

## PHONOLOGICAL SIGNALLING OF TOPIC MOVEMENT IN CONVERSATIONAL DISCOURSE: AUTOMATIC TAGGING IN SPOKEN LANGUAGE CORPORA

Christine Cheepen  
University of Hertfordshire, UK.

### 1. INTRODUCTION

This paper focusses on some of the phonological features of unscripted, conversational discourse, and their relevance to the overall structure of that discourse. The identification of the features and consequent discussion is conducted in the context of recommendations for automatic tagging of such features in computationally stored spoken language corpora. The particular feature of conversation which is the main focus of this paper is the phenomenon of mutually congruent evaluation, which functions to section the discourse in terms of topic movement, and which promotes the main aim of conversation - speaker solidarity.

Over the past decade there has been an increasing interest among the spoken language research community in the collection and exploitation of large, computer-stored corpora. This has resulted largely from the new potential offered by technological developments in computing - particularly the vastly increased storage capacity, and rapid retrieval which are now possible. A number of large, computer-stored language corpora are now available - notably the British National Corpus, which is the outcome of the biggest language data collection exercise ever undertaken.

However, the collection and storage of vast amounts of data is only the first step in the development of a fully usable computational system for analysis. To make the data properly accessible to the research community, the material must be classified in some way - both in terms of overall discourse category, and in terms of the internal features which characterise that category. To this end, various systems of data tagging have been devised, in an attempt to perform a 'once and for all' basic analysis on the raw data, which will then be automatically available with the data to any researcher accessing the corpus. Much of the effort in this area has focussed on a narrow kind of grammatical tagging, so that individual lexical items in a corpus are stored along with their part of speech descriptions.

In parallel with this, and specific to spoken language corpora, systems are being developed to tag spoken items with phonological features, which clearly carry meaning in the speech event, but which have no, or only very inadequate, orthographical equivalent in a transcription. Like the grammatical tagging, however, this tends to operate at a very low level, on an item by item basis, so that variations in pitch, amplitude etc. which are heard on the taped material are mapped onto the orthographical transcription, and are then available for researchers along with the transcription.

To leave data tagging at this level is, however, to miss out on the full potential of large, computer-stored corpora. To make maximum use of such large amounts of easily available data it is

## PHONOLOGICAL SIGNALLING OF TOPIC MOVEMENT

important to push the functionality of the computer systems to their full extent, and to provide for study of the higher levels of language, i.e. those structures which operate over large sections of spoken text (equivalent, say, to the paragraph and above). This more general kind of analysis and data tagging can then be used as the basis for the development of further analysis on the structures which operate at the level of the whole discourse, and which characterise different discourse genres.

Ideally, the inclusion in a stored corpus of analytical material which is relevant to the overall structure of the discourse should be done as the corpus is built, so that future research which draws on the corpus does not involve the researcher in unnecessary repetition of the more basic kinds of analysis.

To map signals of large discourse structures onto textual transcriptions we cannot, of course, rely on intuition. Our evidence for the existence of a particular structure must rest on signals which are made explicit in the speech event, textually or otherwise. This means that substantial groundwork must be carried out in the initial stages, in order to identify these signals operating in real-world contexts. This present paper is a first step in that groundwork.

## 2. DISCOURSE STRUCTURES IN CONVERSATION

The aim of the research described here is to identify some of the phonological features of conversation which operate to signal large structures in the discourse as a whole. I have taken a primarily top-down approach, using as a starting point my earlier work on the lexical signalling of discourse structure (Cheepen 1988) and tying this in with observations of the more low-level features (e.g. turn-taking mechanisms, adjacency pairs etc.) identified by conversation analysts such as Schegloff, Sacks and others.

The main discourse components of conversation have been identified as *introduction*, *speech-in-action*, *story* and *closing* (Cheepen 1988, Cheepen & Monaghan 1990). The bulk of conversation consists of *story*. This should not be confused with what is commonly called *narrative* (following Malinowski's 1923 definition) - which is a monologic event, performed by a 'story teller' for an 'audience'. Although such narratives can and do occur in conversation, they are comparatively rare. The kind of story which characterises conversation is essentially a collaborative event, where both (all) participants contribute to the telling by guessing what comes next, asking questions to elicit further details, and (typically) contributing evaluations throughout the telling. Stretches of *story* are punctuated throughout conversational discourse by stretches of *speech-in-action* where the participants collaborate to discuss their immediate surroundings.

The major discursive feature of *speech-in-action* is *evaluation* where the speakers give their opinions of various items in the surroundings. *Evaluation* is also a major component of *story*, where it serves to express the speakers' opinion of the story in an overall sense. By this, I mean a form of evaluative summary of the preceding section of the discourse, which is typically shared by both participants, and serves to establish their joint opinion of the status of what has just been said.

*Story* contains other components - *state-event-state sequence*, *temporal location* and *participant specification*<sup>1</sup>, and these sections often also contain evaluative material of some kind, but it does not have the same status and function in the discourse as what I am calling *evaluation*. I will return to this point later in section 4.

### 3. LEXICAL SIGNALLING OF DISCOURSE COMPONENTS

The discourse components I have referred to above are all clearly signalled in the speech event, and in many cases the signalling is done by lexical items. In some cases, simple tagging in terms of lexical signals is sufficient for most research purposes, and this is particularly the case with the component *temporal location*, which is signalled by a small, fixed set of lexical items, e.g. "this morning", "last week", "on Thursday" etc. Identification of this component is made even simpler, because inspection of a variety of conversational transcriptions shows that this kind of occurrence is to be found *within 20 words* of the beginning of the story, and this makes it an ideal feature for tagging in corpora.

The other components of *story* are, however, slightly more problematic to characterise by purely lexical signals, though they do certainly have lexical (and grammatical) signals which can be identified in the transcriptions. (For a more detailed discussion on lexical and grammatical signalling, see Cheepen 1994.) The most problematic is *evaluation*, because although it is, in the vast majority of cases, lexically signalled, the number of possible lexical realisations is extremely large, and in practical terms this makes a purely lexical tagging of this component of *story* an impossible task. The identification of certain phonological features of evaluation is, however, a more promising candidate.

### 4. FUNCTIONS OF EVALUATION IN SPOKEN DISCOURSE

I have said earlier that evaluation can be thought of as an overall judgement of what was talked about in the preceding section of the discourse. The notion of overall judgement is important here, because evaluative material of many kinds proliferates throughout conversation (and indeed other discourse genres, both spoken and written). Evaluative, or judgemental material is present everywhere, and whatever is mentioned will usually be evaluated in some way e.g. "there are three good reasons..." "I saw a lovely cat yesterday" "Mr Kinnock was very unhappy about the results of the election". These are instances of some level of evaluation, but they are not what I am concerned with here. I am, rather, concerned with an essentially overall kind of evaluation, shared by co-conversationalists, which occurs primarily at two points in conversation, and which I have outlined in section 2 above.

Firstly there is the evaluation which occurs towards the end of a dialogic story-telling, which sums up the story-telling in terms of what the discourse participants feel about it, and which operates in part to draw a line under that part of the discourse so that the participants can move on to other

<sup>1</sup>These can be thought of as "what happened?", "when did it happen?" and "who was involved?"

things. Secondly, there is the evaluation which occurs (even more frequently) in the stretches of *speech-in-action* which punctuate conversation, and during which the speech participants take turns to comment on the 'here and now' of their immediate surroundings and to exchange and share opinions of those surroundings. While this kind of evaluation is extremely easy to recognise when looking at a transcription, it is exceptionally difficult to characterise in terms of its lexical realisations, because they are potentially so varied.

Because conversation is essentially a dialogic speech event, the phenomenon of frequent speaker change is clearly observable at all times throughout the discourse. It is clear from observation of a wide range of conversational transcripts that this change over from one speaker to another is a major mechanism for promoting the collaboration between participants which is necessary for the continuation of the discourse and the smooth progression from one topic to another. Where evaluation occurs, it is always the case that (in a two-party conversation) both participants will make some kind of contribution. Often this will involve both speakers producing an evaluation (in sequence, or sometimes overlapping), but sometimes the turn taking is observable as an 'invitation' by one speaker to the other, who is then expected ('obliged') to provide the evaluation. As the overall goal of conversation is to create and maintain social bonds between the participants, it is usually the case that where both speakers produce evaluations, they will be matching ones, showing agreement between the speakers. However, within this general context of agreement, there are sometimes non-problematic disagreements which do not disturb the general flow of the discourse<sup>2</sup>. I will return to these non-problematic disagreements in section 6 below.

### 5. PHONOLOGICAL SIGNALS OF EVALUATION

The lexical signalling of an evaluative consensus between speakers is typically of a loosely echoic nature, so that an utterance such as "that's dreadful" from one speaker, is likely to be followed by "yeh, dreadful" (or sometimes "yeh, awful") from the next speaker. This kind of overt matching by lexical tokens is also manifested phonologically, by a matching of the participants' speech patterns. This matching occurs over quite large sections of the discourse - certainly larger than the syllable, and more usually over whole utterances. This matching of speech patterns will be illustrated in the examples used in sections 5.2 and 5.3 below.

#### 5.1 Phonological features

The features I have considered in the analyses which follow are the same as those used by Crystal & Davy (1969). They include phenomena such as:

- pitch* - the categories here are 'high' or 'low', and 'ascend' or 'descend'
- loudness* - the categories here are 'fortissimo' (very loud), 'forte' (loud), 'piano' (quiet) and 'pianissimo' (very quiet)
- more gradual changes towards these extremes are 'crescendo' and 'diminuendo'

<sup>2</sup>For an account of more fundamental kinds of disagreement, which do have an impact on the flow of the discourse, see Cheepen and Monaghan 1990, Ch. 3.

# Proceedings of the Institute of Acoustics

## PHONOLOGICAL SIGNALLING OF TOPIC MOVEMENT

- speed* - the categories here are 'allegro' and 'allegro' (fast and faster)  
and 'lento' and 'lento' (slow and slower)  
- more gradual changes towards these extremes are 'accelerando'  
and 'rallentando'
- rhythm* - the categories here are 'staccato' (clipped) and 'legato' (drawled)

In all these cases, the identification of features must be seen as instances which are extreme for the speaker in question. The norm for any speaker is the range (in whatever category) which is readily observable throughout the bulk of the discourse in question. The examples used have their phonological features noted in square brackets, immediately below the relevant section of the dialogue.

### 5.2 Signals of evaluation in speech-in-action

#### [Example 1]

- A: I can't stay long I've only got six cigarettes  
B: oh that's dreadful  
[lento, piano, dimin. to pianiss., descend to very low, legato]  
(Cheepen & Monaghan 1990, Tape 'Xmas 83')

In this example, the evaluation clearly occurs at a point when the speakers are involved in talking about the here-and-now of their immediate surroundings. Note that the evaluations (all underlined) are characterised by slow, quiet speech, which is low-pitched and has a smooth (drawled) rhythm.

#### [Example 2]

- A: looks awfully nice  
[lento, descend to v. low, v. legato]  
B: looks better than the last time you saw it  
[lento, piano, descend to v. low, legato]  
A: true yes  
[lento, piano dimin. to pianiss., low, legato]  
(Cheepen & Monaghan 1990, Tape 'Celia')

Again, the same phonological features can be observed in this second example of *evaluation in speech-in-action*. Let us now consider some examples of *evaluation* as it occurs in *story*.

### 5.3 Signals of evaluation in story

#### [Example 3]

- A: I like the first one the first one was much better where you saw him  
stroking the fox  
[forte, accel.]  
and then you see him go in and pick up the raw things \*to  
eat\*  
[rall., staccato]

# Proceedings of the Institute of Acoustics

## PHONOLOGICAL SIGNALLING OF TOPIC MOVEMENT

- B: \*yeh\*  
[forte]  
A: and you kept getting kre . and then you see him cut the lemon and you  
(gasp) and he . that was  
[v. staccato]  
euhhhhh horrible (laugh)  
[fortiss., legato]  
that was really grotesque I thought  
[rall., dimin. to piano, lento, descend to v. low, legato]  
B: was a bit  
[pianiss., v. low, lento, legato]  
(Cheepen & Monaghan 1990, Tape 'Celia')

Note that the quiet, slow, low-pitched, drawled features of the evaluations correspond with those which occur in the evaluations in *speech-in-action* - see examples 1 and 2 above. These features do, of course, occur in other discoursal structures in conversation, and some even cluster together at other points. A clustering of, for instance, piano, low pitch, and lento, can be occasionally observed in other sub-sections of stories apart from evaluation, but this tends to happen where some kind of dramatic effect is intended, and instead of the drawled, legato rhythm which characterises evaluation, we find instead a noticeable staccato rhythm, usually delivered in a 'whispery' manner, quite unlike the 'dying away' delivery of evaluation.

### [Example 4]

- A: we had twenty cards this morning -  
[piano, whisper]  
one of them was from . Philip Anne Rebecca and Penelope  
[v. low, piano, lento, staccato, whisper]  
(Cheepen & Monaghan 1990, Tape 'Xmas 83')

The extra feature of staccato, combined with the whispery delivery, acts in this case to add emphasis to what is said, rather than to sum up by evaluation. Note that this example is taken from the beginning of a story, and includes the *story* components of *temporal location* - "this morning", *state-event-state sequence* - "we had twenty cards ..... one of them was from", and *participant specification* - "Philip Anne Rebecca and Penelope".

Example 3 illustrates two important features of dialogic story-telling. First, as the notation shows, the evaluation (underlined) is characterised by the same phonological features as occur in evaluations in *speech-in-action* - the speakers deliver their utterances in very low pitch, with very reduced amplitude, and slow down their normal speed of utterance, while eliminating possible pauses by drawling syllables and words. Second, as the extract shows, speakers use evaluative material frequently throughout their talk - note the opening utterance "I like the first one....". Clearly this is evaluative, in that it is giving an opinion, but as the rest of the extract shows it is not an 'evaluation' as I am discussing it here - a general summing up of the story. It is, in fact, operating more in the nature of a headline, preparing the way for the speaker to continue, and as

the notation shows, this is made clear by an acceleration in speed, which works to hold the floor for the speaker in an extended turn at talk.

### 6. TOPIC MANAGEMENT AND EVALUATION IN CONVERSATION

Example 3 is a good illustration of how speakers in conversation tend to link or embed stories in lengthy story-telling sessions. As stories are inevitably topic-based (that is, concerned with a particular conversational topic) it is usual to find that any particular story section has what can be thought of as a 'general topic', and that speakers move from sub-topic to sub-topic within this general topic area. Crow (1983) refers to the phenomenon of "topic shading", where one conversational topic 'leads naturally' to the next. Examination of a range of conversations shows that this is indeed the case, and furthermore that what speakers tend to do with these topics and sub-topics is to tell stories about them. Each sub-topic (embedded story) is structured (in terms of discourse patterning) in the same way as the overall *story* section. The beginning of the extract, with its referential items "I like the first one...", indicates very clearly that this is part of an ongoing general topic.

This repetition of the discourse patterning of large story sections in the smaller stories within those sections leads inevitably to a consideration of the purpose of such structuring, and to ask how this functions to progress the discourse as a whole. To answer this, we must remember the 'why' of conversational discourse - the overall aim of such speech events. Given that conversation (in the specialised sense of the term as used here) is essentially a non-goal based speech event (that is, it has no 'transactional' aim in terms of the world outside the encounter), its only purpose is the social bonding of the speakers.

The building of bonds between individuals is, in the main, a matter of sharing the same viewpoint about a range of matters (possibly transactional as well as social, or interactional). To a large extent this means that the individuals will agree with one another. When we look at a speech event such as conversation, which promotes (even demands) the verbal expression of social bonds, it is clear that structuring the discourse to provide regular and frequent opportunities for participants to deliver evaluations is an ideal way of allowing speakers to match up their opinions, and to make their agreement explicit. As I have pointed out, evaluative material occurs throughout conversation - at its lowest level (the adjective) it peppers nearly all the utterances in the discourse - and not all these shades of opinion will be shared by the conversational participants. However, the summing-up evaluation which occurs at the end of stories tends to be much more emphatic in that it is much more clearly signalled, firstly by both participants contributing their version of the evaluation, and secondly by the particular phonological features I have indicated earlier - low (to very low) pitch, reduced speed, reduced amplitude and very smooth, drawled rhythm.

Although the overwhelming tendency for speakers in conversation is to match their evaluations, there is also, as I indicated in section 4 above, an occasional non-problematic disagreement - or, more accurately, a 'surface disagreement' - which occurs in evaluation. These instances can be observed to be non-problematic, and purely surface disagreements, if we look at what happens

# Proceedings of the Institute of Acoustics

## PHONOLOGICAL SIGNALLING OF TOPIC MOVEMENT

next in the discourse<sup>3</sup>.

[Example 5]

A: our . previous next door neighbours down in Sutton Road . have just sent us a card  
by post  
[forte, lento, low, strong stress]

B: mm -  
[pianiss., v. low, legato]

A: by post -  
[forte, lento, low, strong stress]

B: well people do .  
[descend to v. low, pianiss., v. legato]  
I've done it myself accidentally you know you put them \*in piles\*  
[higher pitch, cresc., accel., staccato]

A: \*yes\* yes that's right it can be  
[accel., higher pitch]  
accidentally done can't it  
[descend., rall., legato]

(Cheepen & Monaghan 1990, Tape 'Xmas 83')

This extract illustrates the surface nature of the disagreement which occurs. Speaker A is clearly expecting a very different kind of evaluation from Speaker B, when she first says "by post", and even tries again for a negative evaluation. Both of Speaker B's evaluations are, however, positive rather than negative - though note that the phonological features of these utterances ("mm" and "well people do") correspond exactly to those found in cases of evaluation where the speakers are clearly in agreement - i.e. low pitch, slow, smooth drawl. As the extract progresses it is clear that the disagreement is not going to cause any kind of trouble between the speakers. What happens is that Speaker A changes her expectation of the appropriate evaluation from negative to positive, and, at the end of the extract, delivers an evaluation ("yes that's right it can be accidentally done can't it"), which matches the opinion of Speaker B, and finally echoes precisely the phonological patterning of Speaker B's original evaluation.

## 7. CONCLUSION

As the preceding examples and discussion have shown, *evaluation* is a very important, and frequently occurring component of conversational discourse. Because it allows the conversational participants to 'show their colours' it functions to promote the overall aim of conversation, which is to reveal personal opinions with a view to building bonds between speakers by sharing those opinions. The analysis I have used in the extracts shows that speakers expend considerable effort on signalling these important points in conversation, both by their lexical choices, and by the clustering of phonological features which they associate with those choices.

<sup>3</sup>For a description of the phonological features which characterise disagreements which do become problematic in the discourse, see Cheepen 'Phonological features of troubles and repairs in conversation' (forthcoming).



## PHONOLOGICAL SIGNALLING OF TOPIC MOVEMENT

The interpersonal goal of using *evaluation* in conversation is to achieve speaker solidarity through agreement, and although it is always a possibility that a conversation will run into serious trouble<sup>4</sup>, this is very rarely the case. As the above extracts show, even an occasional mismatch in evaluation can be (and usually is) collaboratively 'wound in' to the ongoing discourse so that the discursive drift, in terms of interpersonal agreement is constantly maintained. While this quality of 'not going to be a problem' is observable (eventually) in the lexical realisations used over a sequence of utterances, it is, even at its onset, clearly evident in the phonological signalling used by the speakers, because the clustered features of non-problematic disagreement echo precisely those of the more generally occurring speaker agreement<sup>5</sup>. Given the regularity of the phonological signals which characterise *evaluation* in conversation, which are constant in both *speech-in-action* and in *story*, and which therefore form such a large part of conversational discourse, my conclusion is that this clustering of phonological features is a most important signal of discourse structure. As such it is essential to the study and analysis of the whole genre of conversational discourse, and should, therefore, be built into the corpora of computationally-stored conversational material at present being amassed.

### 8. FUTURE WORK

Having identified evaluation signalling (phonological and lexical) as a component of conversational discourse which is suitable for tagging in stored corpora, it is now necessary to devise some form of tagging notation which is appropriate for this level of analysis. Future work will focus on the development of a suitable tagging system for phonological and lexical signals of evaluation, in parallel with a detailed analysis of other components of conversational discourse - such as temporal location and participant specification, which are sub-structures within conversational stories - and which may, if their signalling is similarly regular, also be potential candidates for inclusion in such a tagging system.

### 9. BIBLIOGRAPHY

- Cheepen, C., 1988, The predictability of informal conversation, Pinter Publishers Ltd.  
Cheepen, C. & J. Monaghan, 1990, Spoken English: A practical guide, Pinter Publishers Ltd.  
Crow, B., 1983, in Craig & Tracey (eds), 1983, Conversational coherence, Sage  
Crystal, D., & D. Davy, 1969, Investigating English Style, Longman  
Malinowski, B., 1923, *The problem of meaning in primitive languages*, in Ogden, C.K., & I.A. Richards (eds), 1923, The meaning of meaning, Routledge & Kegan Paul  
Sacks, H., E. Schegloff & G. Jefferson, 1974, *A simplest systematics for the organisation of turn taking for conversation*, in Schenkein, J., (ed), 1978, Studies in the organisation of conversational interaction, New York, Academic Press

<sup>4</sup>It would in that case turn into some other kind of discourse genre - perhaps a row - or the communication may break down entirely.

<sup>5</sup>It may be that this is a case where what Malinowski (1923) calls "the bonds of antipathy" are in operation. Malinowski's observation was in the context of his study of "phatic communion", but this phenomenon corresponds very closely to what I am calling "conversation".

