

# PITCH ACCENTS, BOUNDARY TONES AND TURN-TAKING IN DUTCH MAP TASK DIALOGUES

Johanneke Caspers

Phonetics Laboratory/Holland Institute of Generative Linguistics  
PO Box 9515, 2300 RA Leiden, The Netherlands; CASPERS@RULLET.LEIDENUNIV.NL

## ABSTRACT

The present paper reports on an investigation of Dutch Map Task dialogues, looking for specific melodic turn-keeping cues. The materials were divided into so-called Inter Pausal Units (cf. [1]), and all IPU boundaries were labeled as instances of either turn-keeping or turn-changing. The pitch accents and boundary tones preceding the IPU boundaries were labeled in the ToDI system [2]. Results indicate one particular melodic configuration that seems typically associated with turn-keeping: an accent-lending rise followed by level high pitch (H\* %). But also the level boundary tone (%) in itself and – to a lesser extent – the rising pitch accent (H\*) seem to function as melodic turn-keeping cues.

## 1. INTRODUCTION

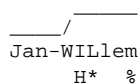
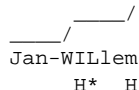
The alternation of speaking turns in every-day conversation is surprisingly smooth (generally one speaker at a time) and very fast (often no pause between successive speakers), which may be explained only in terms of a highly complex system of interacting factors, comprising syntax, semantics, pragmatics, prosody, visual cues etc. The relevance of intonation for the turn-taking process is recognized, but the knowledge of its precise function in the structuring of conversation is still fairly limited. The question asked in the present investigation is whether specific pitch accent types and/or boundary tone types function as turn-keeping cues in Dutch dialogue. To answer this question a corpus of Map Task dialogues (task-oriented spontaneous dialogues) was analyzed.

Within the so-called Conversation Analysis, the turn-taking mechanism is approached as a system of rules adhered to by the interlocutors. Sacks, Schegloff and Jefferson [3] set out the lines along which numerous empirical studies of natural conversation have been carried out, using detailed transcripts of natural conversation as a tool for finding out more about the systematics of conversational interaction. A central notion is the ‘Transition Relevant Place’, a point of possible completion of the present turn, where turn-taking becomes relevant (but not obligatory). Interlocutors *project* these points of possible turn completion before their actual occurrence – hence the often jointless transition between consecutive speakers – and my interest is in the role intonation may play in this projection of upcoming turn endings. Only recently, CA has taken an interest in speech melody as a factor in the organization of conversation (e.g. [4,5,6,7,8]). Note that the CA studies are qualitative in nature, which means that no frequency counts are given for specific variables, and that data are

presented in the form of detailed examples.

There is some evidence from the analysis of natural dialogue that pitch accents play a significant role in the projection of possible turn endings ([7,8] for German and English respectively). It seems that the shape of a potentially final pitch accent (plus the non-prominent tones that follow) can signal whether the current speaker wants to continue speaking or give the floor to someone else.

Furthermore, perception experiments [9] provided evidence for the relevance of a certain melodic configuration for the turn-taking system in Dutch. In these experiments an accent-lending rise followed by level pitch up to the end of the utterance (H\* %) was compared to an accent-lending rise followed by a final rise (H\* H%). Below stylized versions of the two contours are given on a short utterance (a proper name; the stressed syllable is capitalized, annotated in the ToDI system (‘Transcription of Dutch Intonation’, cf. [2]):

1.   
Jan-WILlem  
H\* %
2.   
Jan-WILlem  
H\* H%

The contours were presented in controlled environments, in which listeners were asked to perform a number of comparison and rating tasks. The results indicated that H\* % cannot readily be interpreted as a signal that the listener should take the next turn. Instead, use of this contour was understood as a sign that the current speaker wanted to keep the turn. The contour ending in a final rise (H\* H%) was interpreted as either turn-keeping or turn-yielding (for more details see [9]).

These findings led to the following research questions. Does an accent-lending rise followed by level high pitch (H\* %) function as a turn-holding cue in natural Dutch dialogue? Are there other melodic configurations with the same function?

## 2. APPROACH

A quantitative analysis of the melodic and turn-transitional structures of natural conversations should provide some insight into the role that intonation plays in the turn-taking system. In order to find out whether a systematic relationship exists between speech melody and turn-taking, we need to locate possible turn transition places in the dialogue, and then exam-

ine these places for the transition type actually occurring and the melodic characteristics immediately preceding these positions. I adopt the approach taken by Koiso et al. [1] who use so-called *Inter Pausal Units* (IPUs) to divide spontaneous dialogues into turn-relevant components. IPUs can be determined objectively and the boundaries between these units can be labeled as instances of either turn-holding or turn-changing. The accent-lending and boundary-marking configurations preceding every IPU boundary were described in terms of the ToDI model. After these labeling stages, the available data were viewed in the light of the research questions posed.

### 3. MATERIALS

Use was made of a corpus of Dutch guided spontaneous conversations, so-called Map Task dialogues (cf. [10]). In these task-oriented dialogues, maps provide a handle on an essentially spontaneous conversation. Two roles can be distinguished: an ‘instruction giver’ and an ‘instruction follower’. The former participant has a map with a route drawn on it and s/he has to explain to the instruction follower which route to draw on his or her unmarked copy of the map. The participants cannot see each other’s maps. Both maps have a number of reference points on it (e.g., ‘old pond’, ‘new pond’, ‘green meadow’) and by introducing small differences between these reference points it is possible to complicate the dialogue to some extent. The materials used for the present investigation were collected by Bob Ladd and Astrid Schepman.

### 4. ANALYSIS

#### 4.1. Inter Pausal Units

The materials were divided into *Inter Pausal Units*. An IPU is defined as “a stretch of a single speaker’s speech bounded by pauses longer than 100 ms” ([1], p. 299). This means that boundaries were drawn in all positions where a pause longer than 100 ms appeared in the signal, and in positions where a change of speaker occurred. In contrast with [1], overlapping stretches of speech were included in the analysis.

#### 4.2. Transition types

All IPU boundaries were characterized with respect to turn-transition type. A division was made between the categories *turn-holding* and *turn-changing*: the former meaning that the same speaker continues after a (turn-internal) pause of 100 ms or more, or after a backchannel uttered by the listener (a short optional utterance like ‘uh-huh’), the latter indicating that a turn change has occurred (i.e., the current listener takes the next turn). Interruptions were categorized as instances of turn-keeping, since the interrupted speaker *intended* to keep his or her turn. For more details the reader is referred to [11].

#### 4.3. Melodic Characteristics

The third stage of the analysis comprised a transcription of the melodic phenomena occurring immediately before each IPU

boundary. As a tool for labeling the melodic phenomena, the recently developed ToDI system was chosen. Starting from the principle that the “potential last accent” (cf. [7]) is of major importance to the turn-taking system, the last pitch accent before every IPU boundary was transcribed, as well as the tone sequence following this accent up to the boundary. The intonation was labeled on the basis of the auditory impression of the pitch curve only. Before every IPU boundary a boundary tone was transcribed, which means that intonation domain boundaries were determined by pauses or speaker changes actually occurring in the material, and not by the syntactic structure of the utterance. Note that, as a result, the boundary tones marked in the current analysis do not necessarily correspond to the boundary tones as defined by ToDI.

### 5. RESULTS

In total, eight complete Map Task dialogues were analyzed, amounting to over 40 minutes of speech. These materials contained 1552 IPU boundaries. It was hypothesized that an accent-lending rise followed by a level high tone (H\* %) would predominantly be followed by a turn from the same speaker. Table 1 presents the frequency of turn-changing versus turn-holding transitions immediately following the H\* % configurations found in the materials (N=140).

IPU transition type	frequency
+ CHANGE	19 (14%)
– CHANGE	121 (86%)
total	140 (100%)

**Table 1:** Absolute (and relative) frequency of turn-changing (+CHANGE) and turn-holding transitions (–CHANGE) after H\*%.

The table shows that in 86% of the cases the configuration H\* % is followed by an utterance by the same speaker (–CHANGE), which confirms the turn-holding character hypothesized of this melodic configuration. However, the amount of available data is relatively small, since only 140 of the 1552 IPU boundaries analyzed were preceded by H\* % (9%). The fact that there is no clear boundary tone in the H\* % configuration – the pitch remains flat until the following IPU boundary and does not drop to a clear low point (L%) or rise to a high point (H%) on the final syllable – may be crucial to the issue of turn-taking. In the ToDI system, nuclear contours ending in % are described as “half-completed”, clearly pointing to the unfinished nature of such melodic configurations.

The following question then presented itself: do *all* configurations ending in % generally function as turn-holding cues? Table 2 presents the number of turn-changing and turn-holding transitions broken down by boundary tone type. Note that the materials contained 94 cases that could not be transcribed (labeled “?”), partly due to considerable overlap in speech.

preceding boundary tone type	IPU transition type				total
	+ CHANGE		– CHANGE		
L%	275	(54%)	234	(46%)	509
H%	254	(45%)	315	(55%)	569
%	58	(15%)	322	(85%)	380
?	37	(39%)	57	(61%)	94
total	624	(40%)	928	(60%)	1552

**Table 2:** Absolute (and relative) frequency of turn-changing transitions (+CHANGE) versus turn-holding transitions (–CHANGE) per type of preceding boundary tone; grey cells contain the data for the level boundary tones (%).

Results show that in the majority of cases an IPU boundary was preceded by a low (N=509) or high (N=569) boundary tone. After an L% boundary tone the turn changes in 54% of the cases, and after a H% boundary tone the current speaker keeps the turn in 55% of the cases. IPU boundaries that are preceded by level non-low pitch (%), N=380) are followed by a transition that can be interpreted as turn-keeping in 85% of the cases (i.e., the current speaker continues after a pause, the current listener backchannels, or the current speaker is interrupted). A post-hoc analysis indicates that the percentage of turn-keeping transitions differs significantly for all three boundary tones: after L% there is a tendency to change the turn and after H% to hold the turn, but after % the effect is much more outspoken: in most cases the turn stays with the current speaker.

The results for the % boundary tone closely resemble those found for the H\* % configuration as a whole (cf. table 1); does this mean that the pitch accent type preceding the % boundary tone is irrelevant? Table 3 gives the frequency of turn-keeping and turn-changing transitions broken down by the different pitch accent types. Since a number of IPUs contained a boundary tone but no pitch accent – the majority of these cases constitute a filled pause ('uhm'), or function as a back-channel (cf. [12]) – the total number of cases in table 3 is considerably lower than in table 2.

Averaging over boundary tone types, a rising pitch accent (H\*) is followed by an utterance from the same speaker in 74% of the cases, whereas H\*L and L\*(H) accents are followed by a change of speaker in roughly half of the cases. A post-hoc analysis reveals that the H\* pitch accent is followed by a significantly higher percentage of turn-keeping transitions than each of the other three types of pitch accents; there were no further differences. Limiting the dataset to the IPUs that contain a – transcribable – pitch accent as well as a boundary tone, ANOVA shows main effects of boundary tone type ( $F_{(2,1081)} = 40.36, p < .001$ ) and pitch accent type ( $F_{(3,1080)} = 3.46, p < .05$ ) on the percentage of following turn-keeping transitions, as well as a significant interaction ( $F_{(6,1072)} = 4.51, p < .01$ ). This means that, in addition to a large effect of boundary tone type

and a much smaller effect of pitch accent type, there is a combined influence of both tone types on the chances for turn-keeping at the following IPU boundary. For lack of space this issue will not be further elaborated upon here.

So far, the available data present no clearly turn-yielding melodic configurations.

## 6. CONCLUSION AND DISCUSSION

The data presented in section 5 provide support for the hypothesis that an accent-lending rise followed by level pitch (H\* %) functions as a turn-keeping cue in Dutch spontaneous dialogues. Furthermore, the data show that when there is no clear boundary tone immediately before an IPU boundary, there is no subsequent speaker change in approximately 85% of the cases, in contrast with low or high boundary tones, which are followed by a turn change in roughly half of the cases. The results indicate that the absence of a clear boundary tone is generally taken as a sign that the current speaker has not finished yet (and wants to continue).

The findings are in concurrence with results presented in [9], where an accent-lending rise followed by a high boundary tone (H\* H%) was found to be ambiguous between turn-keeping and turn-yielding, whereas an accent-lending rise followed by level pitch (H\* %) was interpreted by subjects as a signal for turn-keeping. The findings also support the intuition of [7] and [9] that intonation does not seem to play an unequivocal role in yielding the turn (in German and Dutch respectively), whereas there does seem to be evidence for the existence of melodic turn-keeping cues.

In their paper presenting a qualitative analysis of British English conversation, [8] conclude that there are specific "TRP-projecting accents" – pitch accents that project a Transition Relevant Place – as well as specific non-TRP-projecting accents. The current findings indicate that Dutch also has a non-TRP-projecting accent, but the data present no evidence for the existence of specific TRP-projecting configurations.

preceding pitch accent type	IPU transition type		total
	+ CHANGE	– CHANGE	
H*	64 (26%)	186 (74%)	250
H*L	392 (52%)	365 (48%)	757
L*	11 (61%)	7 (39%)	18
L*H	34 (53%)	30 (47%)	64
?	43 (43%)	58 (57%)	101
total	544 (46%)	646 (54%)	1190

**Table 3:** Absolute (and relative) frequency of turn-changing transitions (+CHANGE) versus turn-holding transitions (–CHANGE) per type of preceding pitch accent; grey cells contain the data for the rising pitch accents (H\*).

Further quantitative research is needed to determine whether English differs from Dutch and German in the presence versus absence of specific melodic configurations that project an upcoming change of speaker, or whether English exploits melodic turn-holding devices but no unique melodic turn-yielding cues, as seems to be the case in Dutch and German.

## 7. ACKNOWLEDGMENTS

This research was funded by the Netherlands Organization for Scientific Research (NWO) under project #355-75-002.

## 8. REFERENCES

1. Koiso, H., Horiuchi, Y., Tutiya, S., Ichikawa, A., and Den, Y., “An analysis of turn-taking and backchannels based on prosodic and syntactic features in Japanese Map Task dialogs”, *Language and Speech* 41, 295-321, 1998.
2. Gussenhoven, C., Rietveld, T. and Terken, J., “ToDI, Transcription of Dutch Intonation”, URL <http://lands.let.kun.nl/todi/>, 1999.
3. Sacks, H., Schegloff, E.A., and Jefferson, G., “A simplest systematics for the organization of turn-taking for conversation”, *Language* 50, 696-735, 1974.
4. Auer, P., “On the prosody and syntax of turn-continuations”, in E. Couper-Kuhlen and M. Selting (eds), *Prosody in conversation*, Cambridge University Press, Cambridge, 57-100, 1996.
5. Couper-Kuhlen, E., “Intonation and clause combining in discourse: the case of ‘because’”, *Pragmatics* 6, 389-426, 1996.
6. Ford, C.E., and Thompson, S.A., “Interactional units in conversation: syntactic, intonational, and pragmatic resources for the management of turns”, in E. Ochs, E.A. Schegloff and S.A. Thompson (eds), *Interaction and grammar*, Cambridge University Press, Cambridge, 134-184, 1996.
7. Selting, M., “On the interplay of syntax and prosody in the constitution of turn-constructural units and turns in conversation”, *Pragmatics* 6, 357-388, 1996.
8. Wells, B., and Macfarlane, S., “Prosody as an interactional resource: turn-projection and overlap”, *Language and Speech* 41, 265-294, 1998.
9. Caspers, J., “Who’s next? The melodic marking of question versus continuation in Dutch”, *Language and Speech* 41, 375-398, 1998.
10. Anderson, A.H., Bader, M., Gurman Bard, E., Boyle, E., Doherty, G., Garrod, S., Isard, S., Kowtko, J., McAllister, J., Miller, J., Sotillo, C., and Thompson, H.S., “The HCRC Map Task Corpus”, *Language and Speech* 34, 351-366, 1991.
11. Caspers, J., “Looking for melodic turn-holding configurations in Dutch”, in H. de Hoop and T. van der Wouden (eds), *Linguistics in the Netherlands 2000*, John Benjamins, Amsterdam, to appear.
12. Caspers, J., “Melodic characteristics of backchannels in Dutch spontaneous dialogues”, *Proceedings of the 6<sup>th</sup> International Conference on Spoken Language Processing*, [this volume], 2000.