THE FUNCTION OF INTONATION IN SPONTANEOUS AND READ DIALOGUE

J. C. Kowtko

CSTR & HCRC, University of Edinburgh, U.K.

ABSTRACT

Independently motivated analyses of intonation and discourse function are employed in a study of intonation function in task-oriented dialogue (taken from the HCRC Map Task Corpus [1]). Results from a comparison of the two analyses performed on single-word utterances in spontaneous dialogue show that intonation contour is a significant factor in utterance function. An examination of read-aloud dialogue reveals similar results with different patterns of correlation.

INTRODUCTION

Recent work on intonation in dialogue tends to follow one of two opposite approaches: it either describes very general discourse functions (e.g. [6] connecting, continuing and segmenting) or it identifies very specific discourse contexts (e.g. [3] on anaphor distribution and turn-taking). In order to make progress in this area, we need to combine these two approaches. This in turn requires an independent description of dialogue context as the basis for a robust account of intonational function. Such an independent description is the conversational games analysis outlined in [4]. It is employed here in a study of intonation function in single-word utterances taken from spontaneous and read-aloud dialogue.

Using the conversational games analysis to represent discourse function, and assuming that intonation plays a significant role in discourse function, we expect to find tune patterns by looking at move x in game y. We also expect to discover a difference between results in spontaneous and read-aloud dialogue. The nature of the difference interests us because we could possibly use the more easily constrained read-aloud dialogue to train speech recognisers.

DIALOGUE ANALYSIS

The analysis proposed by Kowtko *et al.* [4] involves interactional exchanges in dialogue, called *conversational games*. Games embody the linguistic interaction, e.g. initiation, response and feedback, which arises from non-linguistic goals. (See [7] for a discussion on the relationship between linguistic processes and underlying action goals.) A game consists of the turns necessary to accomplish a conversational goal. The components of games, *moves*, are defined in terms of speaker intention and dialogue function.

The data used in this study are based upon dialogues arising from the map task [1]: One person has a map with a route marked on it and has to tell another person how to draw the route onto a similar map. Neither participant can see the other's map.

The nature of the task is such that the speakers' intentions are clear to the analyst most of the time. Kowtko *et al.* report that one expert and three naïve judges agree on an average of 83% of the moves classified in two map task dialogues.

Six games appear in the dialogues: Instructing, Checking, Querying-YN, Querying-W, Explaining, and Aligning. They are initiated by the following moves: INSTRUCT (Provides instruction), CHECK (Elicits confirmation of known information—tests speaker's knowledge), QUERY-YN (Asks yes-no question for unknown information), QUERY-W (Asks content, *wh*-, question for unknown information), EXPLAIN (Gives unelicited description), and ALIGN (Aligns hearer's knowledge or beliefs—checks alignment of position in task).

Six other moves provide response and additional feedback: CLARIFY (Clarifies or rephrases given information), REPLY-Y (Responds affirmatively), REPLY-N (Responds negatively), REPLY-W (Responds with requested information), ACKNOWLEDGE (Acknowledges and requests continuation), and READY (Indicates intention to begin a new game).

Since the map task involves one participant instructing the other concerning how to draw the route, the conversations naturally consist of many Instructing games. The conversational games analysis allows for nesting of games and looping of response and feedback moves within games. The prototypical game consists of two or three moves: initiation, response, and optionally feedback. The large majority of games (84% from a sample of 3 dialogues, $n = 65$) match the simple prototype. Games that do not match the prototype are still well-formed, having extra response–feedback loops, nested games, or extra moves. Very few games (less than 2%) break down as a result of a misunderstanding or other problem.

Conversational game structure offers a taxonomy which specifies both the function and context of an utterance, as a specific move at a specific point within a specific game. This facilitates the study of the function of intonational tune since the tune reflects an utterance's conversational role.

INTONATION ANALYSIS

Single-word utterances which comprise whole conversational moves from the HCRC Map Task Corpus ([1] containing Scottish English) have been analysed in terms of intonation contour. A set of five intonational tunes was determined to best represent the data: High Level (Hi; level tune high in the speaker's local pitch range), Low Level (Lo; level tune low in the speaker's local pitch range), Fall (F; simple fall in pitch overall), Rise (R; simple rise in pitch overall), Fall-Rise (F-R; distinct falling pitch followed by a rise). The complementary sixth tune, rise-fall, did not appear in the data examined.

FUNCTION OF INTONATION

A study comparing intonation contour and dialogue function was performed on single words which individually comprise conversational moves and intonational phrases in the HCRC Map Task Corpus (120 from spontaneous dialogues and 120 from dialogues read aloud by the original participants using transcripts of the spontaneous dialogues). The data come from three whole dialogues. They involve 1 male and 5 female speakers who do not know each other, and total 12 minutes of spontaneous speech and 11 minutes of read speech. They sound fairly natural in the read-aloud condition and have good quality audio recordings.

All single-word moves in the three dialogues were considered as data points. The only words removed from the study are those for which the pitch trace failed (e.g. croaky, unintelligible utterances) and those which form partial intonational phrases.

The words used in the study are *almost, aye, ehm, mmhmm, no, okay, okey-dokey, right, rightee-ho, uh-huh, yeah, yes,* and *yup.* They appear as 6 of the conversational moves [4]: ALIGN, REPLY-Y, REPLY-N, REPLY-W, ACKNOWLEDGE, and READY. Each word is transcribed intonationally as high level (Hi), low level (Lo), rise (R), fall (F), or rise-fall (F-R). Their discourse function is represented by the framework of conversational games. The intonation of these words was compared with discourse function in terms of the move type and the game in which

it occurs.

Spontaneous Results

Results are shown with three intonation categories: rise + high level (R+Hi), low level + fall (Lo+F), and fall-rise (F-R). The motivation for this clustering is two-fold. Firstly, the level tunes may be phonetic variations of underlying pitch accent glides. Many of the low level tunes, for example, exhibit a detectable fall in pitch of approximately 8 Hz, but were categorised as level overall. Secondly, the results show some clustering of intonation categories, especially between low level and falling tunes.

From the spontaneous dialogues, significant correlations[1] emerge between intonation and discourse categories. Table 1 displays the results in terms of intonational tune associated with conversational move in a game. ALIGN moves significantly rise in pitch. The majority of REPLY-Y tunes fall, and the category is significantly falling and low level. REPLY-N moves, although small in number, exclusively fall in pitch. Most ACKNOWLEDGE moves are rising or high level. The two READY moves fall in pitch.

Looking at ACKNOWLEDGE moves in their game context adds some clarity to the pattern of tunes. In *Instructing* games, rises and high levels form a majority. Moves in *Querying-W* games show a similar significant pattern of rising tunes. Results in *Querying-YN*, *Checking*, and *Explaining* games are insignificant although the first two tend toward falling and low level, and the third toward low level tunes.

These correlations indicate that discourse function, as defined by move type, is a principal factor in determining intonation contour.

Read Results

Results of the analysis of read aloud dialogue also show significant correlations between intonation and discourse

---

Table 1: *Spontaneous Intonation v. Conversational Move*

Move	R+Hi	Lo+F	F-R	#
ALIGN	†7	1		8
REPLY-Y	6	†23		29
REPLY-N			4	4
REPLY-W	1			1
ACKNWL.	41	32	3	76
READY		2		2

categories. They reveal a pattern somewhat different from that in spontaneous speech. The tunes in some discourse categories have shifted toward low level and falling tunes. Table 2 shows a summary.

ALIGN moves significantly rise. REPLY-Y moves are significantly falling and low level. REPLY-N moves are all falling and low level tunes. The majority of ACKNOWLEDGE moves are falling and low level. READY moves fall in pitch.

Examining the moves in their game contexts reveals a pattern different from the overall one only in the case of ACKNOWLEDGE moves. All of the rising tunes except one occur within *Instructing* games. Here the data are shared between the rising and falling categories. Moves in *Querying-W* games tend toward falling and low level tunes. ACKNOWLEDGE moves in *Querying-YN* and *Checking* games are all falling or low level. The moves in *Explaining* games are exclusively level in pitch, and most of these are high level tunes.

Again, the significant correlations and clear trends found in the results of read dialogue suggest that intonation contour is an important factor in determining discourse function.

Discussion: Spont. v. Read

One of the most noticeable differences between spontaneous and read dialogue is a shift in the read condition toward low level and falling tunes in the response moves (REPLY-Y, REPLY-N, REPLY-W, and ACKNOWLEDGE). This shift is consistent with general knowledge about reading style and various

---

Table 2: *Read Intonation v. Conversational Move*

Move	R+Hi	Lo+F	F-R	#
ALIGN	†8			8
REPLY-Y	5	†24		29
REPLY-N			4	4
REPLY-W		1		1
ACKNWL.	28	47	1	76
READY		2		2

comparisons of the two modes of speech (e.g. [2] which states that at boundaries, read speech contains more falling tunes).

The increased presence of low level and falling tunes in read dialogue is most visible in ACKNOWLEDGE moves. There is a shift from a majority rise and high level pitch pattern to a fall and low level one. ACKNOWLEDGE moves in *Querying-W* games change from a significant pattern of rising tunes to a trend toward falling and low level tunes.

CONCLUSION

Discourse structure correlates with intonation contour, showing that intonation is a significant factor in dialogue function. The similar patterns in results from spontaneous and read dialogue encourage us regarding the use of read speech data to train speech recognition systems. They suggest that read-aloud dialogue could provide a partial substitute for that of spontaneous dialogue. We must, however, be careful in accepting read dialogue for analysis because of the increased number of low level and falling tunes in read speech. The increased number of these tunes could be related to the speaker's level of confidence while performing the task. In reading the transcript and acting out the map task dialogue, the participant may feel more confident than while originally doing the map task. This possibility is left to future research.

ACKNOWLEDGEMENT

REFERENCES

[1] Anderson, A H, M Bader, E G Bard, E Boyle, G Doherty, S Garrod, S Isard, J Kowtko, J McAllister, J Miller, C Sotillo, H Thompson, and R Weinert (1991), "The HCRC Map Task Corpus," *Language and Speech*, 34: 351-366.

[2] Blaauw, E (1994), "The contribution of prosodic boundary markers to the perceptual difference between read and spontaneous speech," *Speech Communication*, 14: 359-375.

[3] Hockey, B A (1992), "Prosody and the interpretation of cue phrases," *Proceedings of the IRCS Workshop on Prosody in Natural Speech*, University of Pennsylvania, IRCS Report No.: 92-37, 71-77.

[4] Kowtko, J C, S D Isard and G M Doherty-Sneddon (1992), "Conversational games within dialogue," Research Paper HCRC/RP-31, Human Communication Research Centre, University of Edinburgh.

[6] McLemore, C A (1991), *The Pragmatic Interpretation of English Intonation: Sorority Speech*, Ph.D. dissertation, University of Texas at Austin.

[7] Power, R (1979), "The organisation of purposeful dialogues," *Linguistics*, 17: 107-152.