TESTING THE MEANING OF FOUR DUTCH PITCH ACCENT TYPES[1]

Johanneke Caspers

Phonetics Laboratory Leiden University/HIL
Cleveringaplaats 1, 2311 BD Leiden, The Netherlands
Tel. +31 71 527 2325, FAX: +31 71 527 2615, E-mail: caspers@rullet.leidenuniv.nl

ABSTRACT

The Grammar of Dutch Intonation ('GDI', [1]), the model I adopted to describe intonational phenomena, provides an inventory of accent-lending and boundary-marking pitch configurations for Dutch, but little is known about the factors influencing the choice within these two categories. The present study aims to provide some insight into this issue, by way of experimentally testing abstract linguistic propositions regarding the meaning of a number of accent-lending intonation patterns in Dutch. Two perception experiments have been carried out to test four form-meaning hypotheses and the results confirm the basic correctness of the semantic proposals.

1. INTRODUCTION

The semantic notions that are associated with intonation are prominence and phrasing (among others, [2], [3],[4]). Speakers use melodic means to direct the listener's attention to the semantically central aspects of the message (prominence) and to aid the listener in segmenting the speech stream at several linguistic levels (phrasing). To enable the listener to locate focused or semantically central constituents, pitch accents are exploited, and the location of boundaries can be melodically marked. Speakers can choose from an inventory of different accent-lending pitch configurations when focusing a constituent, but little is known about the reasons a speaker has for selecting a specific pitch accent type, or what this choice conveys to the listener. In this contribution a hypothetical meaning analysis of four Dutch pitch accent types is presented (section 2), as well as a description of the experimental approach chosen to verify the semantic propositions (sections 3 and 4). The paper concludes with a discussion of the obtained experimental results (section 5).

2. A SEMANTIC ANALYSIS OF FOUR PITCH ACCENT TYPES

Four simple melodic shapes were selected for experimental verification, each containing one pitch accent. In figure 1 the four shapes are illustrated, superimposed on a three-syllable word with an accented second syllable (capitalized). The letters and numbers correspond to the perceptually relevant pitch movements distinguished in the GDI ('1': accent-lending rise, '2':

boundary-marking rise,[2] 'A': accent-lending fall, 'E': accent-lending half fall, 'Ø': high declination line, '0': low declination line, '&' indicates that movements are combined on one syllable).
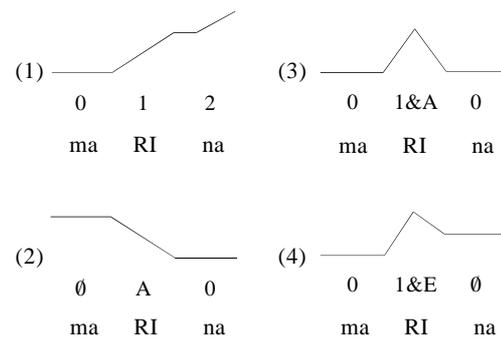

Figure 1. Stylized examples of the four accent-lending shapes on the proper name 'Marina'.

Generally, there is agreement that the function of a pitch accent is to focus the hearer's *attention* on specific parts of the utterance, and that the type of pitch accent contains additional information, expressing the communicative '*intention*' of the speaker ([5]). As Ward & Hirschberg ([6]) have shown, the 'intentional' meaning aspect cannot be defined truth-functionally, which means that true-false distinctions cannot be used as experimental technique. According to Pierrehumbert & Hirschberg ([7]) 'intention' pertains to the information status of the prominent element in the 'background' shared between speaker and listener. Following this line of thought, meaning hypotheses were developed for four Dutch pitch accent types on the basis of two theoretical analyses of Dutch melodic shapes: the structuralistic form-meaning units of Keijsper ([8]) and the autosegmental tones of Gussenhoven ([9]). Below the four melodic shapes (in autosegmental as well as GDI notation) and their meaning hypotheses are given (S = speaker, V = 'Variable', for focused element):

1) L*HH% with low onset/'12' = 'testing': S leaves it up to the hearer to determine whether V belongs to the background or not ('is V part of the background?')

2) H*LL% with high onset/'A' = 'selection': S selects a V from the background (which means that V was present in the background at the moment of speaking)

3) H*LL% with low onset/'1&A' = 'addition': S

adds a V to the background (which means that V was not present in the background at the moment of speaking)

4) half completed H*L/'1&E' = 'addition' plus 'this information was predictable' (V is added to the background, but that should come as no surprise)

A further distinction was made: each pitch accent can be used either 'literally' or 'metaphorically' (cf. [9]). This difference in 'orientation' is operationalized as 'speaking about a person' (literal orientation) versus 'addressing a person' (metaphorical orientation).

3. EXPERIMENT I

One-word utterances carrying the four different pitch accent types were embedded in a specific context and presented to listeners in an appropriateness test. In the contexts, the background shared between speaker and listener is stipulated, creating a specific setting for the manipulation of this background. For each of the four abstract meaning categories different contexts were created for the two types of 'orientation' (S speaks about a person versus S addresses a person), and there are three versions of every meaning-orientation combination, resulting in a total of 24 different contexts. All contexts are situated in a school and the speaker ('S') is the teacher. I give two examples here, translated into English:

a) You are in a meeting with colleagues about the problem of ongoing thefts from the cafeteria cash register. A number of aspects of the thefts are being discussed and suddenly everything is clear to you; it hàs to be her: "Marina"

In context a) the orientation is literal. The person referred to is not present in the shared background at the moment of speaking, which leads to the prediction that contour type '1&A', expressing 'addition', will be the best fitting melodic shape in this context.

b) On the agenda are a number of classroom presentations; after a few introductory remarks about the first subject you address the pupil concerned: "Marina"

The addressing of the pupil (metaphorical orientation) is present in the background, since everybody knows that Marina will be giving a presentation, so the prediction for context b) is that contour type 'A', 'selection', will be the most appropriate melodic shape.

As a prerequisite for verifying the intentional meaning of spoken realizations of the test utterances, the adequacy of the contexts was independently evaluated in a paper-and-pencil experiment ([10]). Results proved that the contexts present an adequate reflection of the intended meaning categories.

The four different melodic shapes were realized by two intonologists on four different proper names. The 24 contexts were presented one by one, followed by a proper name which was realized with each of the four different intonation patterns. Sixty-four subjects were asked to project themselves into the speaker role (a school teacher) in each situation, and to rate the four versions of the target utterance from best to worst fit in their particular context. The data were presented to the subjects via an interactive computer program.[3]

The four different intonation contours are assumed to correspond to the four abstract meaning categories (cf. section 2), and therefore a positive association
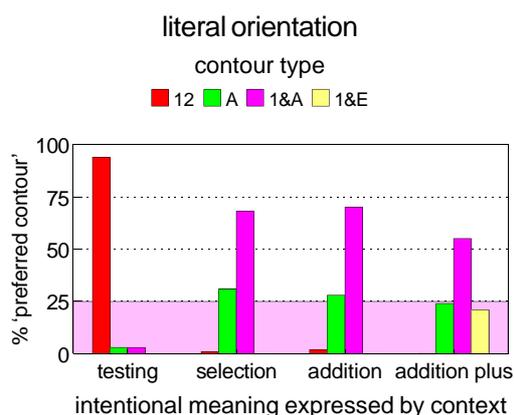

Figure 2. Percentage of preferred contour type in the literal orientation, broken down by context type.

between the contour type corresponding to the meaning captured in a specific context and the judged appropriateness of that contour type was expected. Results are presented in figures 2 and 3.

The contexts designed to capture the abstract meaning of 'testing' receive a large number of responses favoring the predicted contour type ('12') in both
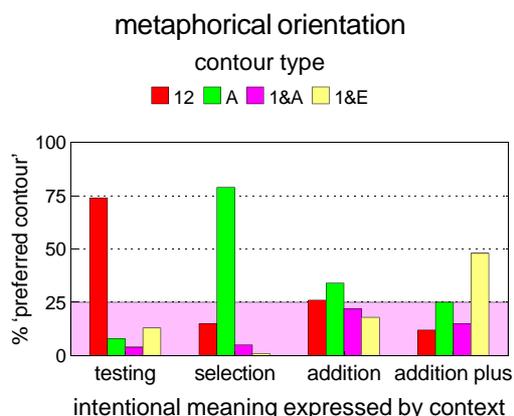

Figure 3. Percentage of preferred contour type in the metaphorical orientation, broken down by context type.

orientations (94% and 74% respectively). The other three types of context behave as predicted in just one

orientation: the 'selection' and 'addition plus' contexts perform well metaphorically (79% and 48% correct), but poorly in the literal orientation (around chance level), and the 'addition' contexts perform well literally (70% correct), whereas the responses in the metaphorical orientation are roughly random.

In the literal orientation (speaker refers to a person) all contexts, except the 'testing' ones, lead to a preference for contour '1&A' (preferred in ca. 65% of the cases). On the assumption that this accent type is the most neutral means of assigning prominence (cf. [8]), it would constitute a convenient escape hatch in cases of doubt. Possibly there is a 'pragmatic' explanation for the finding that contour types 'A' and '1&E' are not preferred where predicted, e.g., interference of what is felt to be 'good manners' or 'polite'. Indeed, contour type 'A' is chosen as second most appropriate contour type in the literally oriented 'selection' contexts (data are not presented). Subjects probably find use of 'A' a little impolite because it indicates that the listener should already have this information.

The literal 'addition plus' contexts did not evoke contour type 4 ('1&E') as the most appropriate type, which may be a result of the hypothesized aspect of *predictability* (cf. [9]). Keijsper chooses a stronger wording: the information is really *superfluous* ([8]). It may be the case that this is the better meaning analysis and that the contexts do not sufficiently express this superfluousness, causing the subjects to divert to other contours, because choosing the '1&E' contour would be impolite (since it exposes the speaker's feeling that what he says is superfluous, i.e. suggesting that the listener is stupid). This assumption is supported by the fact that the one context containing the term 'annoyed' performed far better than the other two. The same pattern is present in the metaphorical orientation, where the 'addition plus' contexts generally performed rather well: here the results are also better for the contexts expressing annoyance. This means that the abstract meaning of the '1&E' contour (and, subsequently, each relevant context) has to be revised.

In the metaphorical orientation the results are more in line with the predictions, except for the 'addition' category. Contour '1&A' is preferred in only 12% of the total number of metaphorically oriented contexts, which indicates that this type of pitch accent is not very suitable to initiate a conversation. Indeed it proved to be very difficult to design contexts where addressing a person would constitute the addition of new information to the background. Normally, there is eye contact between speaker and addressee, which renders 'addition' an illogical manipulation, since the communicative situation between speaker and hearer is part of the background already.

4. EXPERIMENT II

The goals of the second perception experiment were to verify the results obtained in the first experiment, and to test the adjusted meaning for contour '1&E'—'addition' plus 'this information is really superfluous'. The same materials were used, except for the 'addition plus' contexts, which were strengthened. For example:

c) A colleague has just asked you which pupil was absent that morning; you mentioned the name Marina, but your colleague apparently has not listened well and asks: "Who?" You reply, somewhat annoyed: "Marina"

Thirty subjects were presented with a proper name bearing one of the four pitch accent types, as well as four different situational contexts (one for each type of intentional meaning), which they had to rank from the most to the least appropriate environment for the proper name in question.[4]

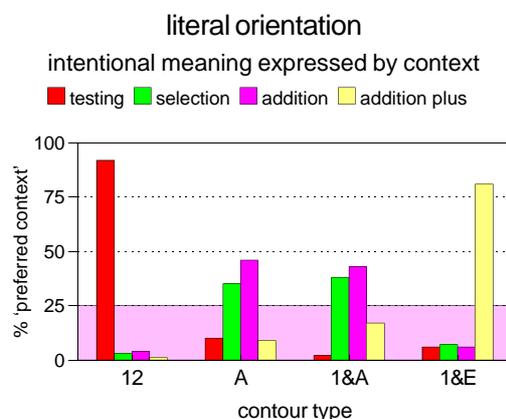It was hypothesized that the results would resemble

Figure 4. Percentage of preferred context type in the literal orientation, broken down by contour type.

those of experiment I, except for the 'addition plus' contexts; the match between the adjusted contexts and contour type '1&E' was expected to be better than in experiment I. Results are presented in figures 4 and 5.
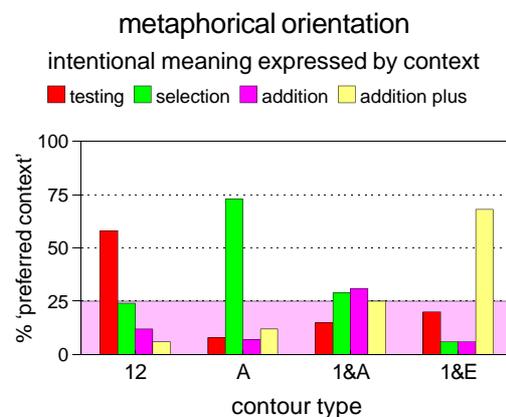

Figure 5. Percentage of preferred context type in the metaphorical orientation, broken down by contour type.

In the literal orientation the best results are again

found in the 'testing' category: when contour '12' is presented, the 'testing' context is preferred in 92% of the cases. The results for contours 'A' and '1&A' are virtually the same, with responses favoring both the 'selection' and the 'addition' contexts (respectively 35% and 43% correct). The results for contour '1&E' are as predicted: the adjustment of the abstract meaning and situational contexts results in a substantial improvement: the number of correct responses rises from 21% to 81%. The outspoken preference for the 'addition' category found in experiment I has vanished, but again there is no clear difference between the 'selection' and 'addition' categories.

In the metaphorical orientation the data are very close to those from experiment I ('12' 58%, 'A' 73%, '1&A' 31% and '1&E' 68% correct responses), except for the '12' contours, that—unexpectedly—perform worse (but still well above chance), and the '1&E' contours, that—as predicted—perform better. The responses to the '1&A' contour are randomly dispersed over the four types of context (cf. experiment I).

5. DISCUSSION AND CONCLUSION

The results of experiment II support those obtained in experiment I. The association between the '12' and '1&E' contours and the matching (adjusted) situational contexts is fairly strong, which indicates that the hypothesized abstract meanings for both contours are essentially correct. The results in the 'selection' and 'addition' categories are less clear; in the literal orientation the 'A' and '1&A' contours and the 'selection' and 'addition' contexts form one group, indicating that there is no difference between these contours when referring to a person, at least not in the way I hypothesized. This may be explained by the fact that a speaker is free to present the focused information as either 'given' or 'new', since the names of the pupils are present in the background, but the fact that the speaker refers to this pupil and not another one may be viewed as new information. When addressing a person (metaphorical orientation), however, there is a clear difference between these groups in both experiments. The 'selection' category performs well, which indicates that 'selection' is an adequate description of the intention expressed by contour 'A'. The poor performance of the 'addition' category may be explained by the fact that adding a communicative situation between speaker and hearer to the background is not a very natural 'intention'.

The results show that different pitch accent types express different 'intentions', as was predicted on the basis of non-experimental analyses of the meaning of intonation ([7],[8],[9]). Results suggest that the dimension of the orientation of a pitch accent should be carefully considered when investigating the contribution of intonation to utterance interpretation. The abstract meaning hypotheses are generally supported, which

indicates that it is possible to express intonational meaning in terms of the information status of a referent with respect to a background, and to test these abstract 'intentions' in an experimental setting.

Footnotes
[1] This research was funded by the Netherlands Organization for Scientific Research (NWO) through the Foundation for Speech, Language and Logic, under project # 300-75-001.
[2] In my opinion a single accent-lending rise ('1') is not suitable utterance-finally: it indicates that the speaker is not finished yet. Since I wanted the tested utterances to be final, I added the final rise '2'. This means that in contour (1) the meaning of '2' is added to the meaning of '1'. At present, an experiment investigating the meaning difference between contours '12' and '1' is conducted.
[3] The experiment can be accessed at:
"http://fonetiek-6.leidenuniv.nl/caspers/le-intro.html".
[4] The experiment can be accessed at:
"http://fonetiek-6.leidenuniv.nl/caspers/le2-intro.html".

6. REFERENCES

[1]   Hart, J. 't, R. Collier & A. Cohen (1990) *A perceptual study of intonation, An experimental-phonetic approach to speech melody*. Cambridge University Press.
[2]   Bolinger, D.L. (1975) *Aspects of language*. Harcourt Brace Jovanovich, New York.
[3]   Cutler, A. (1991) Prosody in situations of communication: salience and segmentation. *Proceedings of the 12th International Congress of Phonetic Sciences*, vol.1: 264–270.
[4]   Ladd, D.R. (1992) An introduction to intonational phonology. In: G.J. Docherty & D.R. Ladd (eds) *Papers in Laboratory Phonology II* [Gesture, segment, prosody], 321-334.
[5]   Grosz, B.J. & C.L. Sidner (1986) Attention, intentions, and the structure of discourse. *Computational Linguistics* 12, 175-204.
[6]   Ward, G. & J. Hirschberg (1985) Implicating uncertainty. *Language* 61, 747-776.
[7]   Pierrehumbert, J.B. & J. Hirschberg (1990) The meaning of intonational contours in the interpretation of discourse. In: P.R. Cohen, J. Morgan & M.E. Pollack (eds) *Intentions in communication*. MIT Press, Cambridge Massachusetts, 271-311.
[8]   Keijsper, C.E. (1984) Vorm en betekenis in Nederlandse toonhoogtecontouren. *Forum der Letteren* 25, 20-37, 113-126.
[9]   Gussenhoven, C. (1984) *On the grammar and semantics of sentence accents*. Foris, Dordrecht.
[10]   Caspers, J. (1996) Testing the semantics of single-accent intonation patterns. In: A. Arregui & C. Cremers (eds) *Meaning on the HIL*. HIL Occasional Papers 1, 14-32.