On the Function of Low and High Boundary Tones in Dutch Dialogue

Johanneke Caspers

Phonetics Laboratory, Universiteit Leiden Centre for Linguistics

Cleveringaplaats 1, 2311 BD Leiden, The Netherlands

E-mail: j.caspers@let.leidenuniv.nl

ABSTRACT

The main question asked in the present investigation is the following: do low boundary tones (L%) signal optional speaker transitions, while high boundary tones (H%) mark obligatory transitions? Optional speaker transitions concern cases where the same speaker could continue, provided that no other interlocutor takes the next turn. Obligatory transitions may involve a change of speaker (question – answer), but they can also implicate turn-holding (comma). In a corpus of Dutch task-oriented dialogues all prepausal stretches of speech were categorized as leading to either obligatory or optional speaker transitions, and these data were compared with the boundary tones occurring immediately before the pause, as well as the transitions directly following the pause. Results suggest that speakers indeed use H% to signal that the listener is expected to do something specific: either take the next turn, or, just the opposite, keep listening, while L% does not invite such specific turn-taking behavior.

1. INTRODUCTION

What could be the function of the high (H%) and low (L%) boundary tones in the Dutch turn-taking system? This question is asked as part of a larger investigation of melodic elements as turn-relevant cues in Dutch dialogue.

Many languages have low and high boundary tones, which means that intonation contours may end in a low final tone (labeled L%), as well as a high tone (H%). According to Cruttenden [1], final low pitch is associated with a non-continuative 'closed' meaning, while final high pitch is associated with a continuative 'open' meaning. A high boundary tone typically occurs in questions, signaling to the listener that an answer should be provided. However, a steep rise in pitch on the final syllable does not signal interrogativity in all cases, since it can also be used as 'comma' intonation, indicating that the speaker is not finished yet [2]. In this case the listener is not prompted for a reaction, but for continued attention.

In the intonation model used in the present investigation ('ToDI' [3,4]), a third boundary tone type is distinguished: the level boundary tone (labeled %). Caspers [5] reports a clear relationship between turn-holding and level boundary tones (%) for a corpus of Dutch dialogue, while no such

relationship was found for the low and high boundary tones and turn-changing or turn-holding. The relevant data are presented in table 1.

boundary tone	following speaker transition		
	change	hold	
H%	188 (52%)	177 (48%)	365
L%	272 (54%)	233 (46%)	505
%	55 (15%)	323 (85%)	378
total	515 (41%)	733 (59%)	1248

Table 1: Absolute (and relative) frequency of speaker changes and speaker holds following high (H%), low (L%) and level (%) boundary tones.[1]

The data show no clear relationship between turn-taking and low or high boundary tones, since both tones are followed by roughly 50% speaker changes and 50 % speaker holds (χ^2 = .47, ins.). In contrast, level boundary tones are followed by turn-holding in 85% of the cases.

Inspired by [6], Herman [7] proposes that H% signals that the phrase to come forms some kind of unit with the previous one, while L% indicates that the information just given can be dismissed from the interlocutors' attention. Translated to the turn-taking system this could mean that H% is associated with *obligatory* aspects of the turn-taking system, which may involve either a change of turn – e.g., the answer to a question – or turn holding – e.g., continued speech from the same speaker after a pause following an incomplete message. In contrast, the low boundary tone L% only marks the completeness of a domain and is therefore associated with *optional* turn-taking, because a speaker change is possible in these cases, but never obligatory [8].

The following hypothesis was formulated: L% signals that the next speaker transition is optional, while H% announces an obligatory change of speaker, or an obligatory continuation of the current speaker.

2. APPROACH

The materials used in the present investigation are Dutch Map Task dialogues dialogues, where one speaker – the 'giver' – has to explain to the other – 'the follower' – how to draw a route on a map; there are deliberate discrepancies between the two maps, leading to many questions,

interruptions, etc. [9]. The task-oriented nature of the dialogues makes it relatively easy to determine the function of specific utterances, which is important when deciding on the optionality of the following speaker turn.

For all possible speaker transition locations it was decided (a) whether an obligatory change of speaker was expected, an obligatory turn-continuation, or an optional change of speaker. In addition (b) the boundary tones immediately preceding these locations were determined, as well as (c) the speaker transition type actually occurring at these points, i.e., 'change' versus 'hold'.

3. METHOD

The materials consist of a corpus of eight Map Task dialogues, amounting to over 40 minutes of speech. The dialogues were divided into so-called Inter Pausal Units ('IPU's, [10]), stretches of a single speaker's speech bounded by pauses longer than 100 ms. These units were chosen as an – objectively determinable – approximation of turn-relevant components. All boundaries between IPUs were labeled as instances of either turn-changing or turn-holding. The boundary tone preceding each IPU boundary was transcribed using 'ToDI' [3]. The reader is referred to [5,11] for more detailed information on the analysis of the corpus.

For the present investigation a prediction was needed with respect to the speaker transition following each IPU boundary. As a first step, the following analysis of the possible dialogue function of the IPUs was made, based largely on syntactic properties:

(i) obligatory speaker change

- after a syntactic (yes/no or wh-) question, or any other initial part of an adjacency pair (e.g., a greeting and a reply)

(ii) obligatory speaker hold

- after a (filled) pause at a location where no syntactic completion point is reached

- after a backchannel (an optional 'go-on' signal, uttered by the current listener [12])

- after a 'cue phrase', e.g., "Nou Xavier" ('Well Xavier')

(iii) optional speaker transition

- after a statement

- after an acknowledgment (non-optional)

4. RESULTS

4.1 Boundary tones and syntactic structure

In table 2 the three categories of predicted speaker transitions are compared with the boundary tones found at the end of the IPUs.

predicted speaker transition	boundary tone type			
	H%	L%	%	total
obl. change	85 (71%)	25 (21%)	9 (8%)	119
obl. hold	35 (9%)	69 (19%)	269 (72%)	373
opt. transition	244 (33%)	407 (54%)	99 (13%)	750
total	364 (29%)	501 (40%)	377 (31%)	1242

Table 2: Absolute (and relative) frequency of high (H%), low (L%) and level (%) boundary tones, broken down by predicted speaker transition (obligatory speaker change, obligatory speaker hold and optional speaker transition; 6 cases could not be classified).

The data show that a large majority of the IPUs leading to obligatory speaker changes – mostly questions – end in a high boundary tone (71%), as was hypothesized. In these cases the H% tone seems to support the function already expressed by the grammatical structure of the utterance, viz., asking a question. When a continuation of the same speaker was predicted ('obligatory hold'), most IPUs end without a 'real' boundary tone (72% level boundary tones). These findings concur with earlier findings [5]; they indicate that obligatory speaker continuations are not typically marked by high boundary tones when no grammatical completion point is reached yet (only 9%). The predicted optional speaker transitions are preceded by an L% in 54% of the cases, which means that there is weak support for the hypothesis that optional transitions are signalled by a low boundary tone; however, a third of the boundaries characterized as 'optional' are preceded by a high boundary tone.

Two questions present themselves: (1) what are the *actual* speaker transitions occurring at the IPU boundaries, and (2) what could be the role of the H% boundary tone within the category of 'optional transitions'?

Table 3 presents the absolute and relative number of speaker changes and speaker holds occurring in the materials for the three types of predicted speaker transition.

predicted speaker transition	speaker transition type		
	change	hold	total
obligatory change	107 (90%)	12 (10%)	119
obligatory hold	28 (8%)	345 (92%)	373
optional transition	375 (50%)	375 (50%)	750
total	510 (41%)	732 (59%)	1242

Table 3: Absolute (and relative) frequency of speaker changes and speaker holds for the three types of predicted speaker transition.

The table shows a clear correspondence between predicted and actual speaker transitions: when an obligatory change is predicted the turn changes in 90% of the cases, while the same speaker continues in 92% of the cases when a continuation is expected; optional transitions are followed by 50% speaker changes.

A hierarchical loglinear analysis performed on the factors *predicted speaker transition*, *boundary tone type* and *speaker transition type* reveals a significant association between *predicted speaker transition* and *speaker transition type* (partial $\chi^2 = 224.38$; $p < 0.0001$), between *boundary tone type* and *speaker transition type* (partial $\chi^2 = 31.04$; $p < 0.0001$), and between *predicted speaker transition* and *boundary tone type* (partial $\chi^2 = 358.47$; $p < 0.0001$), but no third-order interaction among the three factors (Pearson $\chi^2 = .52$, ins). The fact that there is no significant combined influence of the structure of the IPU and the boundary tone on the type of speaker transition occurring ('hold' vs. 'change') suggests that there is no independent effect of the boundary tone on the following speaker transition. However, the number of H% boundary tones within the category of optional speaker changes is quite high (N = 244), and this incites further investigation of the possible function of boundary tones within the category of 'optional speaker changes' (i.e., acknowledgments and statements).

4.1 Boundary tones and pragmatic structure

The group of IPUs labeled as 'optional' consists of a small number of acknowledgements (N = 115) and a larger group of statements (N = 635). It is possible that the H% boundary tone marks some obligatory aspect of the turn-taking system within the group of statements as well: a final rising tone may change a statement into a question, leading to an automatic speaker change, but it can also be meant as a 'comma', which would be followed by continued speech from the same speaker. The question then is: how to distinguish between these two functions of H%? Syntactic information is irrelevant here, but the task-oriented nature of the dialogue provides a more pragmatic source of information. The task of the two speakers is clear: giving instructions and carrying them out, respectively. It is therefore relatively easy to determine whether a particular statement is part of a longer instruction, or whether it is meant as a question.

For example: the utterance "Dan ga je recht naar beneden" ('Then you go straight down') spoken by the instruction giver and ending in a high boundary tone is clearly intended as part of an instruction that is still to be completed. In contrast, the following example illustrates the use of the high boundary tone as a question marker. In a reaction to a particular instruction the follower says: "Dus vanaf het begin om de woonwagen heen" ('So from the beginning around the trailer'); spoken with a high final tone this utterance can only be interpreted as a question.

A further group of statements ending in a high boundary tone could not be classified as either 'comma' or 'question'. Two examples are given, uttered by the instruction giver: "Laat mij nou maar gewoon vertellen!" ('Let me just simply tell!'), and "Maar dat zal wel weer niet kloppen" ('But that will certainly again be wrong'). In both cases, the H% seems to function as some sort of exclamation mark.

Table 4 presents the number of speaker changes and holds following statements ending in an H% tone, characterized as 'comma', 'question' or 'exclamation', and broken down by speaker task.

IPU function	speaker task	speaker transition type		total
		change	hold	
comma	giver	37 (25%)	110 (75%)	147
	follower	2 (50%)	2 (50%)	4
question	giver	13 (81%)	3 (19%)	16
	follower	22 (100%)	-	22
exclamat.	giver	8 (62%)	5 (38%)	13
	follower	7 (70%)	3 (30%)	10
total		89 (38%)	123 (62%)	212

Table 4: Absolute (and relative) frequency of speaker changes and speaker holds following syntactic statements ending in a high boundary tone, for the three IPU functions distinguished (comma, question or exclamation), broken down by speaker task.

The table shows that the H% on statements may be characterized as comma intonation in the large majority of cases (69%), and the same speaker indeed continues speaking in 74% of these cases, providing support for the idea that H% marks an obligatory aspect of the turn-taking system (viz., obligatory hold). These utterances are almost exclusively produced by the instruction giver. A smaller group of statements ending in H% is categorized as a question (18%), produced by both 'giver' and 'follower'; in most of these cases (92%) the turn (obligatorily) changes. For the small group of exclamations the turn changes in two-thirds of the cases.

A hierarchical loglinear analysis performed on the factors *speaker task*, *IPU function* and *speaker transition type* reveals a significant association between *IPU function* and *speaker transition type* (partial $\chi^2 = 34.09$; $p < 0.0001$), between *speaker task* and *IPU function* (partial $\chi^2 = 42.73$; $p < 0.0001$), but no third-order interaction (Pearson $\chi^2 = 7.04$, ins). A logit analysis reveals that the 'comma' ($z = -3.57$) and 'question' ($z = 2.98$) categories have significant – and opposite – effects on the type of speaker transition following the IPU ('change' versus 'hold').

There are 39 cases where an instruction ending in H% (i.e., where the high final tone presumably marks the wish of the speaker to continue after the IPU boundary) is followed by a change of speaker. These cases seem in conflict with the view that the high boundary tone signals obligatory turn-keeping here. However, inspection of the individual cases revealed that they can all be labeled as 'neat' interruptions: the instruction giver seems to be interrupted in the middle of a larger instruction; the follower 'misuses' the pause realized by the giver as an opportunity to take over the turn. For example: "Ja ho es effentjes, je stuurt me net omhoog!" ('Just wait a second, you just sent me up!'). This type of interruption can be contrasted with 'real' interruptions, where the 'interruptor' starts speaking in the middle of an utterance from the current speaker [5,11], usually leading to some sort of competition for the turn [13]. It seems that in the present 39 cases the instruction giver loses the turn, in spite of the signal that (s)he wanted to

keep it. This finding is not completely surprising, since a lot of problems are to be solved while conducting a Map Task dialogue, as a result of the deliberate discrepancies between the two maps.

5. CONCLUSION

The general results indicate that boundary tones more or less reflect the syntactic structure of the utterances they occur on (cf. table 2):

1. most syntactic questions end in a high boundary tone (H%) and are followed by a change of speaker

2. most syntactically incomplete structures end in a level boundary tone (%), i.e., without a 'real' boundary tone, and are not followed by a change of speaker

3. the majority of statements end in a low tone (L%) and are followed by speaker changes as well as speaker holds

These data are in line with the view that a high boundary tone signals that the phrase to come forms some kind of unit with the previous one [7], leading to an *obligatory* change of speaker when occurring on a syntactic question, while a low boundary tone signals that some kind of endpoint is reached [7], leaving open the *opportunity* for the other speaker to take over the turn. The level tone marks the fact that no grammatical endpoint is reached yet [5], which in terms of the turn-taking system may be translated into 'please do not take over the turn'.

The data further indicate that, in addition to signaling obligatory speaker transitions for syntactic questions, high boundary tones may be used to signal a specific dialogue function for 'neutral' syntactic structures as well. Within the large group of statements found in the materials, those ending in an H% boundary tone can be analyzed as leading to either obligatory turn-changes – the listener is expected to provide an answer to a question – or obligatory turn-holds – the listener is expected to keep on listening because there is more to come after the next syntactic boundary. In either case the following grammatical unit is marked as essential for a sensible progress of the conversation.

FOOTNOTE

1   The data deviate minimally from the data presented in [5], because a small number of errors were corrected.

REFERENCES

[1] A. Cruttenden, *Intonation*, Cambridge: Cambridge University Press, 1997.

[2] J. Caspers, "Who's next? The melodic marking of question versus continuation in Dutch," *Language and Speech*, vol. 41, pp. 375-398, 1998.

[3] C. Gussenhoven, T. Rietveld and J. Terken, "ToDI, Transcription of Dutch Intonation, First edition," URL http://lands.let.kun.nl/todi/, 1999.

[4] C. Gussenhoven, "Transcription of Dutch Intonation (ToDI)," In *Prosodic Typology and Transcription: A Unified Approach*, Sun-Ah Jun, Ed. Oxford: Oxford University Press, forthcoming.

[5] J. Caspers, "Local speech melody as limiting factor in the turn-taking system in Dutch," *Journal of Phonetics*, in press.

[6] J.B. Pierrehumbert and J. Hirschberg, "The meaning of intonational contours in the interpretation of discourse," in *Intentions in communication*, P.R. Cohen, J. Morgan and M.E. Pollack, Eds., pp. 271-311. Cambridge MA: MIT Press, 1990.

[7] R. Herman, "Phonetic markers of global discourse structure in English," *Journal of Phonetics*, vol. 28, pp. 466-493, 2000.

[8] H. Sacks, E.A. Schegloff and G. Jefferson, "A simplest systematics for the organization of turn-taking for conversation," *Language*, vol. 50, pp. 696-735, 1974.

[9] A.H. Anderson, M. Bader, E. Gurman Bard, E. Boyle, G. Doherty, S. Garrod, S. Isard, J. Kowtko, J. McAllister, J. Miller, C. Sotillo and H.S. Thompson, "The HCRC Map Task Corpus," *Language and Speech,* vol. 34, pp. 351-366, 1991.

[10] H. Koiso, Y. Horiuchi, S. Tutiya, A. Ichikawa and Y. Den, "An analysis of turn-taking and backchannels based on prosodic and syntactic features in Japanese Map Task dialogs," *Language and Speech,* vol. 41, pp. 295-321, 1998.

[11] J. Caspers, "Looking for melodic turn-holding configurations in Dutch," in *Linguistics in the Netherlands 2000*, H. de Hoop and T. van der Wouden, Eds., pp. 45-55. Amsterdam: John Benjamins, 2000.

[12] J. Caspers, "On the function of the late rise and the early fall in Dutch dialogue: a perception experiment," *Proceedings of the 7th International Conference on Spoken Language Processing*, Denver, pp. 537-540, 2002.

[13] B. Wells and S. Macfarlane, "Prosody as an interactional resource: turn-projection and overlap," *Language and Speech,* vol. 41, pp. 265-294, 1998.