Topic-Focus Articulation in PDT: Prosodic Characteristics of Contrastive Topic

Kateřina Veselá, Nino Peterek and Eva Hajičová

Abstract

This contribution (which was a basis for the poster Some Observations on Contrastive Topic in Czech Spontaneous Speech presented by the authors at CIL 17; the abstract of the poster is published in the volume of Abstracts of CIL17, Prague 2003, p. 361) is a preliminary, draft version of a more detailed and modified paper prepared for submission for Interspeech ICLSP 2004. We are most grateful to Julia Hirschberg for her detailed comments on the first draft of this paper, which have helped considerably to the readability of this contribution; some of her substantial remarks and suggestions will be embodied in the final version.

0. Introduction

The objective of this report is to present results of our investigations that attempt to bridge and relate some relevant outputs of two of the projects of the Center for Computational Linguistics at the Faculty of Mathematics and Physics, Charles University in Prague, namely the project of deep (underlying) level annotation of a Czech corpus (Prague Dependency Treebank) and the project in speech recognition. In particular, the meeting point of the two projects we discuss in our paper is the signaling of some aspects of topic-focus articulation (information structure, henceforth TFA) in speech by means of prosody.

The main motivation for our study is the effort to find out, how the investigation of the intonation contour of Czech sentence and of other characteristics of spoken language may help to look for criteria relevant for distinguishing some of the TFA values (Topic-Focus Articulation, or, in other words, information structure) in the annotation framework of the Prague Dependency Treebank (PDT; for a brief characterization of PDT, see below Sect. 1). In addition to this primary aim, a comparison of this type may lead to some other interesting observations, which may uncover some important features of signalization of meaning by means of suprasegmental means of expression in spoken language in general.

Our approach, which can be characterized as 'phonological' (Daneš 1957, p. 6), starts from the interpretation of meaning, which is only then a guide for the classification and use of the observable phonetic data. We are concerned with the relationships of contextual boundness in utterances (see below, Section 1.3); therefore the interpretation we base our analysis on uses methods of sentential and suprasentential syntax. We are aware that the categories that are an outcome of two layers of description, namely the syntactic and phonetic one, not always overlap and are not always relatable in a simple way. On the other hand, such an approach opens a space for an independent view on the methods of both descriptions from the point of view of their most relevant use in the analysis of sentential and textual relations.

We start by a short overview of our approach to the understanding and description of topic – focus articulation of the sentence (Sect. 1.), passing over to a description of the data we have used for our analysis (Sect. 2.). Sect. 3. then gives an overview of our processing and analyzing of the data, accompanied by a more detailed classification of the topic-focus phenomena. The results of the analysis are presented in Sect. 4. and 5.

1. The Annotation Scenario of the Prague Dependency Treebank

1.1 The Prague Dependency Treebank

The Prague Dependency Treebank (PDT in the sequel) is an annotated corpus of Czech texts (the texts are taken from the Czech National Corpus, the first release of which contains a hundred million word occurrences in journalistic fiction and other texts). At present, its scenario contains three layers of annotation: the morpho-syntactic layer (about 1100 tags are in actual use), the layer of dependency structures representing the surface shape of sentences (the so-called analytic layer) and the underlying syntactic layer (so-called tectogrammatical tree structures, TGTSs).

1.2 The tectogrammatical layer of annotation

The TGTSs are based on the theoretical framework of the functional generative description (FGD) proposed by Petr Sgall in the sixties and developed further by him and his followers until the present time (see esp. Sgall et al. 1986). The tectogrammatical level of FGD is distinguished from other theoretical descriptive frameworks esp. by the following two features: (i) it is dependency based, with coordination as a 'third' dimension of the trees, and (ii) the information structure of the sentence, its topic-focus articulation (TFA), is claimed to be an integral part of the tectogrammatical level, i.e. a distinction belonging to the underlying level of linguistic description.

Starting from these assumptions, the TGTS can be characterized as follows:

(a) only autosemantic words are represented as a separate node; the correlates of function words are just indices of lexical labels

(b) in case of (surface) deletions, nodes are restored in TGTSs,

(c) the condition of projectivity is met (i.e. no crossing of edges is allowed),

(d) types of dependency (tectogrammatical functions, 'functors') such as Actor/Bearer, Patient, Addressee, Origin, Effect, different kinds of Circumstantials are assigned,

(e) basic features of TFA are introduced.; for this purpose, a specific attribute has been established for every node of the TGTS, which may take on one of the following three values (see below in Sect. 1.3. for the primary opposition of contextual boundness):

T for 'contextually bound' (prototypically in Topic),

C for 'contrastive (part of) Topic',

F for 'non-bound', (typically in Focus)

1.3. Topic-focus articulation in a theoretical description

The articulation of the sentence into T(opic) and F(ocus) is based on what is from a *cognitive* point of view understood as the "given-new" strategy; the *semantic* basis of this articulation is the relation of aboutness (for a formal; treatment, see Peregrin 1994): a prototypical declarative sentence asserts that its Focus holds (or does not hold) about its Topic: F(T) or non-F(T).

The study of TFA (or, in more general terms, of the information structure) belongs nowadays to the most topical issues of research in theoretical linguistics. The views vary, but it has been widely accepted that though a bipartition of the sentence into topic and focus (and whatever terms the particular theories

may apply) is basic, a more detailed differentiation within the two parts (which is preferable to distinguishing more than a single basic dichotomy) should be introduced.

In the theoretical framework of FGD, within both T and F, an opposition of **contextually bound** (*b*) and **non-bound** (*nb*) nodes (as two primitive notions) is distinguished. This opposition is understood as a grammatically patterned dichotomy, rather than in the literal sense of the term; an *nb* element may be 'known' in a cognitive sense (from the context or on the basis of background knowledge) but structured as non-bound, new, in Focus; see (1) contrasted with (2). In the relevant sentences of both examples the capitals denote the intonation center, if read aloud (see Hajičová 2003 on contextual boundness and discourse patterns).

(1) She had separated from her first boyfriend with no great pain. With the second it was worse. … She LOVED him, and he was …
(Kundera 102)

(2) From the moment she ran into Josef at the Paris airport, she's been thinking of nothing but HIM. … In the bar, he was older and more interesting than the others, funny and seductive, and he paid attention only to HER.
(Kundera 98)

Both in (1) and (2), the pronouns refer to cognitively 'known' persons, i.e. known from the previous context; however, only in (1) the third sentence is structured in such a way that these elements are contextually bound (i.e. 'spoken about'), while in (2) both sentences are structured in such a way that 'him' and 'her' are contextually non-bound, they belong to the focus of the given sentences, as documented by placing the pitch on them if the segment is read aloud; this interpretation is strengthened by the fact that both pronouns are introduced by a focusing particle (*nothing but* in the first sentences and *only* in the second).

The notions of **topic** and **focus** (and the scopes of the focusing operators) can be derived from the distribution of contextual boundness with the individual nodes of the structural (dependency-based) tree and from the underlying order of the nodes of the tree. Examples (3) and (4), taken from the PDT, illustrate the topic/focus bipartition; the presupposed context in which the sentences with the given bipartition and cb/nb assignment are supposed to be uttered, is represented here by the question. The sentences are supposed to be pronounced with a non-marked position of the intonation center, i.e. with its placement at the end of the sentence; for the reasons of perspicuity of the examples, we denote (here and in the sequel) the position of the intonation center by capitals, be it in a unmarked, final position (as in (3)) or in a marked one (as in (5)). In the schematic representations of the sentences (numbered here with primes) the index *b* denotes the given element as contextually bound, elements with no index are considered to be contextually non-bound. The focus part of the sentences is printed in italics. Function words such as prepositions and auxiliary verbs do not have a node of their own on the underlying level assumed for our formal framework, and thus in our schematic notation they are included in brackets.

(3) V noci ze soboty na neděli skončil ve vojenském prostoru Ralsko sjezd MAJORU.
Lit.: At night from Saturday to Sunday ended in military area Ralsko meeting-Nom. of-majors
(Question: What happened during the night from Saturday to Sunday?)
topic: v noci. ze soboty na neděli
focus: skončil ve vojenském prostoru Ralsko sjezd majorů

(3') (In) noci-b (ze) soboty-b (na) neděli-b skončil (ve) vojenském prostoru Ralsko sjezd MAJORŮ.

(4) Rota nováčků při příležitosti sjezdu vyčistila vojenský prostor OD MUNICE.
Lit.: Squad of-novices at occasion of-meeting cleared military area from munition.
 (Question: What did the squad do at the occasion of the meeting in the military area?)
topic: rota nováčků při příležitosti sjezdu vojenský prostor
focus: vyčistila od munice

(4') Rota-b nováčků (při) příležitosti-b sjezdu-b *vyčistila vojenský-b proctor-b (od)MUNICE.*

A detailed empirical study of more complex sentences has led us to an introduction of the notion of **contrastive topic**, i.e. a contextually bound item that stands in contrast. Hajičová, Partee and Sgall (1998, p. 151) introduce the notion of contrastive (part of) topic in connection with the occurrences of the so-called focusing particles (focalizers) in topic, see (5); they use the index *c* to mark an element in such a position.

(5) (Who criticized even MOTHER TERESA as a tool of the capitalists?) JOHN criticized even Mother Theresa as a tool of the capitalists..

(5') *JOHN* criticized-b even-b Mother-c Theresa-c (as a) tool-b (of the) capitalists-b..

As our empirical analysis of the sentences in the Czech National Corpus has shown, the notion of contrastive topic should not be restricted to cases with focalizers in topic, see (6) and compare it with (7).

(6) (Preceding context: Kde se mluví česky?) Česky se mluví v ČESKU.
(Where is Czech spoken?) Czech is spoken in CZECHIA.

(6') Česky-b (se) mluví-b (v) *ČESKU.*

(7) (Preceding context: Mluví se česky v Česku nebo na Slovensku?) Česky se mluví v ČESKU, na Slovensku se mluví SLOVENSKY.
(Is Czech spoken in Czechia or in Slovakia?) Czech is spoken in CZECHIA, while in Slovakia, one speaks SLOVAK.

(7') Česky-c (se) mluví-b v *ČESKU*, na Slovensku-c se mluví-b SLOVENSKY.

To substantiate an introduction of some opposition in a linguistic description, one has to look for an operational test. Following up Koktová's (1999) observation that short forms of pronouns in Czech cannot be used in certain positions in topic, we use a criterion based on the opposition of strong and weak (short) personal pronoun. (Note: In Czech, strong personal pronouns differing in form from the weak pronouns exist for genitive, dative and accusative of the second person singular: *tě* vs. *tebe*, *ti* vs.*tobě*, *tě* vs. *tebe*, respectively, of third person singular masculine animate and neuter: *ho* vs. *jeho*, *mu* vs. *jemu*, *ho* vs. *jeho*, respectively, and of the reflexive personal pronouns *se* vs. *sebe*, *si* vs. *sobě*). This criterion may be well illustrated by (8) in comparison with (5) above: if *Mother Theresa* is replaced by a masculine noun in the question (*Mirka Dušína*), then a strong form of pronoun should be used in the answer.

(8) (Kdo kritizoval i Mirka Dušína jako nástroj kapitalismu?) HONZA kritizoval i jeho.
Or: I jeho kritizoval HONZA.
(Who criticized even Mirek Dušín as a tool of capitalism?) JOHN criticized even him.
Or: (lit.): Even him criticized JOHN

(8') HONZA kritizoval-b i-b jeho-c.
Or: I.b jeho.c kritizoval.b *HONZA*.

It is the main motivation of our present study to identify some further criterion for the oppositions of topic, contrastive topic and, as the case may be, of focus in the prosodic characteristics of utterances. For this purpose, we have examined three small corpora of naturally occurring speech (described in Sect. 2) and analyzed them in order to find whether the character of intonation contours distinguishes between the three notions (Sect. 3). The results of our analysis are discussed in more detail in Sect. 4. and Sect. 5. and then summarized in Sect. 6.

2. Spoken data

2.1. Description

In our research we used three files of annotated spoken data from a spoken corpus compiled in the Institute of Formal and Applied Linguistics, Charles Univesity, Prague. These three small corpora are recordings of three talk-shows of Czech TV called Na plovárně (On a bathing place), moderated by a well-known and highly intelligent moderator Marek Eben; in these shows, he interviews three personalities of Czech cultural and sport life, namely Vladimír Komárek, a painter, Magdalena Kožená, a concert and opera singer, and Petr Jirmus, a pilot of an acrobatic plane. In the table, the three interviews are indicated by the surnames of the persons interviewed. It should be noted that the moderator does not interrupt his partners very often, he remains in the background.

We decided to work with spontaneous dialogues, because it made it possible to look at main different realizations of the basic TFA categories – from one short word to longer and more complicated phrases. Even though we could not compare particular cases so well, because every context was unique, we were able to examine many possible problems caused by the wide variability such complex categories induce.

The dialogues were recorded with 22kHz sample frequency and with 16-bit resolution. We generated for each dialogue the smoothed F0-contours (fundamental frequency contours) with the help of the Edinburgh Speech Tools (EST 2000) software. The F0-contour extraction was done with the time step 0,01sec and for the 40Hz-400Hz frequency range.

Table 1 describes the turn structures of the three dialogues. The dialogue identification is followed by the number of speakers, the whole length of the dialogue, the number of speaker turns and the mean number of moderator's and visitor's utterances over all turns.

Table 1

dialogue name	No. of speakers	duration (sec)	No. of turns	the mean length of turns (sp1/sp2)
Plovárna Komárek	2	1298,33	60	1,903/9,069
Plovárna Kožená	2	1307,15	70	2,500/6,822
Plovárna Jirmus	2	1308,14	138	1,985/3,357

Table 2 shows the word statistics and the percentual proportions of moderator's (sp1) and visitor's (sp2) talk time for each dialogue.

Table 2

dialogue name	No. of words	the proportion of the speakers (sp1/sp2) (%)	the mean No. of words per utterance (sec)
Plovárna Komárek	3791	27,5/72,5	11,59
Plovárna Kožená	3363	18,3/81,7	10,44
Plovárna Jirmus	3636	36,5/63,5	9,83

2.2. Segmentation

The dialogues were transcribed manually with the help of the Transcriber program (Transcriber 2002), see Fig.1.

The following segmentation of the text has been performed:

2.2.1 Segmentation into **phonemes** and **words** was carried out automatically by the HTK software (HTK 1999). The phone boundaries are set by Viterbi forced alignment; phone models trained for Czech phone models are used. The boundary of a word begins with the beginning of the first phoneme of the word and ends with the end of its last phoneme. The pauses between words and non-articulated sounds are set apart.

2.2.2 Segmentation into **speech segments** has been done manually. The transcribers inserted synchronization marks at those places where the end of the sentence was followed by a pause. These synchronization marks define the speech segments. In the case of longer sentences or sentences without the final pause the transcribers inserted synchronization marks at the nearest longer pause. The main criterion for the segmentation consisted in the length of the pause. However, this purely phonetic articulation cannot be used for our purposes, because a longer pause in a spoken text may be connected with some uncertainty of the speaker and his searching a suitable expression or an appropriate continuation of the discourse. Such a hesitation, of course, may occur in any point of the discourse, not only on the boundaries of utterances but also in the middle of them. In addition, a segmentation of discourse by means of pauses influences also the tempo and rhythm of the discourse – longer and more complicated discourses will tend to be segmented in a more detail than the shorter ones, and more simple discourses will tend to include into more than one utterance one speech segment. Thus such a segment may correspond to units of different kinds: several utterances, a single utterance, a clause, a phrase or a part of a phrase.

2.2.3 Segmentation into **turns** is carried out by hand according to the following principle: a turn is characterized as a textual segment uttered by one speaker. This is not so simple, of course, and leads to a certain distortion, because some turns consist only in expressions or sounds of confirmation; these sounds actually do not lead to an interruption of the monologue of the speaker, the continuity of which is very important. Overlapped segments were excluded from our study because of our inability to separate overlapped voices.

3. Processing and analyzing the data

3.1.   Modifications of the segmentation

As already mentioned, we had to make some additions and modifications of the scheme described in Sect. 2.2.

The word boundaries remain as characterized in Sect. 2.2.1, because their definition is motivated phonologically. This has influenced also our characterization of utterance events, see below.

For the reasons mentioned above in Sect. 2.2.2., we had to modify the specification of the boundaries of speech segments for our purposes. In addition to the phonetic criterion the labelers added marks for the boundaries of utterances: the delimitation of an utterance is crucial for our analysis, because it serves as the basic unit for the determination of TFA.
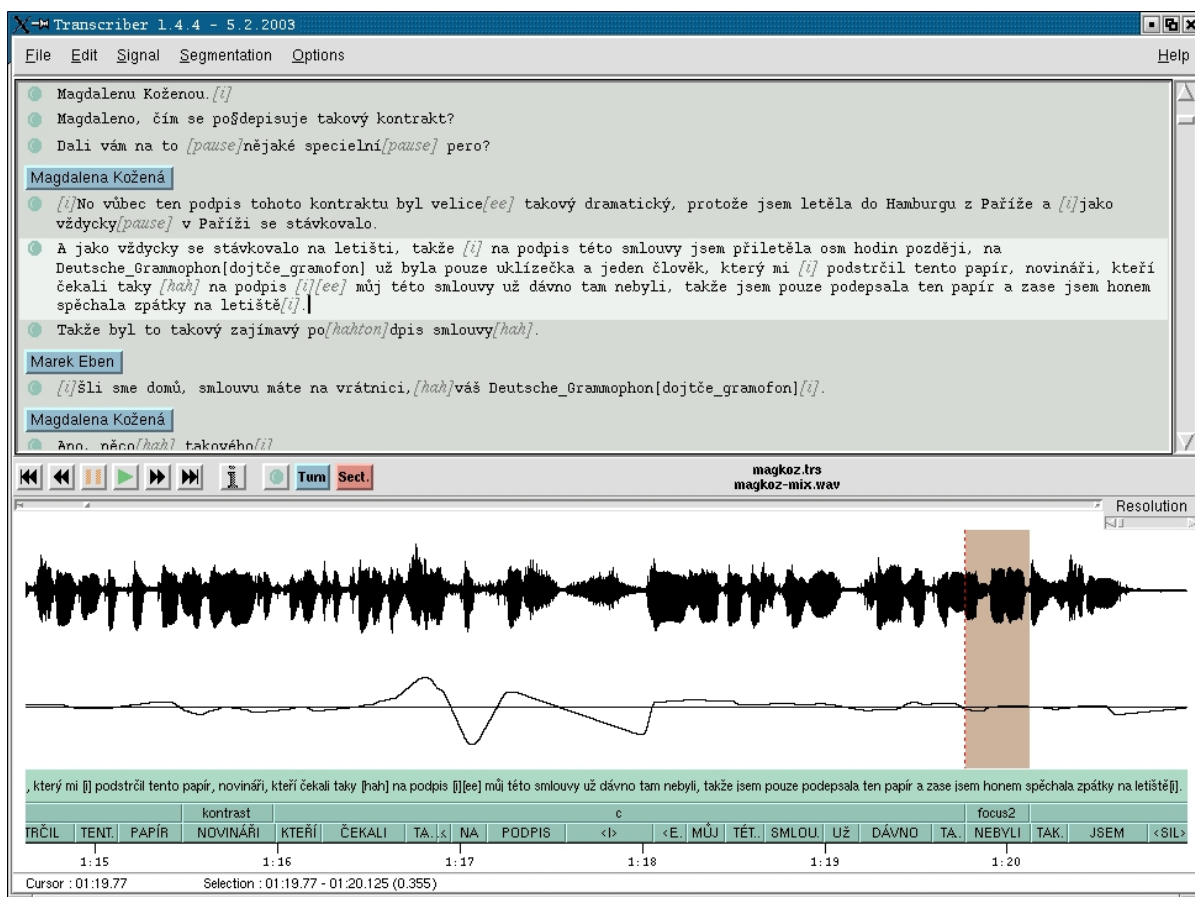
Fig. 1. The Transcriber software window.
In the Transcriber software window the top section displays sentence transcriptions. The middle part of window displays a waveform and an F0 contour. The bottom part displays three layers of our segmentation – the sentence segmentation, the contrast-focus segmentation and the word segmentation.

Note. The difference between a sentential segment and an utterance lies in the fact that a sentential segment (i.e. the segment of the text from a fullstop to a fullstop) may contain also coordinated clauses, which in turn contain several coordinated utterances. However, syntactic analysis and also the TFA analysis treats clauses in coordination relation as independent utterances. In contrast to this, subordinated clauses are treated as one of the means of expression of dependency relations, i.e, on a par with complementations of the main clause.

3.2 Sectors of an utterance

3.2.1 General description

The segmentation of utterances into sectors is guided by structural and semantically relevant criteria rather than by the phonetic shape of these sectors. A sector is a node or a subtree of the tectogrammatical tree structure (which represents a complementation of the main verb or of some dependent node); this node/subtree may be rendered in the surface shape of the sentence by a single word, a prepositional or prepositionless nominal group, or by a whole clause in case of a complex ('subordinate') sentence.

In our research we did not use any segment shorter than one word. In addition to the semantical reasons (described above) we suppose the length of a whole word to be sufficient to show a relevant difference between the word-level acoustic characteristics and those of prosody of selected sectors, which are, in our opinion, much more hard to separate on the layer of the syllables or the phones.

For evident reasons, we left aside in our analysis those cases when a complementation was rendered by a clause; dependent clauses have an inner structure and phonetic realization of their own and are 'recursively' articulated into more subtle parts.

The simplest (and, of course, ideal) cases for our analysis and for the assignment of the TFA values are those when a complementation is expressed in the surface shape of the sentence by a single noun (with or without a preposition). In our corpora, these cases are in majority.

However, when isolating the cases of contrast from those of focus proper, we must also deal with more complex nominal groups, for which we had to formulate additional rules. A complex nominal group, especially if it is a longer one, has a tendency to be articulated into a contextually bound (*cb*) and non-bound (*nb*) part and the boundary between these two parts has to be taken as relevant (see Sect 1.3 above). Contextually bound attributes occur rather rarely and the presence of a *cb* modifier in the surface shape of the sentence is either required by the grammatical structure or the same (or semantically close) adjectives precede different nouns that are in the relation of a contrast – they 'prepare' the ground for a contrast. Such *cb* members of an utterance depending on a non-bound or a bound, but contrastive node are phonetically suppressed, which is usually reflected in the fall of the vocal tone when they are pronounced. The boundaries of sectors within nominal groups were therefore drawn between the governor and its cb attributes.

In case of the presence of a focalizer (a focusing particle such as *jenom* 'only', *dokonce* 'even', *taky* 'also' etc.), these focalizers introduce a contrastive topic in the topic part of the sentence or the focus proper in the focus of the sentence. Since the function of these particles is to signal these TFA characteristics, utterances with focalizers are also included in our material.

3.2.2 Types of sectors

Three main types of sectors are differentiated in our material (see above in Sect.1 3 for the theoretical prerequisite of such distinctions):

(a) Non-contrastive topic: The category of topic can be paraphrased as 'what is the sentence about' (see Sect. 1.3 above). From the TFA point of view, in unmarked cases the initial position of a Czech sentence may be occupied in two ways: by a contextually bound non-contrastive item OR by a contrastive topic. Our starting hypothesis was that these two cases should be distinguished in the intonation contours, so that if we succeed to distinguish in a sufficiently clear way the realizations of topic and contrastive topic, we can speak about a typical contrastive contour. Thus the delimitation of topic is the starting point of the delimitation of contrastive topic. For this purpose, we have selected again 100 utterances with an evidently well-recognizable topic, based on the criterion of a repetition of the expression from the preceding utterance (ex. (9)) or a clearly describable link to the preceding utterance (ex. (10)).

(9) (Kdy jsi ho viděl?) Viděl jsem ho VČERA.
(Where have you seen him?) Lit.: I-saw Aux him yesterday

(9').[já-b] ho-b viděl.b *VČERA*

(10) (Nakonec přece jen dojedeme do Nairobi.) V hotelu Stanley bydlíval HEMINGWAY.
(Finally we arrive at Nairobi.) In the hotel Stanley used to live Hemingway.

(10') (V) hotelu-b Stanley-b *bydlíval HEMINGWAY.*

(b) Contrastive topic: Since contrast is a rather unclear and oscillating category, contrastive topics have been determined by means of a number of criteria, the main being the (univocal) contrastive character of their 'linking' in the text. In the first phase of our analysis, we have therefore tried to exclude vague or somewhat doubtful cases. The quantitative limit was to obtain at least 100 cases of contrastive topics, which was the lowest number of cases we wanted to have for the purpose of further analysis. The transcribed sentences were annotated manually by labelers who were instructed to use the following criteria and apply them in the following order (from the most significant to the least significant):

(i) Textual criteria: Contrastive topics constitute first of all a category which to a great and important extent contributes to the structuring of a text and which is one of the most important TFA factors (if not the most important one) for the representation of the intersentential links. Thus the most important role in our decisions was played by the relations between utterances and between their initial parts. One of the most decisive indicators of contrast was an insufficient predictability of the initial part of the utterance from the preceding context (ex. (11)), or a participation of this part in structures representing an enumeration (ex. (12)).

(11) (Včera se hrál zápas mezi Brnem a Ostravou.) Domácím se dařilo ZE ZAČÁTKU. Hostům se povedl až druhý POLOČAS.
(Yesterday the match between Brno and Ostrava took place.) Lit.: Hosts.Dat were-successful at beginning. Visitors.Dat succeeded only second half.
'The hosting team was successful at the beginning. The visiting team succeeded only in the second half.

(11') Domácím-c se-dařilo.b (ZE) *ZAČÁTKU.* Hostům-c se-povedl-b *až druhý POLOČAS.*

(12) (Loni jsme renovovali celý dům.) Omítku. jsme natřeli NA ŽLUTO, střechu jsme TAKY vyspravili, všechna okna jsme vyměnili ZA DVOJITÁ.
(Last year we renovated the whole house.) Lit. The plaster (we) painted yellow, the roof (we) also fixed, all windows (we) replaced by double (ones).

(12') Omítku-c. (jsme) natřeli-b (na) *ŽLUTO,* střechu-c (jsme) *TAKY* vyspravili-b, všechna okna-c (jsme) vyměnili-b (na) *DVOJITÁ.*

(ii) Semantic criteria: Contrast is in itself a phenomenon founded semantically, including such phenomena as antonymy, cohyponymy and the semantic part-and-whole relation, which all signal a contrastive relation between two parts and motivate the use of a contrastive topic as a syntactic category (ex. (13))

(13) (První pohled na nový model Toyoty potvrzuje, že se jim jejich záměr podařilo naplnit.) Karosérie je velmi ATYPICKÁ., motor je DVOULITROVÝ, …
(The first glance at the new model of Toyota confirms that they managed to accomplish their intentions.) Lit. The body is very untypical, the motor is two-litre, …

(13') Karosérie-c je-b *velmi ATYPICKÁ.*, motor-c je-b DVOULITORVÝ, …

(iii) Structural criteria: Within the boundary of a single utterance, one of the signals of a contrastive topic is also the syntactic structure of the given utterance: An important signal is the type of the syntactic relation (dependency) of the given node to its governor (i.e. to the governing node): participants (i.e. arguments of the verb), at least in Czech, have a stronger tendency to be in the position of a contrastive topic than free modifications (i.e.adjuncts, such as temporal or local settings). Another important aspect for contrastive topic in Czech is the position of the given element in the sentence: the initial position is very typical , but there are also other signals such as a non-projective characteristics of the surface word order (long-distance dependency, discontinuous structure), see ex. (14).

(14) Karla se nám podařilo poslat do NĚMECKA.
Lit. Karel-Accusative we managed to send to Germany.

(14') Karla-c (se) nám-b podařilo-b *poslat (do) NĚMECKA.*

(c) Focus: As a rule, focus is understood as a part of the sentence that is signaled by including the bearer of the intonation center and by its unmarked position at the end of the utterance, i.e. also by a characteristic final cadence. Most theories of TFA assume that every utterance must have a focus. Since one of the main tasks of our analysis was to attempt to demonstrate a difference in the contours of focus sectors and contrastive topics, we have singled out for our analysis of foci those utterances in which a contrastive topic was identified. This was a simple and at the same time a transparent criterion from the point of view of comparison of differences in the intonation contours of the two categories (i.e. both

occurred in the same sentence).

Focus segments, however, may also occur in a marked position, i.e. either at the beginning (see ex. (8) above repeated here as (15) for convenience) or in the middle of an utterance (see ex. (1) repeated here as (16)).

Or: I jeho kritizoval HONZA.
(Who criticized even Mirek Dušín as a tool of capitalism?) JOHN criticized even him.
Or: (lit.): Even him criticized JOHN

(15') HONZA kritizoval-b i-b jeho-c.
Or: I.b jeho.c kritizoval.b *HONZA*.

(16) (She had separated from her first boyfriend with no great pain. With the second it was worse. …) She LOVED him, and he was …

(16') She-b *LOVED* him-b

Such a marked order of elements was called by Mathesius (1947, p. 241) a subjective order. An analysis of such utterances may help us to find out, whether an eventual difference between the intonation contours of contrast and focus is not conditioned just by position in the utterance (initial vs. final) and whether the characteristics obtained are not typical just for a specific sentential position rather than for a TFA category. This has led us to look at cases of focus in marked positions. This is a rather rare phenomenon in Czech, though, and therefore we have analyzed all such cases. In our data we called them Sfocus (subjective focus). In this way, it was also possible to follow the difference between the intonation center and the so-called emphasis – a secondary marking of a certain part of the utterance.

3.2.3 A more detailed classification of the basic types

The category of **focus** has been further divided into two types: (i) the respective sector is the final sector of the given utterance (focus1), see ex. (17), (ii) the sector is the focus of an utterance that stands in a coordination relation within a larger whole (focus2), see ex. (18). The reason for such a subdivision was to make it possible to find out the influence of a final cadence or half-cadence on the examined intonation contour.

(17) Tom včera navštívil svou SESTRU.
Lit. Tom yesterday visited his sister.

(17') Tom-b včera-b *navštívil svou-b SESTRU*

(18) Tom včera navštívil svou SESTRU a koupil KYTICI.
Lit. Tom yesterday visited his sister and bought a flower.

(18') Tom-b včera-b *navštívil svou-b SESTRU* a *koupil KYTICI*.

According to the strength of the contrast, the category of **contrast in topic** has been subdivided into three types; the decision of the assignment of one of these categories was based on the size of the set of alternatives from which the contrastive topic has been chosen, its means of expression and the way of the choice of the contrastive topic from this set (see Hajičová, Sgall, Veselá 2003).

Contrast 1: the strongest type of contrast, when an enumeration of elements of a set of alternatives continues, these elements are in the same semantic class and in an analogous position in the sentence, ex. (19).

(19)   (Dalšími uchazeči o pořadatelství světového šampionátu v roce 2002 jsou Korejská republika a Mexico,) přičemž Korea.c již svou žádost *PŘEDLOŽILA*.

(Further candidates for organizing the World Championships in 2002 are the Republic of Korea and Mexico,) whilst Korea has already SUBMITTED its application.

(19') (přičemž) Korea-c *již* svou-b žádost-b *PŘEDLOŽILA.*

Contrast 2: a weaker type of contrast, the contrastive topic is the first choice of an element of the set of alternatives and it carries information about the articulation of this set, ex. (20).

(20) (Na výlet jela celá třída.) Učitel připravil ITINERÁŘ, kluci stavěli STANY, dívky VAŘILY.

(Lit. At trip went whole class.) The teacher prepared the itinerary, the boys were building the tents and the girls were cooking.

(20') Učitel-c *připravil ITINERÁŘ*, kluci.c *stavěli STANY*, dívky.c *VAŘILY.*

Contrast 3: the weakest type of contrast, the contrastive topic is selected from the 'hypertheme' of the discourse segment, and the main signal of contrastivity is the non-derivability of the contrastive element from the (immediately) preceding context, see ex. (12) above, repeated here as (21).

(21) (Loni jsme renovovali celý dům.) Omítku. jsme natřeli NA ŽLUTO, střechu jsme TAKY vyspravili, všechna okna jsme vyměnili ZA DVOJITÁ.
(Last year we renovated the whole house.) Lit. The plaster (we) painted yellow, the roof (we) also fixed, all windows (we) replaced by double (ones).

 (21') Omítku-c. (jsme) natřeli-b (na) *ŽLUTO*, střechu-c (jsme) *TAKY* vyspravili-b, všechna okna-c (jsme) vyměnili-b (na) *DVOJITÁ.*

4. Outputs of the analysis

4.1   Tables

We present first the results in a tabular form, followed (in Sect. 4.2) by the description of the factors measured.

Table 3

type of sector	# of occur.	length (msec)	mean value (F0)	beginning (F0)	end (F0)	range (F0)	rise | fall (F0)	difference (F0)	std. Deviation of difference
Topic	97	615,6	137,02	134,31	135,91	30,78	14,79	1,61	34,29
contrast1	31	617,7	136,73	123,86	142,39	45,48	32,59	18,53	48,33
contrast2	30	614,3	149,54	137,06	156,43	40,63	32,67	19,37	39,37
contrast3	50	809,4	138,34	122,65	142,48	48,37	37,28	19,83	34,85
Focus1	32	837,8	126,19	129,92	115,50	49,49	30,82	-14,42	52,46
Focus2	83	669,2	131,22	136,21	122,03	44,57	30,20	-14,18	46,93
Sfocus	25	646,4	136,69	139,25	126,12	40,18	29,73	-13,13	37,57

4.2 Quantified parameters

Table 3 presents the results of the phonetic analysis of our data, selected from the collections of authentic speech records described in Sect. 2 above. We summarize there the maximum of data considered potentially relevant from some point of view. All these values are taken from the F0 curve, because we take the intonation contour to be the most important factor of the intonation of Czech sentences. All F0

values presented here are calculated as mean values of all sectors of the given type. Both tables reflect aggregated data (across all the speakers); in the expanded and modified version of our paper we follow the data also for individual speakers.

The following values have been calculated:

1. the number of occurrences of the given sector type

2. maximum – the maximal value of F0 measured on the given sector

3. minimum – the minimal value of F0 measured on the given sector

4. mean – the mean height of F0 on the given sector

5. beginnings of the sectors – the value of F0 of the initial point of the sector

6. ends of the sectors – the value of F0 of the final point of the sector

7. length – length of the sector in milliseconds

8. difference – the overall tendency of the sector – the difference between the initial and the final values of F0 of the given sector

9. range – the difference between the maximal and minimal value of F0 of every sector

10. rise | fall – the difference between the maximal and lesser of the initial and final F0 values of the given sector

11. standard deviation of difference square – root of difference variability

4.3 Results

4.3.1 The height of the tone at the beginning and at the end of the sector: an overall tendency

From the point of view of interpretation, the height of the tone at the beginning and at the end of the sector is not relevant. It depends on the overall position of the voice of the speaker, and as such it is important rather for a description of the subjective factors that are at play in the discourse (see Sect. 5.).

The data of Table 3 provide several properties that can be generalized in the following way:

The overall tendency of a sector is given by the difference of the F0 values of the initial and final point of the given sector. This difference shows whether and to what extent the given contour is rising or falling.

The column "difference" indicates quite clearly that the material can be subdivided into three groups: the sectors of the type topic occur at the point near to zero, the sectors of the focus type in the negative values and the sectors of the type contrastive topic in the positive values. We can thus say that the typical contour of contrast is a rising one, the typical contour of focus the falling one and that topic has a more or less constant contour.

As for the characteristics of focus, the absolute values of the mediums /media in different categories indicate that the falling contour of focus remains evident irrespective of the placement of focus in the utterance. The values for the final and initial position of focus and for focus in an unfinished or coordinated utterance are very close.

There are also no evident differences between the individual groups of contrastive topic. The resulting values rather indicate that the strength of contrast does not have an influence on the phonetic shape of the sentence.

4.3.2 The length of the sector

The length of a sector (the third column of Table 3) is given here only for informative reasons; as we have mentioned above, the segmentation of the discourse into sectors is based on semantic criteria and thus the sectors may have very different lengths. The length rather has an influence on the values of other parameters: in case the sectors are longer, where the length may have an influence on the tempo of speech and in this connection also on the intonation contours, a question may arise whether the values measured in the initial and the final points of the sector and its maximum and minimum are well representative.

4.3.3 The range of the sector

The parameter of the range of the sector is defined by the difference of the values of the maximum and the minimum. The values of the range reflect the tendencies mentioned above: the range of all sectors lies between forty and fifty Hz, with the exception of the range of the topic sectors, which is significantly lower. The sound realization of the contextually bound sectors is thus "flatter", with smaller digressions in the vocal tone, which may be connected with their lower importance and their role of a kind of "fillers" in the sentence structure.

4.3.4 Combination and hierarchization of the parameters

In order to identify most useful parameters for the distinguishing of particular types of sectors it was necessary to combine the measured parameters and to evaluate these combinations. Most of these combinations have been described in the preceding sections together with the characterization of the individual parameters; below we only summarize the findings:

Range

The parameter of range was established from the values of the minimum and the maximum, i.e. the value of their difference. This indicates the overall span of the given sector, regardless of the positions of the minimum and the maximum.

The overall tendency

The overall tendency, i.e. the difference between the initial and the final value of the sector, indicates whether and to what extent the given sector is rising or falling. This parameter is important not only because it can be well and relatively easily captured, but also because the rise|fall is an important semantically relevant phenomenon in the intonation contour of a Czech sentence (e.g. for the recognition of sentential modality).

Rise|fall

The parameter of rise|fall of the sector is defined as the difference between the maximal and lesser of the initial and final F0 value of the given sector. We tried to find out, whether the whole rise/fall value or only rise/fall of the F0 value to/from the highest value of the segment is more relevant for rising/falling identification.

The distribution of rise|fall values is rather regular – focus and contrast have about 30 Hz (with the small exception of contrast3), topic has rise|fall value about 15Hz.

Hierarchization of parameters

On the basis of the results presented above we can attempt at a hierarchization of the measured parameters and their combinations according to their ability to help to distinguish between individual types of sectors. It is quite evident that the most important of the values is the difference between the initial and final value, which perspicuously differentiates the types of focus, contrast and topic. These values, of course, are incident to the values of the beginnings and ends of the sectors and the rise|fall of the sector. The parameters concerning the values of minimum and maximum can be used as an additional information about the shape of the contour in the given sector.

4.3.5 Distribution of the overall tendency values

We focused our observation on details of the difference between the initial and the final value of the sector and we were interested the in number of occurrences of segments with characteristic values – rising for contrast, falling for focus and the constant contour for topic. The distribution of the values of the difference in layers differing by 10 Hz is given in Table 4.

Table 4

	topic	contrast1	contrast2	contrast3	focus1	focus2	Sfocus	k	f
-100 – 90	-	-	-	0		2			2
-90 – 80	-	1	-	0	2	1		1	3
-80 – 70	-	0	-	0	1	3			4
-70 – 60	1	0	-	0	1	4			5
-60 – 50	2	0	-	1	1	1			2
-50 – 40	3	0	-	0	2	3			5
-40 – 30	2	0	1	0	2	4	3	1	9
-30 – 20	3	1	0	0	2	9	4	1	15
-20 – 10	12	1	4	5	2	11	3	10	16
-10 – 0	17	3	2	8	9	16	4	13	29
0 – 10	36	7	8	8	2	14	4	21	20
10 – 20	8	6	4	10	3	4	2	20	9
20 – 30	2	2	4	5	3	1	2	11	6
30 – 40	2	1	2	1	0	3		4	3
40 – 50	3	2	2	1	1	1		5	2
50 – 60	1	2	2	2	2	0		6	2
60 – 70	2	1	0	4	-	1		5	1
70 – 80	1	0	0	2	-	-		2	
80 – 90	-	1	1	1	-	-		3	
90 – 100	-	1	-	1	-	-		2	

Table 4 shows a relatively high variability, but the largest number of occurrences is around zero for topic, in the range of 0Hz-20Hz for contrast and in the largest range –30Hz-10Hz for focus. The domination of falling F0 tendency is thus confirmed for focus as well as the domination of rising F0 tendency for contrast and the domination of constant F0 tendency for topic. The relatively high variability is mostly caused by subjective factors (described below).

14

5. Deviations

5.1 Technical factors

Few errors could have been introduced into our statistic by some wrong word boundary detection (usually due to a wrong boundary detection of plosive phones) or by some wrong F0-values extraction at voiced consonants and at speech segments boundaries where a small F0-contour flattening occurs by smoothing.

5.2 Subjective factors

Quite understandably, the mean values of the measured values are influenced by the manner of speech used by individual speakers. In this domain, there are only relatively few phenomena that can be expressed in a quantitative way, out of which the following ones can be mentioned here:

(i) The mean height of the speaker's voice

The mean height of the tone used by the particular speaker has an influence on all non-relativized values in our calculations, esp. the values of F0 of maximum and minimum of individual sectors and on the initial and final values of F0 in the given sectors. The speaker4 has a significantly medium height of voice, which is quite understandable because she is a female. However, even among the male voices (speaker1, 2 and 3) there is big dispersion in the parameter of the height of the voice. It is thus evident that the mean values of individual types of sectors reflect also the proportion in which individual speakers contribute to them. To make our survey more transparent, we present in Table 5 some details about this proportion: for each speaker, in the left column there is the overall number of occurrences of the given sector and in the right column we give the percentual share that particular speaker has for the given type of sectors. Since the sectors have been determined on the basis of their syntactico-semantic properties, we did not take care of a proportional representation of individual speakers, as Table 5 clearly documents. Therefore higher values of mean values will be found with those types of sectors in which the share of speakers with a higher voice is greater.

Table 5

type of sector	speaker1		speaker2		speaker3		speaker4		sum
topic	12	12,4	27	27,8	25	25,8	33	34,0	97
contrast1	1	3,2	13	41,9	10	32,3	7	22,6	31
contrast2	4	13,3	3	10,0	7	23,3	16	53,3	30
contrast3	9	18,0	13	26,0	20	40,0	8	16,0	50
focus1	8	25,0	6	18,8	5	15,6	13	40,6	32
focus2	7	8,4	25	30,1	31	37,3	20	24,1	83
Sfocus	5	20,0	7	28,0	7	28,0	6	24,0	25

(ii) An overall 'tendency' of the utterance

The overall 'tendency' of the utterance, i.e. the difference between the initial and the final value of F0, is another important factor. This factor reflects the tendency of the speaker in the course of the utterance either to raise or to fall in his voice, and how sharply. The third column of Table 6 indicates that with most speakers the utterance ends roughly at the same level of F0 as it starts, or, more precisely, it ends a little bit higher. The fifth and the sixth columns indicate that the utterances begin and end below the level of an

average utterance, while the magnitude of the difference varies significantly. This fact, however, is related to a greater extent to the range of the utterance. With the exception of speaker1, with whom the utterance falls down more sharply, the other speakers exhibit no differences with regard to the utterance tendency. It can be then assumed that this value has no influence on possible digressions within the frame of the measured sectors.

Table 6

speaker	mean (F0)	difference (F0)	range (F0)	length (msec)	beginning rel. (F0)	ending rel. (F0)
speaker1	118,44	-9,18	164,39	575.30	-5,94	-15,02
speaker2	99,51	1,39	136,57	426.06	-3,20	-2,13
speaker3	132,74	2,29	205,33	810.84	-15,50	-13,60
speaker4	159,49	3,94	214,15	655.87	-11,71	-7,77

(iii) The range of the utterance

The medium of the range of the utterance of one speaker demonstrates how big the digressions of the vocal tone the speaker uses for the signalization of the speech sectors, if his/her utterance is more "articulated" or rather "flat". The values in the third column of Table 6 indicate that in this respect speakers varied to a large extent. A question arise how far these individual difference influence the average values of the range of the measured sectors. If the speakers are subdivided into a group with a lower range (speaker 1 and 2) and with a higher range (speaker 3 and 4), the proportion with a majority of the types of sectors is around 50%-60% for the latter group. The only exception is the category of contrast2, in which over 75% of occurrences belong to speakers with a higher range. The mean value of the range of the measured sectors does not differ from other sectors. It is thus possible to assume that the individual use of the range of the tone of the utterance has not a significant influence on the values investigated.

(iv) Other factors:

In addition to the height of voice and other quantified features, also purely individual tendencies can be observed, that is a specific way of realizations of the individual sectors.

From the point of view of the evaluation of the data gained from the individual speakers, one should not neglect their role in the dialogue. The dialogues we have analyzed are based first of all on the contributions of the invited speakers; speaker 1 (he is the moderator of the all the programmes) only poses questions and controls the course of the dialogue. His utterance are significantly shorter, which can be seen in Table 1 (the average number of utterances in his replicas is about 2), his share in the whole discourses is thus much lower than that of the other speakers (about 30% of the total number of utterances). It is also necessary to take into account that almost 30% of his utterances are interrogative sentences, the intonation contours of which is very different from those of indicative sentences from which our sample was selected. These factors certainly have an impact on the apparent difference in the prosodic characteristics of his speech and the speech of other speakers. The high frequency of the falling tendency of his sectors probably reflects a lower frequency of coordinated structures and the falling cadence of Czech wh-questions.

6. Conclusions

6.1. On the basis of the results presented above we can attempt at a hierarchization of the measured parameters and their combinations according to their ability to help to distinguish between individual types of sectors. It is quite evident that the most important of the values is the difference between the initial and final value, which perspicuously differentiates the types of focus, contrast and topic. These values, of course, are incident to the values of the beginnings and ends of the sectors and the rise/fall of the sector. The parameters concerning the values of minimum and maximum can be used as an additional information about the shape of the curve in the given sector.

6.2. The column "difference" indicates quite clearly that the material can be subdivided into three groups: the sectors of the type topic occur at the point near to zero, the sectors of the focus type in the negative values and the sectors of the type contrastive topic in the positive values. We can thus say that the typical contour of contrast is a rising one, the typical contour of focus the falling one and that topic has a more or less constant contour.

6.3. The parameter of the range of the sector is defined by the difference of the values of the maximum and the minimum. The values of the range reflect the tendencies mentioned above: the range of all sectors lies between forty and fifty, with the exception of the range of the topic sectors, which is significantly lower. The sound realization of the contextually bound sectors is thus "flatter", with smaller digressions in the vocal tone, which may be connected with their lower importance and their role of a kind of "fillers" in the sentence structure.

References

Braun, Bettina and Robert Ladd (2003), *Prosodic Correlates of Contrastive and Non-Contrastive Themes in German*. Poster presented at Eurospeech 2003, Geneva.

Buráňová Eva, Hajičová Eva and Petr Sgall (2000), Tagging of very large corpora: Topic-Focus articulation. In: *COLING Proceedings*, 139-144. Saarbrücken, Universität des Saarlandes.

EST (2000). Alan Black, Paul Taylor, *Festival Speech Synthesis System & Edinburg Speech Tools*, University of Edinburg, 2000

Hajičová Eva (2002), Topic-Focus Articulation in the Czech National Corpus. In *Language and Function*. Ed. Hladký, J. Praha, 2002, 185-194.

Hajičová Eva (2003), *Contextual boundness and discourse patterns*. Presented at the 17th Int. Congress of Linguists, Prague, 2003; to be published in the Proceedings of the Congress.

Hajičová Eva; Partee, Barbara; Sgall, Petr (1998), *Topic-focus articulation, tripartite structures, and semantic content*. Dordrecht, Kluwer 1998.

Hajičová Eva and Petr Sgall (2001), Topic-Focus and Salience. In *Proceedings of the 39th Annual Meeting of ACL*, Toulouse, 2001, 268-273.

Hajičová Eva, Sgall Petr and Kateřina Veselá (2003), Information Structure and Contrastive Topic. In *Annual Worskhop on Formal Approaches to Slavic Languages*. Ed. Brown, Wayles et al. Ann Arbor, Michigan Slavic Publications.

HTK (1999). Steve Young, Dan Kershaw, Julian Odell, Dave Ollason, Valtcho Valtchev, Phil Woodland, HTK Book, Entropic Ltd, 1999

Koktová Eva (1999), *Word-Order Based Grammar*. Berlin, Mouton De Gruyter.

Mathesius Vilém (1947), O tak zvaném aktuálním členění větném in Mathesius, Vilém: *Čeština a obecný jazykozpyt*. Praha, Melantrich 1947, s. 234-242.

Peregrin Jaroslav (1994), Topic-focus articulation as generalized quantification. In: Bosch P. and R. van der Sandt, eds.: *Focus and natural language processing*. IBM Working Paper 7. Heidelberg: IBM Deutschland, 379-388. To be printed in Prague Linguistic Circle Papers 4.

Pražský závislostní korpus [Prague Dependency Treebank] (2001). http://ufal.mff.cuni.cz/pdt

Sgall Petr, Hajičová Eva and Jarmila Panevová (1986): *The Meaning of the Sentence in Its Semantic and Pragmatic Aspects*. Ed. by Mey, J. L. Dordrecht, Reidel; Prague, Academia 1986.

Transcriber (2002). Claude Baras, Transcriber Software, DGA,2002
http://www.etca.fr/CTA/gip/Projects/transcriber,