PROSODY-PRAGMATICS INTERFACE IN THE SITCOM DISCOURSE

Magdaléna Bilá, Alena Kačmárová

University of Prešov in Prešov
ul. 17. novembra 1, Prešov, Slovakia, 08078

The paper presents the outcomes of the analysis of the tonic stress placement and its interpretation in terms of a speaker's communicative intention in the original version of the episode of the sitcom 'Friends'. Phonetic and pragmatic analyses of an utterance represent a combination of langue and parole phenomena and that of quantitative and qualitative approaches. The research intention is to explore quasi-spontaneous conversational speech of American variety of English, and to establish in what manner the occurrence of tonic stress is reflected in pragmatic interpretation of the selected audio-visual text. The findings show that the genre of pre-scripted text does not necessarily inhibit and/or determine the mere vocal realization and pragmatic reading.

Key words: prosody, tonic stress placement, speaker meaning, communicative intention, illocutionary force

1. INTRODUCTION

Austin's beneficiary input in perceiving language was the idea that whatever is uttered at both carries a particular meaning and represents a particular activity. This is to say that an utterance conveys a certain force which will subsequently result in a particular activity. Leech [7] calls attention to the fact that it is advisable to differentiate between an illocutionary force and an illocutionary act. He [7] understands an illocutionary force as a communication plan/aim implied in a speaker's utterance; his understanding of an illocutionary act is such that it is the fulfilment of that particular communication aim. An illocutionary force can be carried out by means of specific language means, though. A certain discourse type needs language means that communicate its functions.

Our focus is audio-visual text, a film dialogue reflecting a conversational routine of a particular contextually embedded language and is both entertaining and attractive for viewers. Audio-visual text as such displays concurrent employment of two communication channels i. e. simultaneous transmission of two sets of signs (verbal and non-verbal). It mirrors linguistic behaviour in a particular communication act and the roles we take up in the social setting. It is created with ambition to give the illusion of natural/idiomatic linguistic and non-linguistic behaviour. This is enhanced by the fact that multiple authorship is at play here. Such collaborative work of a whole team may have a potential to add to the naturalness and credibility of film dialogue (*cf* [9]), which is why a common viewer may well perceive a film dialogue as spontaneous speech. In actuality, this is pre-scripted text acted out by actors (on term pre-scripted text, *cf* [11]).

The objective of our research is to arrive at how actual speakers/actors interpret pre-scripted text in terms of vocal realization and how they interpret the intentionality of the text (on term intentionality, *cf* [11]). The research question is what the pre-

scripted text means to its speakers. The research plan consists of three steps: 1/ the identification of emphatic stress placement in the target language material, 2/ the identification of communication functions in the target language material, 3/ the interrelation of emphatic stress placement and communication functions. Hence our discussion focusses on prosody-pragmatics interface in pre-scripted text.

The discourse selected for the contemplation is the conversational interaction in the sitcom *Friends*, specifically in the episode *The One When No One Proposes* (Episode 1, Season IX, 2002/03). The synopsis of the episode is as follows: Rachel is in the post-delivery room; she just gave birth to baby Emma. In the room, Joey looks for some tissues for upset Rachel, picks up Ross's jacket, and an engagement ring falls to the floor. He kneels to pick it up and turns to Rachel, still on his knees and still holding the ring. Rachel thinks this is his proposal of marriage and accepts. In the rest of the episode, Joey strives to explain the misunderstanding (Friends, online document). Such language material represents present-day language currently in use by a native speaker of English that is idiomatically rich, dialogically structured with multiple-speaker styles.

2. DEFINING KEY CONCEPTS

It is axiomatic that a speech signal is a continuum. Therefore, as addressees we need to be able to spot points of importance in the continuous stream of speech. As addressers, we must highlight information units in our messages so that addressees do not have difficulty interpreting what they hear, deciding how they relate to what has just been said, or anticipating what the addresser is possibly leading up to [5]. "In natural speech people use a variety of prosodic means to convey to their interlocutor which elements of the utterance are especially important" [8].

In speech, "... emphasis ... may be signaled in many ways including use of special stress and intonation patterns, choice of words, choice of grammatical patterns, etc." [1; 408]. Grammatical and lexical markers of emphasis do not operate by themselves; quite on the contrary, such effect is achieved only if they are complemented with prosodic means. As long as we refer to different degrees of a certain quality, express preference, stress the importance, invite or express surprise, etc. with greater insistence than that found in neutral patterns, we make a particular unit of information more prominent. What seems to matter in identifying a particular unit as prominent is a greater acoustic intensity of the word or syllable pronounced, higher pitch of voice as well as prolonged sound duration. The three components of speech manifested by stress, pitch and sound duration, seem to correspond to three modulations of voice: force modulation, tone modulation and temporal modulation respectively. Altogether they make the essence of intonation. Tench [10; 1] points out that "...we have all made an observation like 'It is not what they said, but the way they said it!'" Tench [10; 56—59] calls attention to the fact that a speaker can make choice between neutral (i.e. unmarked) tonicity and marked form in the tonicity system. With reference to Halliday [In: 10; 8], in a neutral form the tonic syllable tends to come last, since the normal structure of the clause presents the given information first and the new last; if,

for whatever reason, the tonic accompanies a non-final lexical item, marked tonicity is the case.

Subconscious use of the above prosodic features is in fact the realization of neutral patterns of intonation; hence, it typifies the neutral tonicity. The features of prosody that operate in such situations are stress, namely word stress, and pitch, specifically organic pitch-range. Stress in its linguistic use is a property of a syllable and as such has a distinctive function in that it places stressed and unstressed syllables in opposition. Neutral tonicity is characterized by linguistic pitch-range, "...which is the range within which the phonologically relevant pitch of the speaker's voice habitually varies in paralinguistically unmarked, attitudinally neutral conversation" [6; 457].

Conscious use of the above prosodic features adds to the marked tonicity, especially with words or in structures that would not otherwise be affected. By and large such linguistic behaviour envisages an emotive communicative situation and anticipates the fact that there is a particular reason for the greater intensity. Attitudinally marked situations require emphatic stress and paralinguistic pitch-range if the desirable effect is to be achieved. The function that emphatic stress fulfils is "... to call the listener's attention to a given syllable or word with greater insistence than is afforded merely by neutral patterns of intonation or lexical stress" [6; 515]. Paralinguistic pitch-range is in fact the adjustment of organic pitch-range for the purposes of paralinguistic communication. It is "... the range of pitch that is exploited for momentary paralinguistic purposes of signaling particular attitudinal information..." [6; 457]. Consequently, emphatic stress and paralinguistic pitch-range are characteristics of an information unit chosen by a speaker for the purposes of a particular communicative situation.

From what has been said it follows that intonation has both linguistic and paralinguistic dimensions. In Tench's words [10; 2], the linguistic dimension "...concerns the message itself: how many pieces of information there are, what information is new"... while the paralinguistic dimension "...concerns the messenger rather than the message: the speaker's state of mind, their degree of politeness and their effort to associate or dissociate from you". As a spoken interaction is a rich source of attitudinally marked, hence emotionally marked utterances, it provides many an example of paralinguistic communication that reflects the actual prominence of a particular information unit over other information units. In speech, the burden of reinforcement is carried by consciously assigning the prominence to a particular information unit through a paralinguistic dimension of intonation.

Each communication act conveys a communication intention, in other words a speaker's intention. This covers intentions such as asserting something, inquiring about something, issuing a command, or evoking emotions. A speaker's intention is present whenever we communicate. Its basic components are the reason for uttering something, the topic of the utterance, and the form of the utterance. A communication intention can serve as a means for specifying a communication function of an utterance.

R. Jakobson argues that every oral or written verbal message or 'speech act' (parole) has the following elements in common: the message itself, an addresser, an addressee, a context (social and historical context in which the utterance is made), a contact (physical channel and psychological connection that obtains between addresser and addressee), and a code, common to both addresser and addressee which permits communication to occur[1].

These six elements of human interaction bear particular language functions. Jakobson [In 2] further provides an arrangement of the six elements and makes them parallel with language functions (see Figure 1).

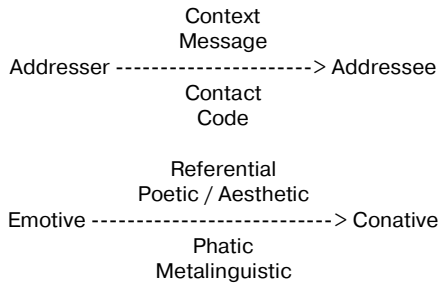

Fig. 1. Elements of human communication and parallel language functions

The above scheme can be interpreted as follows:

1) In the utterance, the Addresser communicates a particular emotion (emotive function).

2) The Addressee decodes a message and responds accordingly (conative function).

3) The Context assists in decoding a message through making available physical environment or social and cultural background (referential function).

4) An utterance/Message has a particular aesthetic value (aesthetic function).

5) Communication serves as a means of building and maintaining Contact (phatic function).

6) Communication Code also serves for discussion on the Code itself (metalinguistic function).

The language material analysed made available referential, emotive, conative and phatic functions. The utterances were assigned these functions based on the following understanding.

A referential communication function is fulfilled in either of the two following cases: a) requesting information and b) providing information. The key criterion is the denotative nature of the utterance in terms of a neutral comment on the state of affairs and subjective evaluation of the truth value of the statement free of emotive involvement.

An emotive communication function is expressed through: a) an addresser's need for emphatic focus conveyed not only prosodically but also through lexis or syntax, or b) an addresser's relationship to the utterance, and his/her way of expressing emotions in terms of marking a high level of the truth value of an utterance, strengthening the

[1] Quoted from Clarke, online document

intensity of a feature, mapping an emotional load on the propositional content, or expressing enthusiasm, involvement, stressing the importance of a piece of information from a speaker's point of view.

A conative communication function is fulfilled when: a) an addresser requires an addressee to perform an activity, b) an addresser causes an addressee to do something, or c) an addresser calls to joint activity.

A phatic communication function is fulfilled when the speaker aims to establish, maintain, or terminate contact with another speaker. Phatic and emotive function is achieved when a greeting fulfils also other than a mere phatic function, i.e. is used to express interest and/or being part of in-group, commonality of purpose.

In our analysis, the key notion is an utterance. Our understanding of an utterance is as follows. An utterance is a part of a turn. While a turn is a content-free element delimited by interactional exchange, an utterance is a content-filled unit delimited by pragmatic meaning and prosodic features. In the target corpus, turns were segmented into utterances based on their meaning, a speaker's communication intention, and perceptual identification of inter-sentential pauses. For example — the initial dialogue between Rachel and Joey has two interlocutors, which is to say two turns. The first turn consists of 4 utterances; the second turn consists of 1 utterance. One turn can have different communication intentions, in other words different communication functions. The communication functions can be assigned as follows (see Figure 2 below):

Rachel: Joey? Oh, my God. Okay. So… I guess we should make it official, huh?
Joey: Look, Rach, I….

	Communication intention	Referential function	Emotive function	Conative function
Joey?	startlement			
Oh, my God.	surprise/ startlement			
Okay.	Expressing agreement			
So, I guess we should make it official, huh?	Appeal to action			
Look, Rach, I….	Effort to provide explanation			

Fig. 2. Example of assigning communication functions to the utterance

3. RESEARCH FINDINGS

The analysis consisted of three steps. Firstly, we perceptually identified emphatic stress placement and its type. Secondly, we assigned communication functions to the utterances. Thirdly, we studied the extent of the interlinkage between prosodic features and communication functions of utterances in the chosen pre-scripted text. In the original American version of the target episode we identified 306 turns and 482 utterances.

In the analysis on the selected suprasegmental, the starting point was Gussenhoven's [4] understanding of the size of the emphatic stress constituent (*cf* focus constituent, Gussenhoven's term) derived from the location of emphatic stress, and the type

of that derived from the distribution within the emphatic stress constituent. The analysis unveiled two types of utterances in terms of emphatic stress placement: neutral and non-neutral (see Figures 3 and 4). Utterances with non-neutral emphatic stress were further discussed in terms of size and distribution of emphatic stress; the former was specified as broad vs. narrow, the latter as marked vs. unmarked emphatic stress placement (see Figures 5 and 6).

Total number of utterances	Neutral utterances	Non-neutral utterances
482	256 (number)	226 (number)
	53.11%	46.89%

Fig. 3. The occurrence of emphatic stress: neutral vs. non-neutral utterances

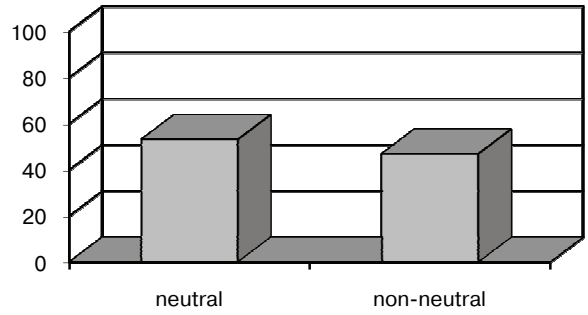

Fig. 4. The occurrence of emphatic stress: neutral vs. non-neutral utterances

Broad emphatic stress	Marked emphatic stress placement	Combined
65	87	74
28.76 %	38.50 %	32.74 %

Fig. 5. The occurrence of types of emphatic stress

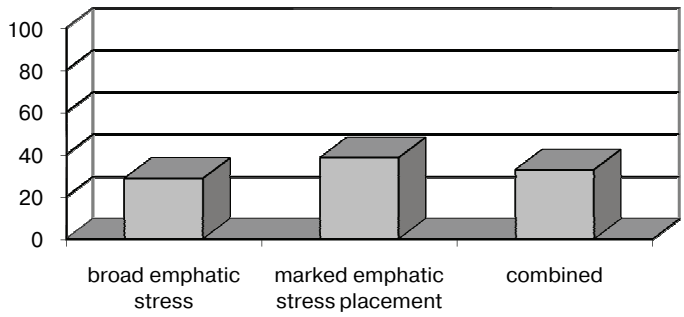

Fig. 6. The occurrence of types of emphatic stress

In the pragmatics analysis, the identification of communication functions was done based on pragmatic meaning, situational and linguistic context of utterances. The analysis of the target corpus yielded the following results (see Figures 7 and 8):

Referential function	Emotive function	Conative function	Phatic function	Phatic + emotive	Total of
272	140	52	9	9	482
56.43 %	29.04 %	10.79 %	1.87 %	1.87 %	100%

Fig. 7. The occurrence of communication functions

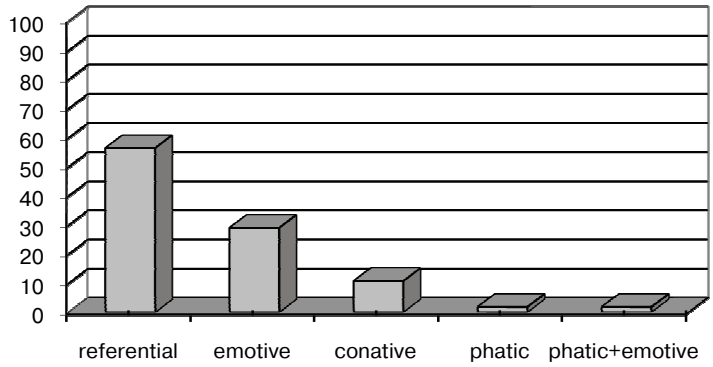

Fig. 8. The occurrence of communication functions

The observation in terms of prosody-pragmatics interface provided the following outcomes. Emphatic stress was spotted in utterances with referential, emotive, and conative communication functions. The occurrence of broad emphatic stress was relatively balanced. The biggest difference was observed between utterances with referential and emotive communication functions on the one hand and those with a conative communication function on the other (see Figure 9).

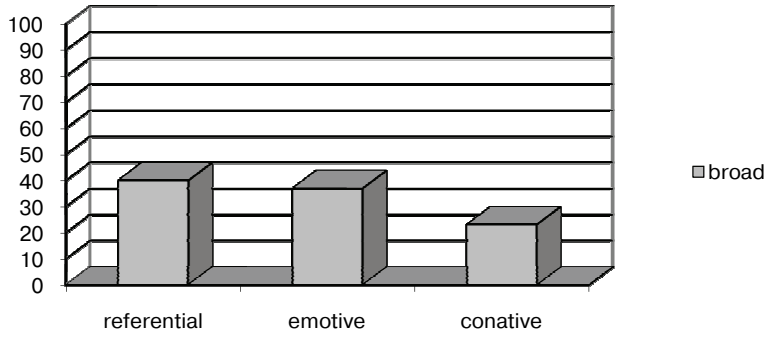

Fig. 9. Broad emphatic stress in utterances with R, E and C functions

The occurrence of marked emphatic stress was balanced between utterances with emotive and conative communication functions. They differ significantly from utterances with a referential communication function in that the occurrence of the two was expressively low. This is illustrated in Figure 10. In total, the occurrence of marked emphatic stress was low. The same applies to the occurrence of the combined type of emphatic stress, which is evident from Figure 11.

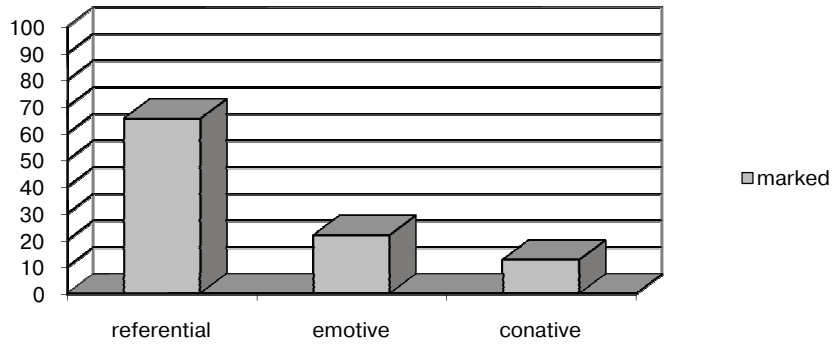

Fig. 10. Marked emphatic stress placement in utterances with R, E, and C functions

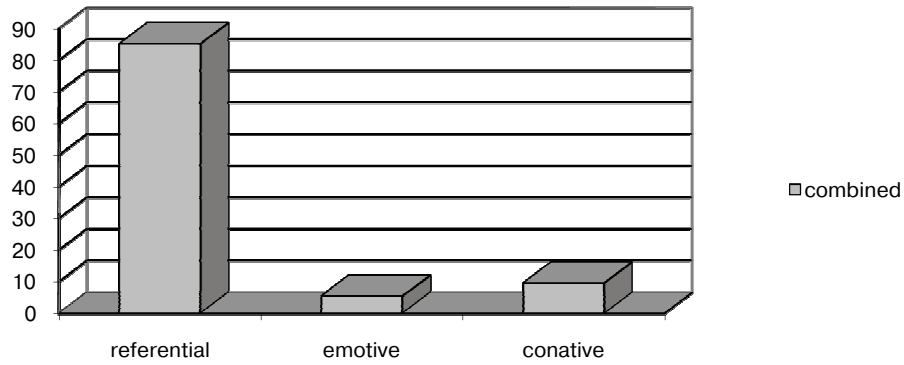

Fig. 11. The combined type of emphatic stress in utterances with R, E, and C functions

Non-neutral emphatic stress appeared in almost one half of the utterances in the analysed corpus. It was most frequent in the utterances with a referential communication function, approximately one third of all utterances (possibly due to the diversity of communication intentions within such a function). It was much less frequent in the utterances with emotive and conative communication functions. Speaking of the types of emphatic stress, the opposition of emotive and conative communication functions versus referential communication function is the case again. In the former, the occurrence of the combined type of emphatic stress outnumbers the other two types (i.e. broad emphatic stress, marked emphatic stress placement occurring by themselves); in the latter, the reverse is true (see Figure 12).

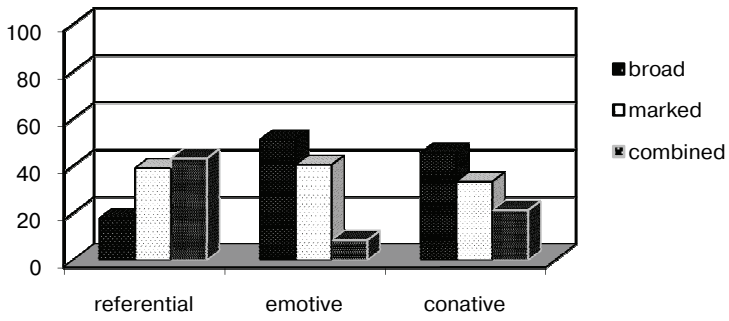

Fig. 12. Emphatic stress types in utterances with R, E, and C functions

4. CONCLUSIONS AND IMPLICATIONS

The objective of our research was to arrive at how actors deal with the text that is primarily written, only secondarily interpreted through actual acting out and through recognizing the intentionality of the text. The discourse selected for the contemplation was conversational interaction in the sitcom *Friends*, specifically in the episode *The One When No One Proposes* (Episode 1, Season IX, 2002/03). The research question was what the pre-scripted text means to its speakers.

Since it is hard to judge the author's intention we aimed to recognize the intentionality of the text through the actors' performance on vocal and pragmatic levels. In the studied language material, we identified 306 turns and 482 utterances that were further explored. During the analysis on the selected suprasegmental we observed 53.11% neutral utterances and 46.89 % non-neutral utterances, i.e. almost fifty-fifty ratio. In terms of types of emphatic stress in non-neutral utterances, in 28.76% of utterances broad emphatic stress was present, in 38.50 % utterances marked emphatic stress placement was noticed, and in 32.74 % utterances combined type of emphatic stress was used. In the target source, utterances with 5 communication functions occurred — referential, emotive, conative, phatic, and phatic-emotive. The interrelation of the two analyses reveals that the observed prosodic trait was present in utterances with three communication functions — referential, emotive, and conative.

The outcomes point to the fact that the actors interpret pre-scripted text with the aim to create the illusion of natural, spontaneous speech. In almost half of the utterances the pre-scripted text was performed with a prominence marker. The logical inference is that the pre-scripted text does not necessarily inhibit and/or determine the mere vocal realization and pragmatic reading. As it were, this is how the illusion of a spontaneous verbal product may as well be generated.

REFERENCES

[1] *Celce-Murcia, M. and Larsen-Freeman, D.* The Grammar Book — An ESL/EFL Teacher's Course, Heinle & Heinle Publishers, 1983.

[2] *Clarke, R.* Jakobson, R. „Linguistics and Poetics", 2005, [cit. 10-09-23]. Available at: http://www.rlwclarke.net/Courses/LITS3304/2004-2005/05AJakobsonLinguisticsAndPoetics.pdf.

[3] Friends. Online document. [cit. 11-09-15]. Available at: http://seriarecenze.webnode.sk/friends.

[4] *Gussenhoven, C. Types of focus in English* [cit. 13-04-15]. Available at: <www.ru.nl/publish/pages/516003/matt.pdf.
[5] *Kenworthy. J.* Teaching English Pronunciation. Longman Publishing Group, 1989.
[6] *Laver, J.* Principles of Phonetics, Cambridge: Cambridge University Press, 1994.
[7] *Leech, G.* Principles of Pragmatics. Longman Group Ltd., 1983.
[8] *Nenkova, A. and Jurafsky, D.* Automatic Detection of Contrastive Elements in Spontaneous Speech. [cit. 13-04-15]. Available at: www.cis.upenn.edu/~nenkova/papers/asru07.pdf.
[9] *Romero-Fresco, P.* A Corpus-Based Study on the Naturalness of the Spanish Dubbing Language: the Analysis of Discourse Markers in the Dubbed Translation of Friends, 2009, [cit. 11-09-15]. Available at: http://www.ros.hw.ac.uk/handle/10399/2237.
[10] *Tench, P.* The Intonation Systems of English, Cassell, 1996.
[11] *Widdowson, H.G.* Text, context, pretext: critical issues in discourse analysis. Oxford: OUP, 2004.

Note:

The research has been conducted within the authors' research activities in the project "Retrofitting and Extension of the Center of Excellence for Linguaculturology, Translation and Interpreting" supported by the Research & Development Operational Programme funded by the ERDF. Project code: 26220120044, Agreement Number: 031/2009/2.1/OPVaV.