On the Alleged Existence of Contrastive Accents

Krahmer, E.J.; Swerts, M.G.J.

Published in:
Speech Communication

Publication date:
2001

ELSEVIER

Speech Communication 34 (2001) 391–405

SPEECH COMMUNICATION

www.elsevier.nl/locate/specom

On the alleged existence of contrastive accents ☆

Emiel Krahmer [a,*], Marc Swerts [a,b]

[a] *IPO, Center for User–System Interaction, Eindhoven University of Technology, P.O. Box 513, NL-5600 MB, Eindhoven, Netherlands*
[b] *CNTS, Universitaire Instelling Antwerpen, Universitertsplein 1, B-2610 Wilrijk, Belgium*

Received 8 June 1999; received in revised form 10 June 2000; accepted 14 June 2000

Abstract

Speakers may use pitch accents as pointers to new information, or as signals of a contrast relation between the accented item and a limited set of alternatives. There is no consensus in the literature whether a separately identifiable contrastive accent exists. Some studies report that contrastive accents are more emphatic than newness accents and have a different melodic shape. In other studies, however, it is maintained that contrastiveness can only be determined by looking at how accents are distributed in an utterance. It is argued that these two contrasting views on contrastiveness can be reconciled by showing that they apply on different levels. To this end, accent patterns were obtained in a (semi-) spontaneous way via a dialogue game (Dutch) in which two participants had to describe coloured figures in consecutive turns. By varying the sequential order, target descriptions ('blue square') were collected in four contexts: no contrast (all new), contrast in the adjective, contrast in the noun, all contrast. A distributional analysis revealed that both all new and all contrast situations correspond with double accents, whereas single accents on the adjective or the noun are used when these are contrastive. Single contrastive accents on the adjective are acoustically different from newness accents in the same syntactic position. The former have the shape of a 'nuclear' accent, whereas the newness accents on the adjective are 'prenuclear'. Contrastive accents stand out as perceptually more prominent than newness accents. This difference in salience tends to disappear if the accented word is heard in isolation.

Keywords: Contrast; Prosody; Dialogue games; Phonetics; Perception

1. Introduction

It has been a matter of considerable debate whether intonation, defined as speech melody, has an intrinsic meaning which a speaker adds to the lexically and syntactically determined meaning of an utterance. Consider the use of pitch accents. Some intonologists would claim that different types of pitch accents indeed express some 'extra' attitudinal information (Gussenhoven, 1983; Keijsper, 1984; Pierrehumbert and Hirschberg, 1990, to mention a few). Others, however, have explicitly argued against a close correspondence between intonation and meaning, either because the relation is one-to-many or even many-to-many ('t Hart et al., 1990), or because it is believed that context largely determines the interpretation of a given intonational pattern, to such an extent that "(...) *the attempt to extract from* [intonation contours] *an element of commonality valid in all contexts must be reckoned a futile endeavour*" (Cutler, 1977, p. 106). To shed some light on this

☆ Speech files available. See www.elsevier.nl/locate/specom.

* Corresponding author. Tel.: +31-40-247-5257; fax: +31-40-243-1930.

E-mail addresses: e.j.krahmer@tue.nl (E. Krahmer), m.g.j.swerts@tue.nl (M. Swerts).

0167-6393/01/$ - see front matter
PII: S0167-6393(00)00058-3

complex matter, the current paper will focus on one issue: the question whether pitch accents can be used to mark a specific aspect of meaning, namely that of *contrast*.

Bolinger (1986) takes contrast to refer to "(...) *cases where one or more individual items are singled out from a larger* (*but limited*) *set as being true regards some relationship whereas others in the same set are untrue...*" Similar characterizations are given by Cruttenden (1986) and Chafe (1974, 1976). A typical example, taken from Chafe (1974), is (1), with small caps indicating an accent:

(1) RONALD made the hamburgers.

According to Chafe (1976, p. 28), what is conveyed by (1) is "(...) *the speaker's knowledge that Ronald, as opposed to other possible candidates the addressee might have in mind, is the right selection for this role*". In that sense, Ronald 'contrasts' with a limited number of alternatives in the listener's consciousness. Similar ideas are formalized in Rooth (1985)'s *alternative semantics* and in Prevost (1995)'s theory of contrast, which is based on Rooth's work.

A major issue is the size and the determination of the set of alternatives. According to some (e.g., Bolinger, 1961; Schmerling, 1976; Prevost, 1995) this problem of restricting the set of alternatives is a main stumbling block in formalizing the notion of contrast (but see van Deemter, 1999). To illustrate, consider the following example, from Blok (1993).

(2) John only SWIMS.

Blok notes that the alternative set, as defined by Rooth (1985), would contain *all* properties that John has. On the basis of similar examples, Schmerling (1976) concludes that contrast is merely a matter of degree, while Bolinger (1961), reasoning along similar lines, concludes that the size of the set of alternatives is inversely proportional to the likelihood of a contrastive interpretation for an accent. However, looking back to example (2), it should be noted that someone uttering (2) is not contrasting swimming with all properties of John, but rather with a restricted set, for instance of sporting activities. In general, and we shall return to this issue later on (in Section 4.2), various ways have been proposed to restrict the set of all alternatives by taking the linguistic *context* into account.

There is a general consensus in the literature that in languages such as Dutch and English, contrastive meaning can be signalled by means of accent distribution, but only to some extent. To illustrate this, consider the following pair of utterances:

(3) Push the RED button.

(4) Pick up the blue CYLINDER.

These sentences differ in that the former can only be used in one type of context whereas the latter is suitable in at least two contexts. What is conveyed by (3) is a request from the speaker to push the red button as opposed to some differently coloured candidate button which the addressee might have in mind. The accent on 'red' is said to have a *narrow focus* because its scope is limited to the word it is associated with. Example (4), on the other hand, may continue something like "*First pick up the blue cube, and subsequently...*" (in which case the accent on 'cylinder' has narrow focus), but it may also be an answer to a question such as "*What should I do now?*". In the latter case the accent on 'cylinder' is said to have *broad focus*, i.e., takes scope over the entire utterance. Ladd (1980, pp. 78–79) claims that often an accent with a narrow focus is assigned a contrastive interpretation, and that to some extent the detection of narrow focus is determined by the *distribution* of accents. If the accent occurs in a non-default position, as in (3), a contrastive interpretation is certified. An accent in default position, of which (4) is an example, is ambiguous without further context between a narrow and a broad focus reading.

The distribution facts (narrow versus broad focus) described in connection with examples (3) and (4) are mostly uncontroversial. However, the existence of *additional* phonological features which distinguish contrastive accents from more 'neutral' accents only marking new information is hotly debated. One reason to expect that such additional features exist is that they could help disambiguate

between broad and narrow focus readings such as in (4). Indeed, some maintain that contrastive accents *are* formally different from other accents, either because the *type of accent* is different for the contrastive cases or because they are more *prominent*. Couper-Kuhlen (1984) and Chafe (1974) mention the existence of a sudden drop in pitch after the contrastive accent, whereas a non-contrastive accent is more likely to be sustained. Pierrehumbert and Hirschberg (1990) suggested that contrastive accents have an $L + H^*$ pattern while novelty accents have an H^* form. Bartels and Kingston (1994) were unable to find support for Pierrehumbert and Hirschberg's suggestion, but found evidence instead that contrastive accents tend to have higher peak heights than novelty accents. The latter finding is in line with the claim that contrastive accents are more 'emphatic', in the sense that they are extra high or boosted (Brown et al., 1980; Ladd, 1983). Others, however, maintain that contrastive accents do *not* exhibit specific intonation features. This is the position taken by intonologists like Halliday and Bolinger. "*As far as we can tell from the behaviour of pitch, nothing is uniquely contrastive*" (Bolinger, 1986, p. 342).

In this paper, it is argued that these two opposing views can be reconciled by showing that they apply on different levels. Given that it is difficult to judge and compare the various observations mentioned above (differences in the definition of contrastiveness, methodological differences, etc.), we propose an experimental approach in an attempt to create a common test bed for the various hypotheses. The study consists of a production test (Section 2) and a perception test (Section 3). The former mainly concentrates on accent distributions as a function of different discourse contexts and also tackles the issue of differences in shape between contrastive and neutral accents. The latter investigates whether we can find perceptual evidence for the alleged difference in prominence between 'normal' and 'contrastive' accents. As will be shown, both the production and perception data reveal that the context of accents is a crucial factor to determine contrastive intonation. We end this paper with a general discussion (Section 4) and some remarks about contrastiveness in other languages (Section 5).

2. Production test

2.1. Method

The production experiment, which tries to mimic natural conversation, consists of a simple dialogue game played by four pairs of subjects. The game is essentially an alignment task of figures played by two subjects, call them A and B, who are separated from each other by a screen. Fig. 1 gives a bird's-eye view on the experimental set-up of the elicitation task, both the initial stage and the stage after A's first move. In each game, both players had an identical set of eight cards to their disposal, every card showing a geometrical figure (square, triangle or circle) in a particular colour (blue, red, black or yellow). Beforehand, the eight cards were divided in two sets of four cards each: one set of four cards was ordered as a stack, the other half

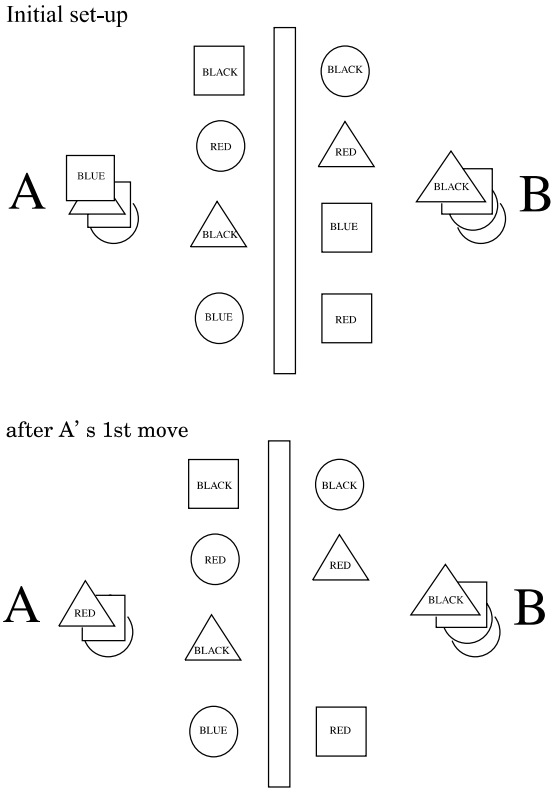

Fig. 1. Bird's-eye view of the experimental set-up. The initial stage and the first move of the dialogue game are displayed.

was unordered. The stacks of A and B were disjoint: thus A's stack corresponded with the unordered set of B, and vice versa. At the onset of the game, one participant, say A, describes the figure on the top of his stack ('a blue square') and this prompts both A and B to place the card with the blue square at the first position of the row of figures under construction. Next, B takes over and describes the object on top of *his* stack ('a black triangle'). Now, both A and B place the card with this object on the second place in the row with figures, and so on. The game is over when both players are out of cards. There are no winners or losers.

The data thus obtained allow an unambiguous operationalization of the relevant contexts. A property (colour or figure) is defined to be *new* (N) to the conversation if it is mentioned in the first turn of the current dialogue game, it is *given* (G) if it was mentioned in the previous turn and finally a property is *contrastive* (C) if the object described in the previous turn had a different value for the relevant property. To avoid a possible confusion, notice that contrast defined in this way is something different from a *correction* (which one might interpret as a kind of meta-linguistic contrast). While corrections are also highly interesting from a prosodic/phonetic (not to mention a semantic) point of view, here we restrict ourselves to the narrower definition of (non-corrective) contrast.

By systematically varying the sequential order of the cards in front of the subjects, target descriptions were collected for the eight speakers in four contexts: no contrast (all new, NN), contrast in the adjective (CG), contrast in the noun (GC), all contrast (CC). Table 1 summarizes the situation.

All [1] utterances of two target descriptions ('blue square' and 'red square') were used for a distributional analysis. To this end, two intonation experts, the second author and one independent intonologist, separately determined on which words in the different utterances they perceived an accent.

[1] There is one piece of missing data. One speaker made a slip and failed to describe a figure in CG context correctly.

Table 1
Examples of the four contexts for the utterance "blue square"

NN	(beginning of game) B: "blue square"
CC	A: "red circle" B: "blue square"
CG	A: "yellow square" B: "blue square"
GC	A: "blue triangle" B: "blue square"

2.2. Results

The results of the distributional analysis are given in Table 2. The first thing to notice is that the two intonation experts agree on the vast majority of cases. On only 4 of the 63 utterances the two disagree.

Table 2 reveals a clear trend: in the NN (no contrast/all new case) both adjective and noun are (nearly) always accented, and in most cases the same holds for the CC (double contrast) cases. When one item is given, while the other is contrasted (i.e., the CG and GC cases), the contrasted item generally is the only accented word. That the NN case always requires a double accent is interesting, since this entails that there is no ambiguity in the data between broad and narrow focus, contrary to what one might expect on the basis of the literature discussed in the introduction. Even though both CG and GC, and NN and CC are strikingly similar, there are two exceptions. First, there is a complete lack of postnuclear accents in the CG case, while occasionally prenuclear accents on the adjective occur in the GC case. Second, CC differs from NN in that there are a number of utterances in the CC context with an accent only on the adjective or the noun. Looking at these exceptional cases revealed that in all cases the speaker made a contrast with his or her *own* last utterance, thereby ignoring their partners' last contribution. Interestingly, all these 'egocentric' speakers happen to end their utterances on a high (H%) boundary tone, whereas the other speakers uniformly employed low (L%) boundary tones. This appears to be in accordance with the general observation that low boundary tones are generally

Table 2
Accent distribution according to two intonation experts (exp_1 and exp_2) on all target utterances "blauw vierkant" and "rood vierkant" (blue and red square) in four contexts: NN (no contrast), CC (all contrast), CG (contrast only in adjective), GC (contrast only in noun). One CG utterance is missing

	Accent on					
	Adj only		Noun only		Adjective and noun	
Context	exp_1	exp_2	exp_1	exp_2	exp_1	exp_2
NN	0	0	0	2	16	14
CC	3	3	2	4	11	9
CG	15	15	0	0	0	0
GC	1	1	11	11	4	4

interpreted as signals of the speaker's intention to give the turn to the other participant. Apparently, in the current experiment, speakers using high boundary tones signal that they want to continue their own train of thought, which leads to what is generally referred to as list intonation.

2.3. Discussion

The results presented above reveal that a contrastive interpretation can be derived from the way the accents are distributed in the elicited utterances. This is true for the two 'types' of speakers: those who uniformly employ low boundary tones and those who systematically use high boundary tones. The distribution results lead one to reflect on the issue whether or not contrastive accents have a specific intonational shape. If one makes the common assumption that a single accent on the noun is ambiguous between a broad focus and a narrow focus reading, then one might expect that a contrastive accent manifests itself most clearly in the noun position. However, for none of our speakers does a comparison of a single contrastive accent on the noun (GC) with a newness accent reveal differences with respect to the type of accent. This appears to be in accordance with the observation, made in Section 2.2, that the data do not exhibit any broad versus narrow focus ambiguities. Interestingly, at first sight the single contrastive accent on the adjective (CG) *is* of a different type than the newness accent on that same syntactic position. However, the single contrastive accent on the adjective is of the *same* type as the accent on the noun. Thus, the difference in type of accent is only apparent, since in the CG context the adjective is associated with a nuclear accent in a non-default position.

Fig. 2, which gives the intonation contours of one of the low-ending speakers (JR), visualizes this phenomenon. It contains JR's realizations of "blauw vierkant" (blue square) in the NN and the CG context. [2] Focussing on the F_0 pattern on the word "blauw" (blue) shows a difference in pitch contour, in particular with respect to the timing of the fall. In the NN context (top), the pitch level is sustained throughout the adjective and only drops slightly at the syllable boundary. In the CG context (bottom), one can indeed observe 'a sudden drop of pitch' *before* the syllable boundary is reached. The latter pattern is phonologically the same as the pitch movement that occurs on the syllable "vier" in the NN context. For reasons of completeness it should be noted that for the high ending speakers, essentially the same is observed: the accent on the noun in the NN context and on the adjective in the CG context both show a pronounced increase in pitch which is sustained (due to the final high boundary tone) to the end of the utterance.

The section above suggests that prosodic context largely determines a contrastive interpretation, while no evidence was found that contrastive accents have a specific intonational shape different from other accents. However, as mentioned in the

[2] The reader is invited to listen to the data at www.elsevier.nl/locate/specom.

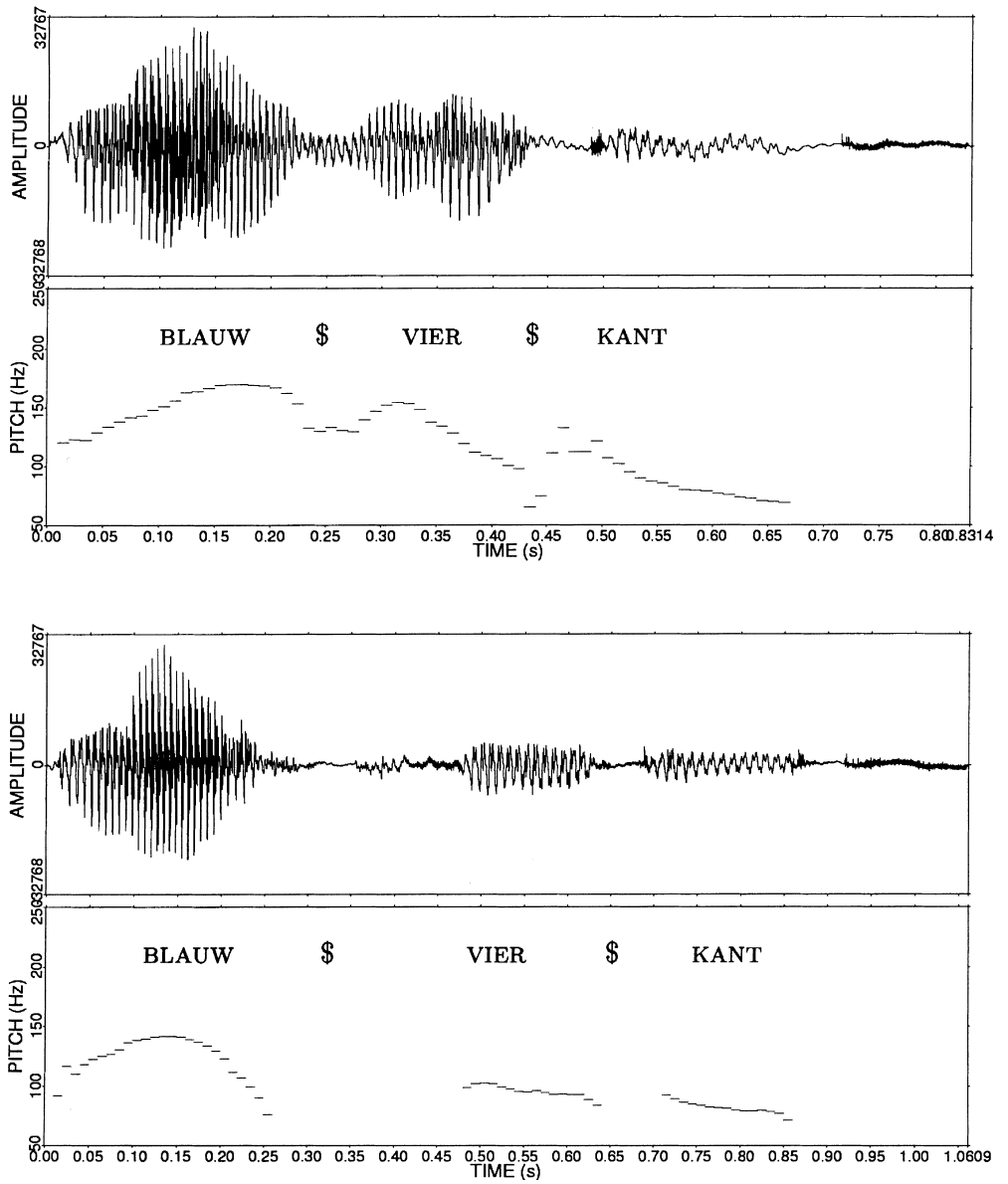

Fig. 2. Waveform and F_0 measurement of two realizations of "blauw vierkant" (blue square) by speaker JR. Top: NN context. Bottom: CG context. $ represents syllable boundaries.

introduction, it has also been claimed that contrastive accents differ from neutral ones in that they are more emphatic. We have set up a perception experiment to test whether contrastive and other accents yield different prominence ratings from listeners.

3. Perception test

3.1. Method

Perceptual data were obtained from sixteen prosodically naive subjects (distinct from the eight

speakers) who participated in an individually performed listening task. They were presented with four lists of 12 pairs of phrases: two lists consisted of utterance pairs produced by speaker JR (a low-ending speaker), the two other lists of pairs by speaker WY (a high-ending speaker). The pairs from both JR and WY data were presented in two conditions: *complete* (entire utterances) and *isolated* (words). In the former, subjects could hear the complete utterances as they were originally produced by speaker JR or WY. In the latter, listeners were presented with one word cut from its context, either the adjective or the noun. The listener's task was always to select the member of the pair which he or she thought was the most prominent: in the complete condition, they were asked to focus on either the noun or the adjective and to determine by forced choice which of the pairs contained the most prominent one. In the isolated condition, they had to select (again by forced choice) the word which they judged to be the most prominent. No specific definition of prominence was given to the subjects. Given that the utterances were elicited in four contexts (NN, CC, GC, CG), subjects had to make $3! = 6$ pairwise comparisons (NN–CC, NN–GC, NN–CG, CC–GC, CC–CG and GC–CG). These 6 pairs were always presented in both orders (A–B and B–A), which leads to 12 pairwise comparisons per list. The listener heard each of these 12 pairs two times in a row, before he or she had to make a judgment on the pair. The interval between two such identical pairs was 2500 ms, the interval between two members of a pair was 500 ms. The next pair of utterances was presented as soon as a subject had made a judgment on the previous pair. The actual listening task was preceded by the presentation of a few test utterances to make the listeners acquainted with the stimulus materials. Both the order of the pairs in a list as the order of the different lists were fully randomized.

3.2. Results

To give an overview idea of the perception results, Figs. 3 and 4 visualize the main results for

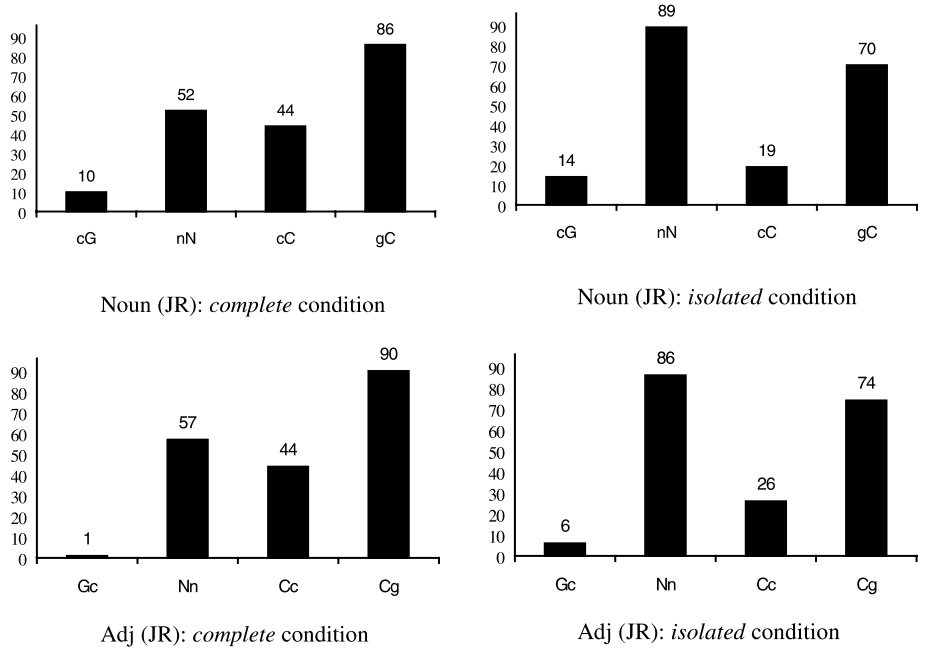

Fig. 3. Results of the perceptual measurements of prominence ratings on Adj and Noun for speaker JR in the isolated and complete conditions. The four contexts CG, NN, CC and GC are defined in the text, for readability the word of interest is capitalized. The maximum value of the *Y*-axis is 96.

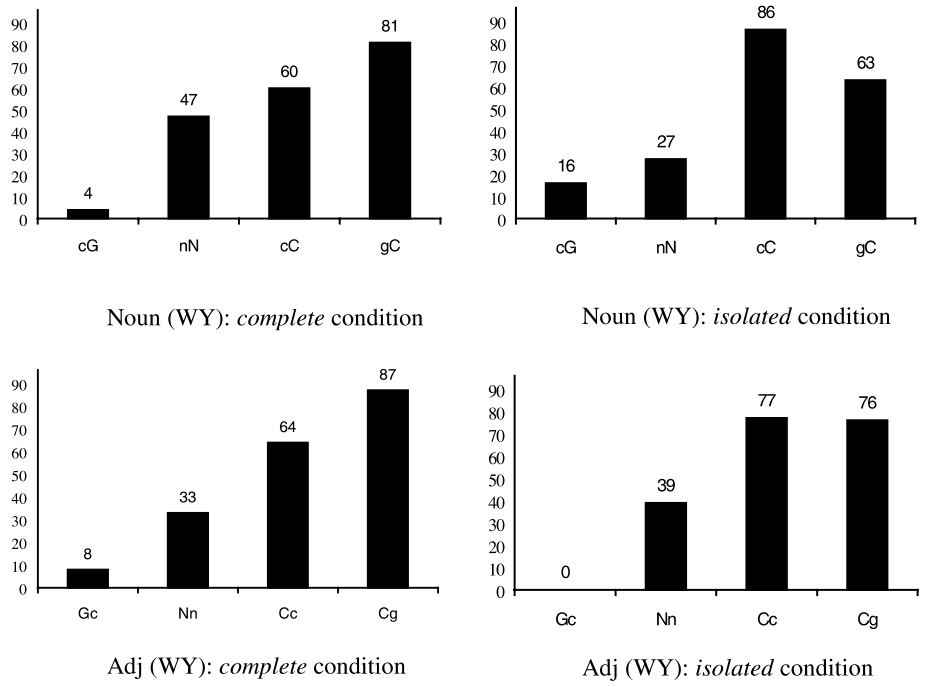

Fig. 4. Results of the perceptual measurements of prominence ratings on Adj and Noun for speaker WY in the isolated and complete conditions. The four contexts CG, NN, CC and GC are defined in the text, for readability the word of interest is capitalized. The maximum value of the *Y*-axis is 96.

speaker JR and WY, respectively. The bar charts give the number of times a word (either the adjective or the noun) was judged the most prominent in a pairwise experiment. The maximum here is 96 (16 subjects $\times$ 2 judgments $\times$ 3 comparisons). Note that these figures only give the overall results, and that the details of the pairwise comparisons are lost here (but see below).

In the complete condition, subjects were presented with pairs of entire utterances and were asked to judge in which of the two the noun or the adjective was most prominent. In the following, capitalized letters indicate the words on which the subject scored (thus: 'Gc' indicated that the subject had to rate the prominence of the adjective in GC context). The results thus obtained for the complete condition are basically the same in all cases. That is: in all four cases, a single contrastive accent (gC or Cg) is judged to be the most prominent, while givenness (Gc or cG) uniformally scores lowest on prominence. The double contrast (cC/Cc) and no contrast/all new (nN/Nn) cases are in between these two extremes, with for WY double contrast being judged relatively more prominent than the no contrast case, whereas the trend is in the opposite direction for JR. Thus both adjective and noun lead to comparable prominence patterns. This is true for both speakers, even though they provided their utterances with different intonation contours (low versus high ending). A χ^2 test revealed that all the different distributions for the data obtained in the complete condition are significantly different from chance (Noun (JR): $\chi^2 = 60.83$, df $= 3$, $p < 0.001$; Adj (JR): $\chi^2 = 84.79$, df $= 3$, $p < 0.001$; Noun (WY): $\chi^2 = 66.04$, df $= 3$, $p < 0.001$; Adj (WY): $\chi^2 = 75.04$, df $= 3$, $p < 0.001$).

However, the overall picture changes dramatically when words are presented in isolation. Within speakers essentially the same pattern for the adjective and the noun can be observed, but this pattern is rather different from the complete

pattern. In particular, for JR the newness accent (Nn/nN) is suddenly judged to be the most prominent, while the double contrast accent (Cc/cC) scores almost as low as the given case (Gc/cG). For WY, the single contrast (Cg/gC) gets comparatively lower prominence ratings. The distributions in the isolated condition are again significantly different from chance level (Noun (JR): $\chi^2 = 86.7,\ \mathrm{df} = 3,\ p < 0.001$; Adj (JR): $\chi^2 = 90.99,\ \mathrm{df} = 3,\ p < 0.001$; Noun (WY): $\chi^2 = 65.28,\ \mathrm{df} = 3,\ p < 0.001$; Adj (WY): $\chi^2 = 83.34, \mathrm{df} = 3,\ p < 0.001$). Also, for each speaker, the distributions for the isolated condition are significantly different from those obtained in the complete condition (Noun (JR): $\chi^2 = 21.92,\ \mathrm{df} = 3,\ p < 0.001$; Adj (JR): $\chi^2 = 15.62,\ \mathrm{df} = 3,\ p < 0.005$; Noun (WY): $\chi^2 = 19.47,\ \mathrm{df} = 3,\ p < 0.001$; Adj (WY): $\chi^2 = 10.45,\ \mathrm{df} = 3,\ p < 0.025$).

3.3. Discussion

In the complete condition, single contrastive accents stand out as the most prominent ones, irrespective of the intonation contour (high versus low ending) and irrespective of the place of the accent in the utterance (adjective versus noun). Similarly, given items are always judged to be the least prominent, while the all new and double contrast cases lie in between the two extremes. It is striking that in the isolated condition a different picture emerges, in that, for instance, the single contrastive accents are no longer perceived as being the most prominent ones. The only difference between the isolated and the complete condition is that in the former condition words are taken from their natural context. This suggests that prosodic context information plays a role in the complete condition (see also Section 4.1), whereas in the isolated condition hearers solely base prominence judgments on acoustic properties of the target word. In particular, informal observation suggests that in the isolated condition listeners might have focused on two prosodic features, i.e., pitch and loudness, to judge prominence. In order to find support for these impressions, the data were subjected to two types of acoustic measurements, i.e. F_0 maxima and sones. The former are commonly assumed to be relatively good correlates of pitch prominence, the latter are used in Moore et al. (1997), to account for perceived loudness, following psychoacoustical studies of loudness such as Stevens (1957) and Zwicker and Scharf (1965). Some were calculated in stationary sections of 10 ms duration, taken from the middle of the target vowels.

Table 3 contains the relevant F_0 values for the 16 target words in Figs. 3 and 4, while Table 4 contains the relevant sone values. Tables 3 and 4 show that the sone and F_0 values

Table 3
F_0 maxima (Hz) on adjective and noun for speaker JR and WY in the four contexts CG, NN, CC and GC[a]

	Accent on noun				Accent on adjective			
	cG	nN	cC	gC	Gc	Nn	Cc	Cg
JR	103	155	128	140	112	170	141	142
WY	171	169	168	166	123	131	133	141

[a] For readability the word of interest is capitalized.

Table 4
Loudness (sones) on adjective and noun for speaker JR and WY in the four contexts CG, NN, CC and GC[a]

	Accent on noun				Accent on adjective			
	cG	nN	cC	gC	Gc	Nn	Cc	Cg
JR	5.3576	7.3193	6.7168	6.9699	7.7499	12.6720	8.8869	13.8095
WY	6.9867	7.4226	9.3835	8.8698	9.3517	13.7298	15.3775	16.5147

[a] For readability the word of interest is capitalized.

are different for the different words. Tables 5–8 list the pairwise judgments for the isolated condition, combined with the respective F_0 and sone differences.

Inspection of these tables indeed reveals a correspondence between perceived prominence and pitch c.q. loudness differences. In general, if one of the utterances is substantially higher or louder than the other item in the pair, it is judged as prominent more often. As the differences get smaller, the perception of prominence tends to disappear correspondingly. In order to make these informal observations more precise, we have experimented with a number of formal models. One complication is that there is no complete picture in the literature of perceptual threshold values for different prosodic features in natural speech data. The thresholds mentioned in psychoacoustic studies tend to be 'just noticeable differences' for completely controlled stimuli, and these are only useful for present purposes in that the thresholds should be well above the psychoacoustic thresholds (see also 't Hart et al., 1990, 35ff). Suppose, for the sake of argument, that we take a conservative 10% difference as a minimal distance for our perceptual judgments, then the following model can be proposed. (Here U_1 and U_2 range over the utterances in the pairwise comparison, $\min(A, B)$ gives the minimum of A and B, $\mathsf{sone}(U)$ gives the sone value of U, and $F_0(U)$ gives the F_0 value of U.)

Table 5
Pairwise comparisons for the Noun (speaker JR) in the isolated condition[a]

A	B	F_0 diff A−B	sone diff A−B	Predicted as prominent	Perceived as prominent
nN	gC	15	0.3494	nN	nN (25*)
nN	cC	27	0.6025	nN	nN (32*)
nN	cG	52	1.9617	nN	nN (32*)
gC	cC	12	0.2531	no pref [b]	gC (32*)
gC	cG	37	1.6123	gC	gC (31*)
cC	cG	25	−0.2531	cC	no pref

[a] Per comparison, the differences in F_0 (pitch) and sone (loudness) are given, followed by the item of the pairwise comparison which was predicted to be the most prominent by the 10% model described in the main text and the item which was perceived as the most prominent. The maximum number of times an utterance could be judged the most prominent is 32 (16 listeners × 2 judgments).
[b] The F_0 difference is only 0.8 Hz below the 10% boundary.
* $p < 0.001$.

Table 6
Pairwise comparisons for the adjective (speaker JR) in the isolated condition[a]

A	B	F_0 diff A−B	sone diff A−B	Predicted as prominent	Perceived as prominent
Nn	Gc	58	4.9221	Nn	Nn (32*)
Nn	Cc	29	3.7851	Nn	Nn (32*)
Nn	Cg	28	−1.1375	Nn	Nn (32*)
Gc	Cc	−29	−1.1370	Cc	Cc (26*)
Gc	Cg	−30	−6.0596	Cg	Cg (31*)
Cc	Cg	−1	−5.0226	Cg	Cg (32*)

[a] Per comparison, the differences in F_0 (pitch) and sone (loudness) are given, followed by the item of the pairwise comparison which was predicted to be the most prominent by the 10% model described in the main text and the item which was perceived as the most prominent. The maximum number of times an utterance could be judged the most prominent is 32 (16 listeners × 2 judgments).
* $p < 0.001$.

Table 7
Pairwise comparisons for the Noun (speaker WY) in the isolated condition[a]

A	B	F_0 diff A−B	sone diff A−B	Predicted as prominent	Perceived as prominent
nN	gC	3	−1.4472	gC	gC (25*)
nN	cC	1	−1.9609	cC	cC (32*)
nN	cG	−2	−0.4359	no pref	no pref
gC	cC	−2	−0.5137	no pref	cC (23**)
gC	cG	−5	1.8831	gC	gC (29*)
cC	cG	−3	2.3968	cC	cC (31*)

[a] Per comparison, the differences in F_0 (pitch) and sone (loudness) are given, followed by the item of the pairwise comparison which was predicted to be the most prominent by the 10 % model described in the main text and the item which was perceived as the most prominent. The maximum number of times an utterance could be judged the most prominent is 32 (16 listeners × 2 judgments).
* $p < 0.001$.
** $p < 0.05$.

Table 8
Pairwise comparisons for the Adj (speaker WY) in the isolated condition[a]

A	B	F_0 diff A−B	sone diff A−B	Predicted as prominent	Perceived as prominent
Nn	Gc	8	4.3781	Nn	Nn (32*)
Nn	Cc	−2	−1.6477	Cc	Cc (30*)
Nn	Cg	−10	−2.7849	Cg	Cg (27*)
Gc	Cc	−10	−6.0258	Cc	Cc (32*)
Gc	Cg	−18	−7.1630	Cg	Cg (32*)
Cc	Cg	−8	−1.1372	no pref	no pref

[a] Per comparison, the differences in F_0 (pitch) and sone (loudness) are given, followed by the item of the pairwise comparison which was predicted to be the most prominent by the 10% model described in the main text and the item which was perceived as the most prominent. The maximum number of times an utterance could be judged the most prominent is 32 (16 listeners × 2 judgments).
* $p < 0.001$.

THE 10% MODEL
if $F_0(U_1) - F_0(U_2) > \frac{1}{10} \cdot \mathsf{min}(F_0(U_1), F_0(U_2))$
then predict U_1 is perceived as more prominent
elseif $\mathsf{sone}(U_1) - \mathsf{sone}(U_2) > \frac{1}{10} \cdot \mathsf{min}(\mathsf{sone}(U_1), \mathsf{sone}(U_2))$
  then predict U_1 is perceived as more prominent
  else predict no preference
  endif
endif

This model first evaluates the pitch difference to predict a prominence preference on the basis of a 10% difference. If the pitch difference is below this threshold, it tests if there is a 10% difference in sone value. If this second test reveals a sone difference below the 10% threshold it predicts that neither of the items in the pair is perceived as more prominent. The results with this model are included in Tables 5–8 and give correct predictions for 21 out of the 24 cases. It is worth stressing that a more ad hoc model can be formulated which gives better predictions: this model is of the same form as the one given above, except that it takes a 9% difference as threshold for pitch and a 0.5 difference in sones as a threshold for loudness. This model makes the correct prediction in 23 out of the 24 cases (96%). It appears that for speaker JR, pitch is the main explanatory factor, whereas for WY, loudness is more important. Given that the second speaker consistently ends his utterance with a high boundary tone, it appears that there is a ceiling effect for this speaker where pitch is concerned. Note that the model does not work, as expected, for the complete condition, where context appears to overrule the prominence judgments

based on acoustic features of the individual words. To give one example (for speaker JR): in a pairwise comparison (complete condition) between nN and gC, the latter is judged significantly more often to be the most prominent one, even though the accent on the noun in the NN condition is both higher and louder. Summarizing: In the isolated condition, the prominence judgment appears to be the composite result of two independent factors: pitch and loudness. That is, if listeners hear a clear difference in pitch between two members of an utterance pair, they tend to choose the utterance with the highest pitch to be the most prominent one. If there is no clear pitch difference, they focus on loudness as a cue.

4. General discussion

The discussion addresses two issues: (1) the alleged existence of contrastive accents (Section 4.1) and (2) the meaning of contrastive intonation (Section 4.2).

4.1. Contrastive intonation

In the introduction two contrasting views on contrastive accent were mentioned: the view that contrastiveness is solely determined by accent distribution and the claim that, besides distribution, the accent on the contrasted item is phonologically different from other accents in terms of accent type and prominence. The data presented here in fact give justification for both positions.

Considering the question whether contrastive accents have a distinctive *shape*, it appears that single contrastive accents on the adjective are indeed prosodically different from newness accents in the same syntactic position. That is, the former are characterized by "a sudden drop in pitch", in line with earlier observations by e.g. Chafe (1974), whereas the latter are more sustained. However, that particular phonological shape appears to be characteristic of 'nuclear' accents in general. This can be illustrated by looking at the shape of the newness and the contrastive accents on the noun; they both occur in default, nuclear position, and are essentially identical. Therefore, the apparent difference between a "new" and a "contrastive" accent on the adjective is actually a difference between a prenuclear and a nuclear accent; because of deaccentuation of the noun, the "contrastive" accent on the adjective has become the final accent in a phrase, and thus gets a nuclear shape. In other words: apparently, the contrastive interpretation is not associated with a specific prosodic shape but rather with the non-default position of the nuclear accent.

Concerning the issue of *prominence* differences, it appears that contrastive accents are perceived as more prominent than newness accents on the same syntactic position. However, this only holds true if subjects can listen to the *complete* utterance. The difference in perceived prominence tends to disappear if the noun or adjective is presented in isolation. This effect might be called *prosodic masking*: an isolated pitch peak is perceived as more prominent than the same peak presented in the context of an intonationally comparable pitch peak. (The Mt. Everest would be perceived as higher when encountered in the low lands than in the Himalaya.) For the prominence judgments, it seems that the prosodic context (whether or not the relevant accent is preceded or followed by another accent) is the major factor contributing to the perception of a contrastive intonation.

Summarizing the intonational results: depending on the perspective from which they are analysed, contrastive accents in Dutch can either be claimed to have special intonational features, both in terms of shape and prominence, or they can be argued not to be essentially different from other accents. On the one hand, one might indeed claim that (1) single contrastive accents have a different shape from other "neutral" accents that mark new information (in particular when focussing on the adjective) and that (2) contrastive accents are judged to be more prominent than newness ones. But, on the other hand, a closer look at the data reveals that (1) the difference in accent type is not so much associated with a contrast-specific prosodic shape, but rather with the non-default position of the nuclear accent. Similarly, (2) the perceived prominence is not so much the result of inherent melodic characteristics of contrastive accents but seems due to the fact that the prosodic

context does not contain other intonationally comparable pitch peaks. Therefore, based on the analyses presented here, our claim is that, in Dutch, a contrastive accent as a separate category does not exist, but contrastive intonation (at a more global level) does.

4.2. Contrastive meaning

It has been argued (e.g., by Cutler, 1977, see also Section 1) that context largely determines the interpretation of a given intonational pattern. Interestingly, recent years have seen an increasing awareness of the influence of context on meaning, and this has prompted various linguists to propose formal, context-dependent interpretations associated with contrastive accents (e.g., Rooth, 1992; Hendriks and Dekker, 1995; van Deemter, 1999; Piwek, 1998). In this paper, a different, more fundamental question was addressed, namely whether a separately identifiable contrastive intonation exists in the first place. Even though no acoustic evidence for a separate contrastive accent was found, the data show that contrastiveness *can* be determined on the basis of intonation. This means that we are now in a position where it makes sense to address the question what the meaning of contrastive intonation is. On a purely pre-theoretical level, it seems reasonable to assume that speakers signal a contrast relation to enhance the hearer's processing of their utterance: by marking the information which is contrastive, e.g., "BLUE square", they seem to say: pick the 'gestalt' of the previous square which we discussed and modify the colour value by setting it to blue (compare Levelt, 1989, pp. 131–132; Pechmann, 1984).

Arguably, the presuppositions as anaphors theory of van der Sandt (1992) is well suited to model this intuitive idea. This is not the place to dig deep in van der Sandt's theory, but in a nutshell its central claim is that presuppositions behave just like anaphors looking for an antecedent. The idea that contrastiveness triggers anaphoric presuppositions can also be found in Piwek (1998), where a synthesis is offered of contrastive and newness accents in that both are assumed to trigger alternative assertions (cf. Rooth, 1992). While Piwek's analysis of contrastive accents is certainly interesting, we disagree with the claim that both kinds of accents trigger presuppositions about alternatives. In our opinion this only holds for contrastive accents. This is confirmed by results from a pilot test which indicated that subjects are able to reconstruct the preceding context only in the case of contrastiveness (Swerts and Krahmer, in progress). We propose to associate a phrase like "a BLUE square" with a presupposition stating that there exists a differently coloured square. According to the theory of van der Sandt (1992) this presupposition needs to be resolved with respect to the preceding context (that is, the context should contain an 'antecedent' for the presupposition). This implies that the direct linguistic context, i.e., the previous turns, should contain a mention of such a square. More precisely, if the utterance ends on a low boundary tone (L%) the antecedent is to be found in the previous turn, if it ends on a high boundary tone (H%) it is to be found in the turn before the previous one (i.e., the previous turn of the current speaker). If this is not the case, the prediction is that the utterance of "a blue square" was infelicitous. In this way, the notion of a limited set of alternatives discussed in the introduction can be formally interpreted as the set of alternative objects mentioned in the last turns. An account along the lines sketched here would also model the finding that neither the distributional nor the perceptional analysis revealed essential differences between the NN and the CC contexts: a double contrast would have very little informative content for the hearer as it would urge her to look for a 'gestalt' (i.e., antecedent) which has a different shape and a different colour, which is nearly tantamount to creating an entirely new object and as such has very little descriptive content.

5. Perspectives: other languages

This study was based on the analysis of Dutch, a *plastic* language (Vallduví, 1991) with a relatively fixed word order and in which deaccenting is common. An interesting question is if and how *non*-plastic languages such as Italian use prosodic

cues to mark contrastiveness. This question is addressed in Swerts et al. (1999), who report on a comparative study of Dutch and Italian using the experimental set-up described in Section 2.1. This comparative study revealed that both Italian and Dutch speakers signal information status prosodically, but in a rather different way. In Dutch, as we have seen above, accent distribution is the main discriminative factor: new and contrastive information are accented, while given information is *deaccented*. A contrastive intonation could be distinguished, however, because contrastive accents generally were the sole accent in the phrase and always had the shape of a nuclear accent even in non-default positions. In Italian, distribution is not a significant factor, since within the elicitated NPs both adjective and noun are always accented, regardless of the information status. This does not entail that information status is not encoded in Italian prosody. For instance, even though given information is *reaccented*, a gradient difference was found in that given accents are always judged to be less prominent (mainly due to a lower pitch range) than other accents. Similarly, contrastive accents are judged to be more prominent, but only in the prefinal case (the noun in the Italian data).

Thus, both Dutch and Italian speakers can use prosodic means to signal givenness, and only the Dutch speakers have the additional possibility to make a systematic prosodic distinction between newness and contrastiveness. That Dutch speakers can make more prosodic distinctions than Italian ones can perhaps be attributed to the fact that the Dutch prosodic repertoire, so to speak, is larger than the Italian. The Dutch speakers in our experiment can either accentuate or deaccentuate a particular word, and, additionally, have the possibility of placing a nuclear accent in a non-nuclear position. As a result, a Dutch listener can rely on two different information sources, namely accent distribution and accent type, to determine whether something is given, new or contrastive. The Italian speakers accentuate or reaccentuate a particular word, but the net effect is the same either way: the word in question receives a pitch accent. Thus Italians cannot use distribution to mark information status in the isolated NPs used in our experiment. However, they *can* use relative gradiency. This is most clear for the accents associated with given words, which consistently receive a low pitch, but also, in the case of prefinal words, there appears to be a slight tendency to associate single contrasts with a relatively high pitch range.

It should be added that Italian, being a non-plastic language, has other means besides prosody of marking information status. For instance, it has a freer word-order than plastic languages such as Dutch, and it is known to exploit this freedom to mark information status. However, the constraints of the experimental paradigm apparently did not offer enough room for Italian speakers to use word-order as an indicator of information status. Therefore it would be interesting to look for an experimental set-up in which speakers have more freedom to describe a particular state of affairs. This might also shed a different light on the reaccentuation/deaccentuation debate, given that Ladd claims that deaccentuation of *complete* NPs within a sentence is quite possible in languages like Italian, which is supported by data from D'Imperio (1997).

Acknowledgements

Krahmer was partly supported by the EC Language Engineering/Telematics Applications Program, project LE-1 2277 (VODIS). Swerts is also affiliated with the Fund for Scientific Research – Flanders (FWO). Thanks are due to Cinzia Avesani, Alan Cruttenden, Elizabeth Couper-Kuhlen, Ralph van Dinther, Bob Ladd, Olga van Herwijnen, Armin Kohlrausch, Maria-Paola D'Imperio, Jan-Roelof de Pijper, Huub Prüst, Saskia te Riele, Jacques Terken, Mariët Theune, Raymond Veldhuis, Carel van Wijk and an anonymous referee for suggestions and various kinds of other assistance. Needless to say, all opinions and errors in this paper are our own.

References

Bartels, C., Kingston, J., 1994. Salient pitch cues in the perception of contrastive focus. In: Bosch, P., van der Sandt, R. (Eds.), Focus and Natural Language Processing,

Proceedings of the Journal of Semantics Conference on Focus, IBM Working Papers, TR-80.94-006.
Blok, P., 1993. The interpretation of focus. Ph.D. thesis, University of Groningen.
Bolinger, D., 1961. Contrastive accent and contrastive stress. Language 37, 83–96.
Bolinger, D., 1986. Intonation and its Parts. Edward Arnold, London.
Brown, G., Currie, K., Kenworthy, J., 1980. Questions of Intonation. Croom Helm, London.
Chafe, W., 1974. Language and consciousness. Language 50, 111–133.
Chafe, W., 1976. Givenness, contrastiveness, definiteness, subjects, topics and points of view. In: Li, C. (Ed.), Subject and Topic. Academic Press, New York, pp. 25–55.
Couper-Kuhlen, E., 1984. A new look at contrastive intonation. In: Watts, R., Weidman, U. (Eds.), Modes of Interpretation: Essays Presented to Ernst Leisi. Gunter Narr Verlag, pp. 137–158.
Cruttenden, A., 1986. Intonation. Cambridge University Press, Cambridge.
Cutler, A., 1977. The context-dependence of intonational meanings. In: Papers from the 13th regional meeting. Chicago Linguistic Society, Chicago, pp. 104–115.
D'Imperio, M., 1997. Narrow focus and focal accent in the neapolitan variety of Italian. In: Proceedings of the ESCA Workshop Intonation, Athens, pp. 87–90.
Gussenhoven, C., 1983. On the Grammar and Semantics of Sentence Accent. Foris, Dordrecht.
Hendriks, H., Dekker, P., 1995. Links without locations. In: Dekker, P., Stokhof, M. (Eds.), Proceedings of the 10th Amsterdam Colloquium. ILLC, Amsterdam, pp. 339–350.
Keijsper, C.E., 1984. Vorm en betekenis van nederlandse toonhoogtecontouren. Forum der Letteren 25, 20–37, 113–126.
Ladd, D., 1980. The Structure of Intonational Meaning. Indiana University Press, Bloomington.
Ladd, D., 1983. Phonological features of intonational meaning. Language 59, 721–759.
Levelt, W., 1989. Speaking, From Intention to Articulation. MIT Press, Cambridge, MA.
Moore, B., Glasberg, B., Baer, T., 1997. A model for the prediction of thresholds, loudness and partial loudness. J. Audio Eng. Soc. 46, 224–240.
Pechmann, T., 1984. Überspezifizierung und Betonung in Referentieller Kommunikation. Ph.D. thesis, Mannheim University.
Pierrehumbert, J., Hirschberg, J., 1990. The meaning of intonational contours in the interpretation of discourse. In: Cohen, P., Morgan, J., Pollack, M. (Eds.), Intentions in Communication. MIT Press, Cambridge, MA, pp. 342–365.
Piwek, P., 1998. Logic, Information and Conversation, Ph.D thesis, Eindhoven.
Prevost, S., 1995. A Semantics of Contrast and Information Structure for Specifying Intonation in Spoken Language Generation. Ph.D. thesis, University of Pennsylvania.
Rooth, M., 1985. Association with Focus. Ph.D. thesis, University of Massachusetts, Amherst.
Rooth, M., 1992. A theory of focus interpretation. Natural Language Semantics 1, 75–116.
Schmerling, S., 1976. Aspects of English Sentence Stress. University of Texas Press, Austin.
Stevens, S., 1957. On the psychophysical law. Psychological Review 64, 153–181.
Swerts, M., Krahmer, E., in progress. Reconstructing Dialogue History (working title).
Swerts, M., Avesani, C., Krahmer, E., 1999. Reaccentuation or deaccentuation: a comparative study of Italian and Dutch. In: Proceedings of the 14th International Congress of Phonetic Sciences 99, San Fransisco, 1–7 August 1999.
Vallduví, E., 1991. The role of plasticity in the association offocus and prominence. In: Proc. ESCOL 7, pp. 295–306.
van Deemter, K., 1999. Contrastive stress, contrariety and focus. In: Bosch, P., van der Sandt, R. (Eds.), Focus: Linguistic, Cognitive and Computational Perspectives. Cambridge University Press, Cambridge.
Zwicker, E., Scharf, B., 1965. A model of loudness summation. Psychological Review 72, 3–26.