary as well as a syntactic parser at its disposal.

The paper is organized as follows: in section 2 I shortly introduce the notions focus and (pitch) accent as I will be using them; as my framework, I will choose the Eindhoven model of Dutch intonation ('t Hart & Cohen, 1973; 't Hart & Collier, 1975) in conjunction with Gussenhoven's (1983) accent placement theory. In section 3 I discuss the rules that connect a domain of focus to an accent on a particular word. The assignment of focus domains is dealt with in section 4. At the end of this section I summarize my proposals in the form of an accent assignment algorithm. In section 5 I present some results obtained in a perceptual evaluation of this algorithm.

2. A two-stage model of accent placement

Work on Dutch intonation at the Institute for Perception Research (IPO) in Eindhoven has resulted in an inventory of elementary pitch movements that make up the occurring Dutch intonation contours ('t Hart & Cohen, 1973; 't Hart & Collier, 1975). The phonetic characteristics of these pitch movements are known precisely, and this knowledge can be used in the synthesis of natural-sounding Dutch intonation contours. It was found that some of these elementary pitch movements cause the syllable on which they are executed to be perceived as accented. I will use the term **pitch accent** or simply **accent** to refer to prominence caused by the presence of such an accent-lending pitch movement. Of course, the intonation model does not predict where in a sentence pitch accents or intonational boundaries will be located, but when these locations are provided as input, the model is capable of generating a natural-sounding contour. In the remainder of this paper I will deal specifically with pitch accent assignment.

It is relatively standard nowadays to view accent placement as a process involving two stages (cf. Ladd, 1980; Gussenhoven, 1983; Fuchs, 1984; Baart, 1987): in the first stage it is decided which constituents of a sentence contain relatively important information (e.g. because they add new information to the background shared by speaker and hearer) and are therefore to be focussed upon; the decision to focus certain parts of a sentence and not focus other parts is based on semantico-pragmatic information and in principle cannot be predicted from the lexico-syntactic structure of a sentence. In the second stage, the exact location of a pitch accent within a focussed constituent is determined; here lexico-syntactic structure does play a crucial role. The following example, cited from Ladd (1980), illustrates these ideas. (In the examples, pitch accent is indicated by means of capitalization.)

(1) even a nineteenth century professor of CLASSICS wouldn't have allowed himself to be so pedantic

In this case, it is probably the speaker's intention to focus on the subject NP; we can say that all the material from *a* to *classics* is [+focus], while the rest of the sentence is [-focus]. Given the speaker's decision to focus on the subject, an accent is placed by rule on the last lexical element within this constituent.

In the following sections, I first discuss the rules that place an accent within a focussed constituent in Dutch, and next turn to the problem of assigning focus to the constituents of a sentence.

3. From focus to accent

As will be clear from the paragraphs above, I assume that accent placement is predictable if the focussing structure of a sentence is known (for discussion see Gussenhoven et al., 1987; Baart, 1987). I adopt Gussenhoven's (1983) idea that accent placement is sensitive to the argument structure of a sentence; however, I replace his semantic orientation by a syntactic one and apply the term **argument** to any constituent which is selected by the subcategorization frame of some lexical head, including subjects.

Input to the accent rules is a binary branching syntactic constituent tree, where apart from syntactic category a node is provided with information concerning its argument status (either argument or not an argument of some lexical head), and where nodes dominating a focussed constituent are assigned the feature [+focus], while nodes dominating unfocussed material are [-focus]. In order to arrive at an accentuation pattern, three rules and a well-formedness condition are to be applied to this input. A first rule (see (2)) applies iteratively to pairs of sister nodes in the input tree, replacing the syntactic labels with the labels *s* (for 'strong') or *w* (for 'weak'), familiar from metrical phonology. By convention, whenever a node is labelled *s* its sister has to be labelled *w* and vice versa, the labellings [s s] and [w w] being excluded for pairs of sister nodes.

(2) Basic Labelling Rule (BLR):
A pair of sister nodes [A B] is labelled [s w] iff A is an argument; otherwise the labelling is [w s].

The function of the *w/s*-labelling is to indicate which element of a phrase will bear the accent when the phrase is in focus: after the application of focus assignment and *w/s*-labelling rules, an accent will be assigned to every terminal that is connected to a dominating [+focus] node by a path that consists exclusively of *s*-nodes.

In (3) I illustrate the operation of the BLR. All left-hand sisters in (3) are labelled *w*, except for the NP *een mooi boek*, which is an argument. Granted a focus on the predicate, accent will be assigned to the element *boek* (there is a path from *boek* to the [+focus] node that consists of *s*-nodes only).

(3) (ik) heb een mooi BOEK gekocht
I have a nice book bought
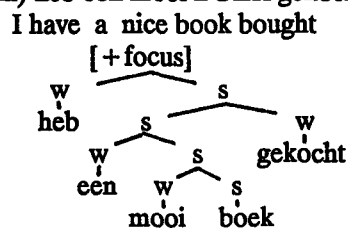
The output of the BLR may be modified by two additional rules. First, the Rhythm Rule accounts for cases of rhythmical accent shift, see (4).

(4) Rhythm Rule (RR, applies to the output of the BLR):
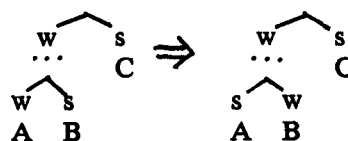
Conditions:
(a) C is dominated by a focus
(b) B and C are string-adjacent
(c) A is not a pronoun, article, preposition or conjunction

In (5), where we assume focus on both the main verb and the time adverbial, the accent pattern on the adverbial has been modified by the RR (the accent which is normally realized on *nacht* has been shifted to *hele*).

(5) (hij heeft) de HELE nacht GELEZEN
  he has the whole night read

```
        w̄‾‾‾‾‾‾s
   [+focus]      [+focus]
   w‾‾‾s          gelezen
   de  s‾‾‾w
      hele  nacht
```
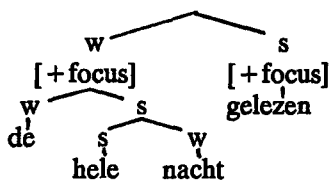
Until now, nothing prevents the label *s* from being assigned to a node which is [-focus]. The following rule, adopted from Ladd (1980) takes care of this case. The rule makes sure that a [-focus] node is labelled *w*; by convention, its sister node becomes *s*.

(6) Default Accent (DA):

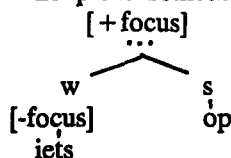$$s \longrightarrow w$$
[-focus]
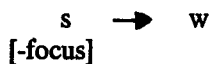
While arguments are normally labelled *s* and therefore likely to receive accent, there are some cases where we do not want an argument to be accented. A case in point are [-focus] pronouns. In (7a) we have an example of a lexical object NP (*een speld*); in (7b) this NP is replaced by a [-focus] pronoun (*iets*). As a result of the DA rule, it is the particle (*op*) that receives the accent in (7b), instead of the object.

(7a) (hij raapt) een SPELD op
  he picks a pin up
        [+focus]

```
        s‾‾‾‾‾‾w
   w‾‾‾s        op
   een  speld
```
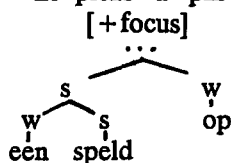
(b) (hij raapt) iets OP
  he picks something up
        [+focus]

```
      w‾‾‾‾‾‾s
  [-focus]    op
   iets
```

In addition to the rules presented thus far, a well-formedness condition is necessary in order to account for the focus-accent relation. It has been noted by Gussenhoven (1983) that an unaccented verb may not be

part of a focus domain if it is directly preceded by an accented adjunct. For instance, in (8a)

(8a) (in ZEIST) is een FABRIEK verwoest
  in Zeist is a factory destroyed

the verb (*verwoest*) is unaccented. There is no problem here: the VP as a whole is in focus, due to the accent on the argument *een fabriek*. Consider, however, (8b):

(8b) (in ZEIST) is een FABRIEK door BRAND
  verwoest
  in Zeist is a factory by fire
  destroyed

This is a somewhat strange sentence. The accent on *door BRAND* arouses an impression of contrast and the verb *verwoest* is out of focus. A more neutral way to pronounce this sentence is given in (8c):

(8c) (in ZEIST) is een FABRIEK door BRAND
  VERWOEST
  in Zeist is a factory by fire
  destroyed

The following condition is proposed in order to account for this type of data:

(9) Prosodic Mismatch Condition (PMC):

```
* [+focus]        * [+focus]
  w‾‾‾s             s‾‾‾w
 +acc  -acc       -acc  +acc
```
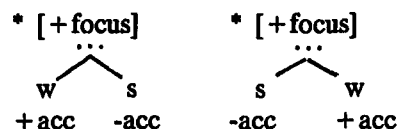
The PMC states that within a focus domain a weak (*w*) constituent (such as *door brand* in (8b,c)) may not be accented if its strong (*s*) sister (such as *verwoest* in (8b,c)) is unaccented.

4. Assigning focus

Assuming that a programme for semantic interpretation of unrestricted Dutch text will not be available within the near future, the following practical strategy is proposed for assigning focus to constituents in a syntactic tree. This strategy is based on the insight that word classes differ with respect to the amount of information that is typically conveyed by their members. The central idea is to assign

[+focus] to the maximal projections of categories that convey extra-grammatical meaning (nouns, adjectives, verbs, numerals and most of the adverbs). In addition, [-focus] is assigned to pronouns. In the case of a coordination, [+focus] is assigned to each conjunct. Finally, [+focus] is assigned to the sisters of focus-governing elements like *niet* 'not', *ook* 'also', *alleen* 'only', *zelfs* 'even', etc. Below I informally present an accent assignment algorithm which combines these focus assignment heuristics with the focus-to-accent rules discussed in section 3:

(1) Read a sentence with its surface structure representation.
(2) Assign the labels *w* and *s* to nodes in the tree, according to the BLR above.
(3) Assign [-focus] to pronouns.
(4) Apply DA: if an *s*-node is [-focus], replace *s* by *w* for this node and *w* by *s* for its sister.
(5) Apply the RR, starting out from the most deeply embedded subtrees.
(6) Assign [+focus] to S, (non-pronominal) NP, AP, AdvP and NumP nodes.
(7) Assign [+focus] to each member of a coordination.
(8) Assign [+focus] to the sister of a focus governor.
(9) Assign [+focus] to every *s*-node, the sister of which has been assigned [+focus] (thus avoiding prosodic mismatch, see the PMC above).
(10) Assign accent to each word that is connected to a dominating [+focus] node via a path that consists exclusively of *s*-nodes.
(11) Stop.

5. Perceptual evaluation

The accent assignment algorithm has been implemented as a Pascal programme. Input to this programme is a Dutch sentence; the user is asked to provide information about syntactic bracketing and labelling, and about the argument status of constituents. The programme next assigns focus structure and w/s labelling to the sentence and outputs the predicted accent pattern.

A small informative text was used for evaluation of the output of the programme. In this evaluation experiment, the predicted accent patterns were compared with the accent patterns spontaneously produced by a human reader, as well as with the accent patterns as predicted by a naive accentuation algorithm which assigns an accent to every content word. Listeners were asked to rate the quality of sentences synthesized with the respective accent patterns on a 7-point scale. As a summary of the results, I here present the mean scores for each of the conditions:

Spontaneous accentuation: 5.2
Sophisticated algorithm: 4.6
Naive algorithm: 3.3

As one can see, human accentuation is still preferred over the output of the algorithm of section 4. Of course this is what we expect as the algorithm does not have access to the semantico-pragmatic properties of an input text, such as coreference and contrast. On the other hand we see that the algorithm which does take syntactic effects on accent placement into account, offers a substantial improvement over a simple algorithm based on the content word - function word distinction.

References

Baart, Joan L.G. (1987): *Focus, Syntax and Accent Placement*. Doct. diss., Leiden University.

Fuchs, Anna (1984): 'Deaccenting' and 'default accent'. In: Dafydd Gibbon & Helmut Richter (eds.): *Intonation, Accent and Rhythm*, de Gruyter, Berlin.

Gussenhoven, Carlos (1983): Focus, mode and the nucleus. *Journal of Linguistics* 19, p. 377-417.

Gussenhoven, Carlos, Dwight Bolinger & Cornelia Keijsper (1987): *On Accent*. IULC Bloomington.

't Hart, J. & A. Cohen (1973): Intonation by rule, a perceptual quest. *Journal of Phonetics* 1, p. 309-327.

't Hart, J. & R. Collier (1975): Integrating different levels of intonation analysis. *Journal of Phonetics* 3, p. 235-255.

Kager, René & Hugo Quené (1987): Deriving prosodic sentence structure without exhaustive syntactic analysis. In: *Proceedings European Conference on Speech Technology*, Edinburgh.

Ladd, D. Robert jr. (1980): *The Structure of Intonational Meaning*. Indiana U.P., Bloomington.