Information Structure

Language, Context, and Cognition

Edited by
Anita Steube

Volume 1

W
DE
G

Walter de Gruyter · Berlin · New York

Information Structure

Theoretical and Empirical Aspects

Edited by
Anita Steube

Contents

Contents

Introduction

"Information Structure: Theoretical and Empirical Aspects" is the first volume in the series "Language, Context, and Cognition". The series was established as the organ of the Research Group "Linguistic Foundations of Cognitive Science: Linguistic and Conceptual Knowledge" at the University of Leipzig, sponsored by the German Research Foundation as Research Group 349.

Several projects in the research group have been concerned with information structure already from 1999 to 2002. They are especially responsible for this volume:

Ewald Lang and Gerhild Zybatow (Marcela Adamíková, Dorothee Fehrmann): *CONTRAST and CORRECTION in the interaction of the lexicon, syntax, information structure and prosody with special attention on the Slavonic languages*
Ingolf Max (Claudia Henning, Werner Wolff, Nils Kürbis): *Multidimensional representations of linguistic and world knowledge: studies on presupposition and negation in discourse*
Anita Steube und Kai Alter (Carla Umbach, Andreas Späth, Stefan Sudhoff): *Intonation and meaning in information structure*
Gerhild Zybatow (Dieter Wirth, Rolf Meyer, Klaus Abels): *Russian information structure*

The first publication of work done in these projects appeared in 2001 under the title "Kontrast – lexikalisch, semantisch, intonatorisch" (Contrast – lexical, semantic, and prosodic), in: "Linguistische Arbeitsberichte" 77, Leipzig. In February 2002, the four projects organized an international workshop on "Contrast and Information Structure". Before this workshop, we sent our results - documented in "Linguistische Arbeitsberichte" 79, Leipzig 2002 under the heading "Contrast" and "Information Structure and Grammatical Modularity" - to our partners in Berlin, Bonn, Budapest, Edinburgh, Eindhoven, Konstanz, Lund, Prague, Stuttgart, Tübingen asking for their comments and for their own approaches to information structure. We are happy to publish three of the feed-back contributions here: Petr Sgall and Eva Hajičová from Prague, Beáta Gyuris from Budapest, and Brigitta Haftka from Potsdam.

  Although there are already syntactic descriptions of nearly all phenomena of German information structure in different models now, and although there is a Minimalist description of Russian information structure of declarative matrix clauses as well, investigations on the impact of information structure on semantics and experiments on modern German and Russian prosody have only just begun. The few phonological parameters delimited so far are highly context dependent. Many more experiments will be necessary. Therefore, this

volume mainly presents models for information structural descriptions, seman-
tic contributions, contributions on the processing of prosody, and evidence
coming from Neuroscience. The papers by Ulrike Toepel and Kai Alter, Clau-
dia Hruska and Kai Alter, and Britta Stolterfoht and Markus Bader document
the close cooperation of the research group with the Max Planck Institute of
Cognitive Neuroscience in Leipzig. The text-structural influence on the topic,
the influence of focus sensitive particles on information structure, and the
topic-focus structure in Russian non-declarative clauses and dependent clauses
is left for research in the second three years' period of the research group
which began in 2003.

We will now introduce the papers of this volume individually.
    First the Praguian Functional Generative Model and the Generative Model
used in Leipzig for the description of German are presented as the background
for the description of modular grammar, information structure, and prosodic
contrast. The contributions on the semantic, syntactic and prosodic compo-
nents of grammar respectively, as in the model proposed by Steube, Alter, and
Späth, then follow. The second part of the book is devoted to the problem of
contrast in a broader sense accompanied by prosodic tests on the realization of
sentences containing German *aber*, and Polish, Czech and Slovak *ale* (engl.
but).

I. Modular grammar, information structure and prosodic contrast

Petr Sgall and **Eva Hajičová** make use of the framework of the Praguian
Functional Generative Description which includes a description of the topic-
focus articulation to explain phenomena connected with contrast and degrees
of its intensity. The semantic basis for the articulation of a sentence into
T(opic) and F(ocus) is understood as the relation of aboutness: in a proto-
typical declarative sentence F holds about its T. Normally the finite verb is F,
the DPs and PPs preceding it are topics, and those following it are focus.
    According to the context (1), only *one of his friends* is in the focus in (2).

(1)   *Whom was your brother visiting yesterday?*
(2)   *My.t brother.t was visiting.t one.f of his.t friends.f. yesterday.t*

Thus, within the Praguian Functional framework the topic-focus articulation is
not conceived of as a separate level of grammatical structure formation.
    The authors take their examples from the syntactically annotated Prague
Dependency Treebank. In Czech, lexical and phonological strategies for ex-
pressing focus and contrast are the so-called strong pronominal forms (as op-
posed to weak pronominal forms as used in bound contexts) and pitch accent.

When in M. Rooth 1985 focus is characterized as a choice from a set of alternatives, F always has a contrastive value. Since - in contrast to German - it is also possible in Czech to use (3) as well as (4) to answer (5), a grammatical value distinguishing CF (contrastive focus) from F seems not to exist. In German, *Buch* in (3) expresses contrastive accent. It answers (5) but (4) does not.

(3)   *Hans hat das BUCH dem Schüler gegeben.*
       Lit.: Hans has the BOOK to the student given.
(4)   *Hans hat dem SCHÜler das Buch gegeben.*
       Lit.: Hans has to the STUDENT the book given.
(5)   *Hat Hans dem Schüler das Buch oder das Bild gegeben?*
       Lit.: Has Hans to the student the book or the picture given?

The distinctiveness of contrastive accent in German as well as in Russian is also shown by Ulrike Toepel and Kai Alter as well as by Grit Mehlhorn in this volume.

   If more than one accented (contrasted) item is present in Czech, they get distributed among T and F. More numerous are the cases with a contrastive part of T. Sentence (7) has two contrasted topics in each of the conjuncts. In Steube, Alter and Späth these constructions are called Bridge Contours with the I-Topic as a contextually known part which semantically is part of the focus domain.

(6)   *Rodiče odjeli na dovolenou a děti svěřili příbuzným.*
       – The parents left for a leave and entrusted their children to their relatives.
(7)   *Syna.c ve čtvrtek.c zavezli.f do PARdubic.f, a dceru.c v sobotu.c do HRADce.f.*

The paper further differentiates several degrees of intensity of contrast and argues that different dimensions have to be distinguished.

Anita Steube, Kai Alter and **Andreas Späth** consider information structure as one of the late pragmatic factors that mark the pieces of propositional structure as being a topic, a background constituent (Haftka's 'anaphoric topic') or the focus constituent or rheme before they are realized in the grammatical module of a modular generative grammar (when the process is analysed as speaking). This means that information structure is considered to be of pragmatic origin but that it will be made visible and understood by the joint interaction of the different grammatical modules. It is not a separate module of grammar.

   Using the lexicon as an interface level between the cognitive model (with information structural pragmatics as one of its parts) and the grammatical model, grammatical computation begins with the semantic level of a two-level semantics. It maps its structures onto a generative surface syntax from where

the structure becomes the input to phonology / prosody. This way the pragmatic features T and F or CF are handed over into grammar and are realized on each grammatical level in the respective way.

The paper uses the model to describe German categorical sentences, correction sentences, and Bridge Contours. The constructions are characterized by their semantic, syntactic and prosodic properties. As there have not been enough prosodic experiments on Bridge Contours, the reader is asked to consider the contribution of Hruska and Alter and that of Mehlhorn on the realization and perceptive value of the contrastive accent in combination with the focus accent in this volume.

Andreas Späth and **Martin Trautwein** discuss what is negated in negative sentences, or, in other words, what negative sentences denote. This is very important for information structure because categorical sentences presuppose objects, times, and places, and the corresponding DPs have referential (i.e. specific, in-group, or partitive) interpretations.

(1)   *What about the child.*
(2)   *The child does not sleep.*

'The child' is anchored in context (1) and therefore presupposed. It exists independently of whether there is a positive or negative predicate on it. And the existing child can only take part in an existing situation. The proposition of which the situation is an instance is negative in (2). It is the propositions which are positive or negative, but not the events. The meaning of (2) is formalized as in (3).

(3) $\exists x \ [[\text{CHILD}, x] \ \& \ [\exists e \ \text{NOT} \ [\text{SLEEP}, x](e)]$

At the same time, (3) expresses the information structure of sentence (2) by representing the presupposed background constituent *the child* as the restrictor and the focus *does not sleep* as the nuclear scope of the construction. When there are no sentence adverbials such as *vielleicht, tatsächlich,* etc. in German sentences, negation also marks the border of the focus domain.

The paper does not consider sentences with a contrastive accent. It rather provides further ontological evidence on situations expressed by negative propositions.

Ingolf Max proposes a two-dimensional explication of the semantics expounded in Steube, Alter and Späth and in earlier papers of the authors. A general two-dimensional form of a sentence S is as in (1):

$$(1) \quad \lambda x_n \ldots \lambda x_1 \; \underset{\beta}{\overset{\alpha}{\maltese \mid \; \mid}}$$

in which: α   represents the proper ordinary meaning of S
            β   represents the background of S
            α and β have a conjunctive form.
            ☼ is the placeholder that marks the position of negation.

By specifying the format in (1) Ingolf Max can explicate many semantic notions and can create a model for a series of linguistic phenomena: semantic presupposition, entailment, assertion; correction sentence, Bridge Contour and the thetic sentence as a sentence with tautological presuppositions. This two-dimensional reconstruction brings to light hitherto unclear notions in information structure.

Beáta Gyuris describes Hungarian Bridge Contours the I-Topics of which are monotone increasing, monotone decreasing, or non-monotonic DPs. The author discusses in which question contexts these sentences are correct. A parallel English example with a non-monotonic DP is (2).

(1)   *Did you talk to five teachers?*
(2)   [*cf* / *To five teachers*] *I* [*f* \ *DIDn't*] *talk.*

After examining Büring's (1997, 1999) theory on Bridge Contours in detail, Beáta Gyuris argues that Kadmon's (2001) theory of discourse congruence is superior. She increases Kadmon's conditions on I-Topics and concludes that there is no evidence to support the binary division within DPs into those which can be licensed by the appearance of the same DP in the preceding discourse only (as in (1) and (2)) and those which can be licensed by other DPs as well. The paper convincingly shows that a declarative sentence with a contrastive topic DP has to satisfy the following three conditions to be licensed:

(i)   The last question under discussion for the declarative and the question preceding it overtly must be members of the topic value associated with the declarative.
(ii)   The two questions must be capable of serving as subquestions of the same superquestion in the discourse (determined by the topic value of the declarative).
(iii)   If the overt question is not identical to the last QUD for the declarative, the declarative must entail a complete answer to the overt question.

Since the phenomenon under review has hardly been described as yet, the paper makes an important contribution to the semantics and pragmatics of Bridge Contours far beyond those in Hungarian.

Brigitta Haftka's paper can be regarded as the volume's reference system on the syntactic order in the middlefield of German categorical sentences (without contrastive pitch accent). Having worked on the problems of German word order and information structure for years, Brigitta Haftka lays down all the observations in the present paper and explains them in a framework. She distinguishes between 'proper topic', 'aboutness topic' and 'theme of the sentence' (which are considered to be synonymous) as well as between anaphoric and stressless constituents called 'anaphoric topics'. The former dominate the latter. So, in order to make a proper topic, a constituent must be familiar in terms of information structure, and it must be predicated on by the rest of the sentence (including the anaphoric topic constituents). This is considered to be a rule of textual structure.

The constituents in the German middle field are ordered by seven constraints which, in turn, are defined by c-command relations. In the last part of the paper the constraints are included in an Optimality Theoretic (henceforth OT) ranking system.

When word formation is achieved in the lexicon and the lexicon is the input to each of the grammatical subsystems, **Rosemarie Lühr**'s paper on the accent pattern of Vedic compounds represents the starting point of the prosodic rules that lead to information structural focussing.

Rosemarie Lühr also uses an optimality based framework to clarify the accent positions of those endocentric and exocentric Vedic compounds that hitherto counted as counterexamples to known rules. But the combination of prosodic constraints with faithfulness constraints concerning the underlying stress of the compound as a whole [MAX (wf)] and of the members as single words [MAX $(w)_{In;}$ MAX $(w)_{Fin}$] was not recognized. Nor was the importance of prosodic constraints understood. Admittedly, there are many lexicalizations, but the following constraint is absolutely clear: The more a vowel which functions as the nucleus is unmarked, the more unmarked it is as the nucleus in stressed syllables, too. As demonstrated in the paper, in compounds this constraint concerns the avoidance of stressed *ú* and *í*. Yet, there are further prosodic constraints.

Prosodically determined stressing and those stresses which fulfill faithfulness constraints can be overridden by contrastive stress. Firstly, the well known stress shifts to differentiate parts of speech must be mentioned. Secondly, contrastive word stress appears within a compound as well.

Thomas Weskott's paper goes over from the more theoretical to the more empirical aspects. It makes use of a scalar notion of information structural markedness to generate empirical hypotheses to which experimental studies have direct access. The notion of markedness is understood in a naïve Jakobsonian way.

Information structural markedness is achieved by deviations from canonical orders: (i) The basic order of arguments is determined by the order in the argument structure of the verbal head; (ii) accent assignment is done in accordance with Cinque's null-hypothesis. A third type of deviation originates in the constraints which context puts on the information structure (topic-comment and background-focus structures) of the sentence. A simultaneous violation of these two orders and of the information structural constraints is additive and increases the degree of markedness. Thus, for the sentence *Der KELLner beleidigte den Gast* the order of background and focus is reversed, and the assumption about the default-accent placement is violated. *Den GAST beleidigte der Kellner* is marked threefold. The scale of markedness is put into a formal framework.

But the scale of markedness should not be considered merely as a descriptive generalization of intuitive judgements about the contextual adequacy of a sentence. It also serves as a basis for predicting processing difficulties of information structural variants relative to a given context. Several experimental studies show that German OVS structures are more difficult to process than canonical SVO structures as long as the context is rich enough.

Claudia Hruska and **Kai Alter** discuss the influence of prosody on sentence perception in online ERPs (event-related potentials). The result of their experiments on dialogues is as follows: As reflected by frontal effects, listeners pay more attention to the positions of expected new information. New information has to be integrated into already given context information maintained in working memory. Inappropriate accentuation in dialogues leads to disturbed comprehension, especially when focus accents are missed. Superfluous highlighting by accentuation, on the other hand, does not lead to a special brain reaction and seems to be ignored in some cases during speech perception.

When the same sentences were presented out of context, the system preferably used the prosodic cues of intonational phrase boundaries for an efficient evaluation of the underlying syntactic structure.

Ulrike Toepel and **Kai Alter** show that it is useful to split up the concept of narrow focus into (i) narrow new information focus and (ii) contrastive focus. New information accents and contrastive accents differ with respect to their acoustic properties in prominence and type. The acoustic analyses show that new information is accented with a bitonal L+H* pattern. For contrastive information, the pertinent pattern is more complicated: L+H*L-. The slightly

differing peak heights might play a role in the interpretation of the focus accents as well, since only in the contrastive condition are there no pre-focal accents. Hence the prominence of the contrastive accent is further enhanced by the surrounding prosodic pattern.

The behavioral and electrophysiological data provide evidence for the human brain's ability to detect these focus accents whenever they occur in an inadequate focus environment.

The authors show that not only the dichotomy of given vs. new information as well as that of broad vs. narrow focus is reflected in electrophysiological data, but also the more subtle difference between new and contrastive information. Prosodically underspecified (focus) accents in discourse generally lead to processing difficulties, whereas overspecified accents do so only under certain task requirements. This agrees with the results of Claudia Hruska and Kai Alter's paper.

The dissertation by **Grit Mehlhorn** is one of the first publications on modern Russian prosody. She provides the fundamental prosodic representations of Russian sentences with New Information Focus (NIF) and with Contrastive Focus in correction sentences as well as in Bridge Contours. In production as well as in perception experiments contrastive pitch accent was identified (in the respective contexts) by the subjects to nearly hundred percent, whereas the NIF was identified only by nearly 55 %. This finding proves that CF is a separate pitch accent with its own function in a Russian sentence. The paper describes the kind of the experiments and the relevant accent patterns. Similarities are found between Russian and German intonation as far as the accent types are concerned.

Britta Stolterfoht and **Markus Bader** focus on scrambling in German. They found two cases where focus structure is not fully determined by syntactic structure but also by context (cf. (1)) and by the type of the moved constituent (cf. (2)):

(1)   focus scrambling
        Maria hat behauptet daß [die Tante]F die Nichten tᵢ begrüßt haben.
(2)   scrambled pronominal objects co-referring to the subject of the main clause
        Maria hat behauptet, daß sieᵢ [die Nichten tᵢ begrüßt haben]F

Stolterfoht and Bader investigate on the scrambled referential DPs in ERP experiments and found effects of focus structural and prosodic revision in (1). But when focus particles are inserted as in (3)

(3)   *Maria hat behauptet, daß die Tanteᵢ nur [die Nichten]F tᵢ begrüßt haben.*

the focus structural effects disappear. A focus particle assigns focus to its adjacent constituent and makes a focus structural revision unnecessary. Further experiments will clarify whether the ERP effects reflect the interaction of syntactic and focus structural processes.

II. Contrast and correction

In a series of papers, **Ewald Lang**[1] deals with the role of contrast in grammar, specifically in terms of conditions between the various modules. As regards the relation of syntax and semantics to prosodically designated information structure, he observes that syntactically constant coordinate constructions like those in (1) – (6) below allow for a range of readings that (i) draw on whether or not the conjuncts display parallel structure in terms of prosody and Topic/Focus articulation, (ii) suggest interpretations, cf. (4) – (6), that come close to those overtly expressed by inherently asymmetric connectives like *obwohl/although, deshalb/therefore* etc., cf. (4') – (6'). Here is a briefly annotated summary of Lang's observations that have stimulated various follow-up studies.

(context): Was machen denn deine Eltern? How are your parents doing?

```
             L*H          H*L L%              L*H          H*L L%
(1)   [[[Mein Vater ↗]ᴛ [ist KRANK ↘ ]ꜰ ᴵᴾ] ᵁ] [Ø [meine MUtter ↗ ]ᴛ [geht ARbeiten ↘]ꜰᴵᴾ]ᵁ]
       My dad is ill↘                        my mom goes out to work\
       <pure contrastive reading, 1ˢᵗ & 2ⁿᵈ conjunct can be reordered
        without change in interpretation>


             L*H          L*H H%     L     L*H          H*L L%
(2)   [[[Mein Vater ↗]ᴛ [ist KRANK ↗ ]ꜰ ᴵᴾ] [und [meine MUtter ↗ ]ᴛ [geht ARbeiten ↘]ꜰᴵᴾ]ᵁ]
         My dad is ill↗                    and my mom goes out to work↘
         <connected contrastive reading. L*H pitch accent at the end of the 1ˢᵗ conjunct obligatory>
```

[1] Ewald Lang 2001: Kontrastiv vs. implikativ: Interpretationseffekte intonatorischer Distinktionen bei Koordination. In: A. Steube, C. Umbach (eds): Kontrast- lexikalisch, semantisch, intonatorisch. Linguistische Arbeitsberichte 77, Leipzig, 113 – 138.
Ewald Lang 2002: Kontrastiv vs. implikativ II: Intonationseffekte einer intonatorischen Distinktion bei elliptischen Sprichwörtern. In: A. Steube (ed): Sprachtheoretische Grundlagen der Kognitionswissenschaft: sprachliches und nichtsprachliches Wissen. Linguistische Arbeitsberichte 79, Leipzig, 187-212.
Ewald Lang (forthcoming): Contrasts in Grammar. Their roles, sources, and ways of realisation.

          L*H   L*H H%   L   L*H   H*L L%

(3)   [[[*Mein Vater* ↗]$_T$ [*ist KRANK* ↗]$_F$ ^{IP}] [*aber* [*meine MUtter* ↗]$_T$ [*geht ARbeiten* ↘]$_F$^{IP}]^U]
      My dad is ill   but my mom goes out to work
      <contrastive reading with added adversative evaluation of the conjuncts>

               L*H   H   H*   H*L L%

(4)   [*Mein Vater* [*ist KRANK* ↗]$_F$ [*aber meine MUtter geht ARbeiten* ↘]$_F$^{U}]
      My dad is ill   but (nevertheless) my mom goes out to work
      <implicative concessive reading, 2nd conjunct via *aber* in contrast to
      assumed inference from 1st conjunct>

(4) has an interpretation that comes close to that in (4') with an overt concessive connective like *obwohl* 'although':

(4')   Obwohl mein Vater krank ist, geht meine Mutter arbeiten.
      Although my dad is ill, my mom goes out to work'

               L*H   H   H*   ↑H*L L%

(5)   [*Mein Vater* [*ist KRANK* ↗]$_F$ [*und meine MUtter geht ARbeiten* ↘]$_F$^{U}]
      My dad is ill   and my mom goes out to work (= (5'))
      <implicative concessive reading, 2nd conjunct via ↑H*L on VP in contrast
      to assumed inference from 1st conjunct>

(5) has an interpretation that comes close to that in (5') with an overt concessive connective like *dennoch* 'nevertheless':

(5')   Mein Vater ist krank   Dennoch geht meine Mutter arbeiten
      My dad is ill. Nevertheless, my mom goes out to work.

               L*H   H   ↑H*L L%

(6)   [*Mein Vater* [*ist KRANK* ↗]$_F$ [*und meine MUtter* ↘ *geht arbeiten*]$_F$^{U}]
      My dad is ill   and my mom goes out to work
      <implicative causal reading, 2nd conjunct via ↑H*L on subject in contrast
      to assumed inference from 1st conjunct>

(6) has an interpretation that comes close to that in (6') with an overt causal connective like *deshalb* 'therefore':

(6')   Mein Vater ist krank (und) deshalb geht statt seiner meine Mutter arbeiten.
      My dad is ill / (and) therefore my mom goes out to work instead

Apart from the connectives, the variants (1) – (6) do not differ in their syntax and truth conditional semantics but do differ remarkably in prosodic patterning

and information structure. Following Lang, the major distinction between contrastive vs. implicative readings crucially rests on the prosodic and information structural symmetry vs. asymmetry of the conjuncts. The relevant distinctive features involved may be summarized as follows:

(1) – (3)
- Subjects marked by ↗ as contrastive Topics
- each conjunct forms an IP within the domain Utterance
- 2nd conjunct starts with low tone
- Topic-Focus articulation parallel in both conjuncts, i.e. $[[Subject]_T [VP]_F]$
- the meaning of 1st and 2nd conjunct interchangeably exemplify (K)

(3) – (6)
- Subjects are deaccented, form part of the Background
- no clearly marked internal IPs within the domain Utterance
- 2nd conjunct starts with high tone
- Topic-Focus articulation in the conjuncts non-parallel:
      narrow F vs. wide F in (4)
      narrow vs. emphatic contrastive F on VP in (5)
      narrow vs. emphatic contrastive F on Subject in (6)
- the meaning of 1st conjunct exemplifies (K), the meaning of the 2nd conjunct relates to an inference drawable from the 1st conjunct

Following Lang, it is the features of parallel structuring that are responsible for the "contrastive" reading in (1) – (3), whereas those of non-parallel structure induce "implicative" readings in (4) – (6), readings which are comparable to those otherwise induced by explicit asymmetric connectives like *obwohl/although, dennoch/nevertheless* or *deshalb/therefore*, cf. in (4') – (6'). What we observe here is a kind of compensatory trade-off between syntax/semantics and contextualized prosody and information structure. The connectives involved determine the way the conjuncts are related to (K) in terms of Discourse Linking. While ø , based on parallel-structured conjuncts, signals just exemplification (1), *und/and* cover a broader range of connections (2, 5, 6), depending on the prosodic make-up of the conjuncts. (3, 4) show the conditions on which *aber/but* may induce either adversativity (3) or concession (4).

Other researchers in the group, notably Carla Umbach, Marcela Adamíková, and Dorothee Fehrmann take up and modify Ewald Lang's approach. **Carla Umbach** herself[2] presents a different semantic explanation of *aber/but* which

---

[2] Cf. also Ewald Lang, Carla Umbach 2002: Kontrast in der Grammatik: spezifische Realisierungen und übergreifender Konnex. In: A. Steube (ed): Sprachtheoretische Grundlagen der Kognitionswissenschaft: Sprachliches und nichtsprachliches Wissen. Linguistische Arbeitsberichte 79, Leipzig, 145-186, and Ewald Lang 2004: Schnittstellen bei der Konnektoren-Beschreibung. In:

has also been made use of in this volume: (i) The contrast evoked by *aber/but* relates to the alternatives evoked by focus. (ii) An *aber/but*–sentence comprises a confirmation and a denial with respect to an explicit question referring to the alternatives under discussion. The meaning of *aber/but* can be characterized as both additive and excluding.

(7)   Adam: *Did John clean up his room and wash the dishes?*
(8)   Ben: *Bill cleaned up his room, but he didn't wash the dishes.*

There is no clear separation of contrast in the sense of adversativity and of concession. The concessive interpretation is induced by causal overinterpretation (which, due to the inherent negation of *aber/but,* results in incausativity, i.e. concession). (9) can be interpreted as (10).

(9)   *It is raining but Mary is happy.*
(10)   *It is raining but it is not the case that Mary is not happy because if that.*

Whereas in Lang's account prosodic symmetry of conjuncts is coupled with adversative and asymmetry with concessive readings, the account of Umbach 2002 and Umbach et al. in this volume sees concessive readings in all of these cases. Still, both Lang and Umbach predict that symmetric and asymmetric cases differ in intonation. **Carla Umbach, Ina Mleinek, Christine Lehmann, Thomas Weskott, Kai Alter, and Anita Steube** examine *aber/but* sentences along the lines of Lang in order to see whether his clear predictions concerning intonational patternings can be verified. Umbach et al. conduct a speech production experiment accompanied by acoustic analyses and perceptual evaluation. The acoustic analysis of the production study illustrate that in German a prosodic distinction between the adversative and the concessive versions in Lang's hypothesis could not be verified. The perceptive evaluation of the experiment did not give decisive insights either, so that a more sophisticated experimental design will be needed to test Lang's and Umbach's prosodic hypotheses.

   Marcela Adamíková and Dorothee Fehrmann have been concerned with the adversative conjunction *ale* (engl. *but*) in Polish, Czech and Slovak. In her paper, **Dorothee Fehrmann** comes to the conclusion that the Polish conjunction *ale* is lexically and syntactically underspecified and thus covers interpretations that in other languages are associated with distinct connectives. Thus, *ale* may either indicate CONTRAST (adversativity) like German *aber* (English *but*) without deletion of the non-focussed material, cf. (11a), or else – in con-

struction with negation – CORRECTION like German *sondern* or English *but* with obligatory deletion (cf. (11b)).

(11a): (Context suggesting a CONTRAST reading)
> A: *How are you getting to the station?*
>
> B: *Peter is taking me there.*
>
> A: *Really? But he hasn't got a car!*
>
> B: *Piotr nie ma samochodu, ale motocykl.*
> 'Peter hasn't got a car, but he's got a motorbike.'

(11b): (Context suggesting a CORRECTION reading)
> A: *Peter has a new Honda.*
>
> B: *Really? Peter's got a new car?*
>
> A: *Listen, dad.*
> *Piotr nie ma samochodu, ale motocykl.*
> 'Peter hasn't got a car, but a motorbike'.

The Polish data under review have a sentence negation in the first conjunct and an elliptic second conjunct. The aim of the paper was to look for prosodic features that might indicate the CONTRAST/CORRECTION distinction. Fehrmann's experiments so far did not show intonation contours that would clearly distinguish the two readings of *ale*. Instead, she found a default continuative intonation contour in Polish. Markers indicating IP boundaries between the two conjuncts were mostly missing. There are two pitch accents in the first conjunct. In the CONTRAST context the one on the XP was slightly more prominent than the one on [$_{Neg°}$ Neg° V°], whereas in the CORRECTION context the prominence marking was mostly the other way round. So far, neither the German intonation contours (with two separate conjunctions *aber, sondern,* (see Umbach et al.) nor the Polish ones show the expected parallelisms. But further research is needed in both languages.

More promising results concerning the problem illustrated by ambiguous data like (11) have been offered by **Marcela Adamíková** (2004). Based on production and perception experiments with native speakers of Slovak, Czech, and Polish, Adamíková investigates the full range of prosodic variations associated with data sets like (11). She takes into account the multifarious role of Negation (Foreground ~ vs. Background ~, focussing ~ vs. non-focussing ~ , sentence ~ vs. replacive ~) as well as the ways in which Focus assignment interacts with the scope of certain operators. Based on this, Adamíková singles out prosodic factors that actually distinguish CONTRAST vs. CORRECTION readings with data like (11) by showing that prosodic disambiguation takes place within the 1st conjunct, which contains the negation operator (normally prefixed to the finite verb). We illustrate two of the relevant patterns A, B by means of the annotated Slovak data in (12a – b):

(12)   (Context: Your husband plans to open a haulage firm. Does he at least have a car?)
                            !H* L%   H   (L+)H* L%
    a.   *Mój muž [ne-má]*F~ASS~ *auto* ↘. *Ale má [motorku]*F~ASS~
          'My husband does not have a car. But he has a motor cycle.'   < CONTRAST >

Pattern A
• non-focussing wide-scope negation
• low boundary tone (↘) at the end of 1st conjunct
• intonational break between the conjuncts

(12)   (Context: I am sure your husband has a car. – You are wrong.)
                            L+H*   L*H   H   (L+)H* L%
    b.   *Mój muž [ne-má]*F~ASS~ *[auto]*F~NEG~, *ale má [motorku]*F~ASS~
          'My husband doesn't have a car but a motor cycle'   < CORRECTION >

Pattern B
• focussing narrow scope negation
• high phrasal tone at the end of the 1st conjunct
• no intonational break between the conjuncts

Hence the prosodic differentiation of CONTRAST (12a) vs. CORRECTION (12b)
readings is shown to correlate with differences in negation scope, focus as-
signment, and falling or raising tones at the end of the 1st conjunct. Though
concerning a facet of adversativity that differs from the one illustrated by (1) –
(6) above, Adamíková's work is on par with Lang's approach in showing that
lacking lexical and/or syntactic distinctions can, to a certain extend, be com-
pensated by prosodic patterns.

English into a homogeneous form. Last but not least I want to thank the Publishing House Walter de Gruyter for initiating the series "Language, Context and Cognition" just at this time.

Leipzig, July 2004   Anita Steube

Books of the research group and of its associated scholars since 2000

Marcela Adamíková 2004: KONTRAST oder KORREKTUR? Prosodische Disambiguierung bei negationshaltigen Adversativkonstruktionen in den Westslavinen. Phil. Diss., Humboldt-University Berlin

Kai Alter 2002: Prosodie zwischen Sprachproduktion und –perzeption. Habilitation Thesis, University Leipzig

Irene Balles 2004: Die altindische Cvi-Konstruktion. Form, Funktion, Ursprung. Phil. Diss., Friedrich-Schiller University Jena

Nicole Dehé 2002: Particle Verbs in English. Syntax, Information Structure, and Intonation. In: Linguistik Aktuell /Linguistics Today 59. Amsterdam, Philadelphia: John Benjamins

Johannes Dölling, Thomas Pechmann (eds): Prosodie – Struktur – Interpretation. Linguistische Arbeitsberichte 74, Leipzig 2000, 306 p.

Johannes Dölling 2001: Systematische Bedeutungsvariationen: Semantische Form und kontextuelle Interpretation. Linguistische Arbeitsberichte 78, Leipzig, 252 p.

Johannes Dölling, Tatjana Zybatow (eds.): Ereignisstrukturen. Linguistische Arbeitsberichte 76. Leipzig 2001, 218 p.

Uwe Junghanns 2002: Untersuchungen zur Syntax und Informationsstruktur slawischer Deklarativsätze. Linguistische Arbeitsberichte 80, Leipzig, 193 p.

Grit Mehlhorn 2002: Kontrastierte Konstituenten im Russischen: Experimentelle Untersuchungen zur Informationsstruktur. Frankfurt/M.: Peter Lang

Sandra Muckel 2002: Wortstellungseffekte beim Satzverstehen: zur Rolle syntaktischer, lexikalischer und prosodischer Informationen. Wiesbaden: Deutscher Universitätsverlag

Andreas Späth 2004: Determination im Satzkontext. Grammatische Voraussetzungen der Nominalreferenz in den artikellosen slawischen Sprachen im Vergleich zum Deutschen. Habilitation Thesis, University of Leipzig

Anita Steube und Carla Umbach (eds): Kontrast – lexikalisch, semantisch, intonatorisch. Linguistische Arbeitsberichte 77, Leipzig 2001, 219 p.

Anita Steube (ed.): Sprachtheoretische Grundlagen der Kognitions-
    wissenschaft: Sprachliches und nichtsprachliches Wissen. Linguistische
    Arbeitsberichte 79, Leipzig 2002, 325 p.
Luka Szucsich 2002: Nominale Adverbiale im Russischen. Syntax, Semantik
    und Informationsstruktur. München: Otto Sager
Martin Trautwein 2003: The Time Window of Language: The Interaction
    between Linguistic and Non-linguistic Knowledge in the Temporal Inter-
    pretation of German and English Texts. Phil. Diss., University Leipzig
Valja Werkmann 2003: Objektklitika im Bulgarischen. studia grammatica 57
    Berlin: Akademie Verlag
Thomas Weskott 2002: The Left Periphery of German Verb-Second Sentences
    and its Interpretation in Context. Phil. Diss., Leipzig 2003
Tatjana Zybatow 2001: Grammatische Determinatoren von Zeit- und Sachver-
    haltsverlauf im Deutschen. Phil. Diss., University Leipzig

Eva Hajičová and Petr Sgall (Prague)

Degrees of Contrast and the Topic-Focus Articulation

1 Introduction

The aim of the present contribution is to present several hypotheses that may help to characterize and identify contrast as such and degrees of its intensity, and to point out how the phenomena connected with contrast can be handled in a descriptive framework that uses a dependency based syntax and includes a description of the topic-focus articulation (TFA) as one of the aspects of the underlying sentence structure. We also illustrate how the chosen descriptive framework is tested in a syntactic annotation of a large corpus. First we examine the chosen description of TFA (Section 1), then the means of expression of contrast are discussed (Section 2), as well as issues of contrast in focus (Section 3), and especially in topic (Section 4).

2 Topic-focus articulation

2.1 Topic, focus, and contextual boundness in a linguistic description

In the theoretical framework of the Praguian Functional Generative Description (FGD, see Sgall et al. 1986, Hajičová et al. 1998), the semantic basis of the articulation of the sentence into T(opic) and F(ocus) is understood as the relation of **aboutness**: a prototypical declarative sentence asserts that its F holds (or, with negation, does not hold) about its T. Thus, the core of the semantico-pragmatic interpretation of a declarative sentence might be based on a formula such as F(T) or, for a negative sentence, as non-F(T), if for the aim of the present discussion issues such as those of intension, lambda calculus and type theory are put aside. Within both T and F, an opposition of **contextually bound (CB)** and **non-bound (NB)** nodes is distinguished, which is understood as a grammatically patterned opposition, rather than in the literal sense of the term. In the underlying left-to-right order, NB dependents follow and CB dependents precede their heads.

In unmarked cases, the main verb (V) and those of its direct dependents that on the surface follow it belong to F, and the items preceding V are parts of T.

In marked (non-prototypical) cases, V can be CB, i.e. in T, or (a part of) F may precede V; usually the intonation centre (sentence stress) then marks F, occupying a marked position. The dependents of nouns primarily are NB.

Let us illustrate this view with a typical example (we understand the intonation center, in the prototypical case expressed by a falling pitch, to be placed at the end of the sentence; in other, marked positions it is denoted by capitals (which we use also in some other cases, to avoid a possible misunderstanding); let us recall that, in our underlying representations, the counterparts of function words are just indices of node labels, not occupying independent syntactic positions:

(1) My.t brother.t was visiting.t/f one.f of his.t friends.f yesterday.t.
       focus: (*was visiting*) *one of his friends* (intonation center on *friends*)

Here and in the sequel, t denotes a CB item, f denotes a NB one, and c is used to denote a contrastive CB item.

The verb in (1) is ambiguous in that it is NB (and thus a part of F) on one reading and CB (a part of T) on another; while the former is an appropriate "full" answer to (2), the latter answers (3).

(2) What was your brother DOING yesterday?

(3) Whom was your brother VISITING yesterday?

This view, the motivation of which has been published several times, makes it possible to analyze similar sentences (with an ambiguous part that may contain other words, not only verbs) with a **single opposition** of T and F. Thus the discrepancy between the single relationship of aboutness and two dichotomies assumed to constitute the information structure (e.g. by Junghanns and Zybatow 1997, ex. (2), p. 290) can be avoided and the T-F articulation (TFA) of the sentence can be assigned a specific position within the system of language (de Saussure's *langue*, Chomsky's *linguistic competence*), namely that of one of the basic aspects of the underlying, **tectogrammatical** representations of sentences (TRs). No separate level of information structure is needed.

The TRs contain no nonterminal symbols; each of their nodes is labelled by a complex symbol composed of a **lexical** and a **morphological** part (values of morphological categories such as number, tense, modalities), and each edge is labelled by the symbol indicating a **syntactic** relation (i.e. the type of the dependency relation).

2.2 TFA and contrast in a large corpus

The approach of FGD makes it possible to capture TFA and contrast in sentences of various degrees of complexity. The chosen descriptive framework, FGD, is being checked with examples taken from the **syntactically annotated Prague Dependency Treebank** (PDT), in which sentences from running text, from the Czech National Corpus (CNC, which contains hundreds of millions of word occurrences in journalistic fiction and other texts) are analyzed by a semi-automatic procedure.

In the PDT scenario, three layers of annotation are present, with TFA and contrast being represented (together with underlying dependency relations) on the underlying syntactic level. The resulting sentence representations have the form of tectogrammatical tree structures (TGTSs), with the following characteristic properties:

(a)   only autosemantic words are represented as separate nodes, with the exception of the coordinating conjunctions (in this point, TGTSs differ from the theoretically based TRs),

(b)   nodes deleted on the surface are restored,

(c)   the condition of projectivity is met (i.e. no discontinuity of sentence parts is allowed),

(d)   tectogrammatical functions ('functors'), i.e. kinds of the dependency relation such as (i) arguments: Actor/Bearer, Patient, Addressee, Origin, Effect and (ii) different kinds of adjuncts (temporal, local, condition, manner, etc.) are assigned as labels of the edges of the tree (or, equivalently, as indices in the labels of the dependents),

(e)   basic features of TFA are introduced (f, t, c, see Sect. 1.1 above).

Let us note that in the present experimental phase, 2000 sentences have been annotated in what concerns their underlying syntactic structure itself ('**large** collection'), with only 200 sentences having been annotated in full detail (the so-called '**model** collection'), and the annotations of 2000 sentences contain a treatment of TFA).

In the sequel, after a more general discussion of the phenomena of contrast, the checking of our descriptive framework on the material from PDT is illustrated by the Czech examples (30), (31), (33) and (34).

3 The means of expression of contrast

Several typical means of expression of contrast can be distinguished:
(i) E.g. in Czech, **strong pronominal forms** are used with certain pronouns; the typical cases of opposition of weak and strong forms are:

Czech   *ho – jeho* 'him-Gen,Acc', *mu – jemu* 'him-Dat',
         tě – tebe 'you-Gen,Acc', *ti – tobě* 'you-Dat',
         se – sebe 'Refl.Gen,Acc', *si – sobě* 'Refl.Dat'.

The strong forms are used to express NB pronominal forms, or CB contrastive forms as *ona* and *jeho* in (4)(a), respectively; they are also used in prepositional case forms and in coordination, cf. *tebe* in (5) and *tobě* in (6).

(4)   (a)   (Petr ji nazval konzervativcem.) Potom.t jeho.c urazila.t ona.f.
               (Petr called her a conservative). Then   him   insulted she.
                                        Then he was insulted by HER.

       (b)   (Petr ji nazval konzervativcem.) Potom.t ho.t opustila.f.
               (Petr called her a conservative.) The she LEFT him.

(5)   Na tebe jsem   se   celý   týden těšil.
              for you I-have-been Refl (the) whole week looking-forward
              I have been looking FORWARD to you for the whole week.

(6)   Tobě   nebo Martinovi to pošlu   zítra.
              to-you or   to-Martin it I-will-send tomorrow
              I will send it to you or to Martin tomorrow.

The corresponding weak ("short") forms are used only as CB, without contrast, cf. *ho* in (4)(b); it should be noted that Czech, a pro-drop language, has a zero form in the Nominative of all the personal pronouns, which occurs as their weak form (this is the case of the counterpart of *she* in (4)(b) or of *I* in (5) and (6)), although the „strong" forms *já, ty, on, my*, etc., may also occur without contrastive function, esp. in colloquial speech.

   In German, English and many other languages (and also in Czech with pronominal forms such as *je* 'them.Acc', *ji* 'her.Acc'), only an opposition of accented and unaccented forms is present as expressing that of contrastive (and NB) vs. non-contrastive use.

   Following up Koktová's (1999) observation that weak forms of pronouns in Czech cannot be used in certain positions in T, we use the opposition of strong and weak personal pronouns as an **operational test** for the contrastive use in T. Thus, *jeho* in (4)(a) is contrasted with *she*; there is no such contrast in (b). However, the application of this test is limited, since not only in coordination or with a preposition, but also when used as NB, in **focus**, is the pronominal form similar to that expressing a contrastive (part of) topic (marked with c), as is the case of *ona* in (4)(a), and also of *jeho* in (7)(b):

(7)   (a)   Jeho.c jsme viděli včera.f. – Him we saw yesterday.
       (b)   Včera.c jsme viděli jeho.f – Yesterday we saw HIM.

(ii) Rising stress (or, perhaps, falling-rising), having the form of L˙H, falls – perhaps optionally – on a contrastive (part of) topic in examples such as *jeho* in (7)(a) or *včera* in (7)(b); cf. also *jeho* in (4)(a). In the sequel we indicate such a "phrasal" or contrastive stress by italics. It would be interesting to check to what extent such examples can be characterized as bearing a **hat contour**, and under which conditions the hat contour can be taken as a criterion for contrastive T.

Steube (2001) examines similar examples in relation to the concept of I-topic, cf. also Jacobs (1997). Specific pragmatic properties accompany such accentuation at least in German (bound with specific illocutionary types and scope effects). If observations of this kind are valid for other languages as well, this would be significant for a further analysis of relationships between TFA, contrast, illocution and scopes of operators (now cf. also Umbach (2001). It would be important to also check examples such as the following, known from preceding dicussions:

(8)   An *American*.c farmer met a CANADIAN.f farmer.

However, it has to be noticed that such (or a similar) rising stress (marked by italics) is also used for an **open continuation** in various cases:

(a) This stress can occur in the middle of a longer sentence (without contrast, marking the ends of certain segments):

(9)   Naši mladší kolegové, kteří nedávno *dostudovali*,   dokončují   své   disertace.
        Our younger colleagues, who recently finished-studies, are-completing their   dissertations.

Perhaps this is the case also in coordinated clauses, e.g. in some languages, cf. the Slovak ex. given here as (10) (presented by Adamíková and Fehrmann 2001 without distinguishing the two kinds of accentuation, the presence of which we would assume):

(10)   Robert nie je HLÚpy, ale LEnivý. – Robert is not stupid, but lazy.

Slovak is far from isolated in such issues. Not only is the situation in Czech similar, but, e.g., also the English equivalent sentence can probably be pronounced as given in (11):

(11)   Robert is not *stupid*, but LAZY.

It would be difficult to speak of hat acccentuation in such examples. Rather, we would understand them as cases in which the two kinds of accent express the focus (an NB item), since it is often supposed (appropriately, as we are

convinced) that each of the coordinated clauses in a compound sentence has its own TFA. The NB value of the rising stress in a non-final clause under coordination might then be understood as being marked, non-prototypical.

(b) Rising stress also occurs at the end of the sentence, similarly as with a yes/no question:

(12)   Here are my *documents...* (... is there everything you are looking for in them?)

We differ from M. Steedman (2000), who has analyzed such sentences, in understanding this kind of stress (often indicated by „..." and having something in common with the intonation prototypically marked by a semicolon) as expressing **focus**, although a marked means of expression is used in such cases.

4 Contrast in focus

Focus as such has been characterized as a choice from **a set of alternatives,** esp. by M. Rooth (1985). This can only be understood so that F as such has a contrastive value. It might be claimed that if a sentence contains a single contrastive item, then this item constitutes the F. Only if more than one contrastive item is present, are they distributed among T and F. Thus, while a single contrast may be seen in sentences such as (13), both *Mike* and *Mary* are contrastive in (14).

(13)   Mike met Mary.

(14)   (Jim and Mike were looking for their classmates.)
       Mike met MARY.
       (*Jim* found ROBIN.)

This would mean that every F is contrastive. However, esp. in German studies, linguists often distinguish between contrastive and non-contrastive focus; perhaps it would be more exact to speak about two degrees of contrast, which are distinguished by the structure of German, as the following observations indicate:
     Sentences such as (15) are appropriate only for contexts which can be characterized by questions similar to (17), rather than by (18), i.e. in (15) *das Buch* is contrastive. On the other hand, (16) is acceptable both after (19) and (20), i.e. in (16) *Schüler* is either contrastive or not; cf. e.g. (in the framework of Optimality Theory) Hye-Won Choi (1996), discussed in Hajičová (2000).

(15) Hans hat das BUCH dem Schüler gegeben .
Lit.: Hans has the BOOK to the student given.

(16) Hans hat dem SCHÜLER das Buch gegeben .
Lit.: Hans has to the STUDENT the book given.

(17) Hat Hans dem Schüler das Buch oder das Bild gegeben?
Lit.: Has Hans to the student the book or the picture given?

(18) Was hat Hans dem Schüler gegeben?
Lit.: What has Hans to student given?

(19) Hat Hans das Buch dem Schüler oder dem Lehrer gegeben?
Lit.: Hans has the book to the student or to the teacher given?

(20) Wem hat Hans das Buch gegeben?
Lit.: Who has Hans the book given?

It seems that in Czech it is quite possible to use (15) as well as (16) as an answer to (17); this would mean that in Czech contrastive (part of) F does not constitute a specific grammatical value. It would than be interesting to investigate what the situation in English and in other languages is.

5 Contrast in topic (on a CB item)

Hajičová et al. (1998, 151) introduce the notion of **contrastive (part of) T** in connection with the occurrences of the so-called focusing particles in T, cf. (21):

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

To see how our criterion with the use of strong pronominal forms works, cf. a parallel Czech sentence with a corresponding noun of the masculine gender, (22). Note that the wording with the weak pronominal form is excluded: *I ho kritizoval Martin.

(22) (Kdo kritizoval i PAPEŽE jako nástroj kapitalismu?) I *jeho* kritizoval MARTIN. (Who criticized even the POPE as a tool of the capitalists?) MARTIN criticized even *him*.

The notion of contrastive T, however, should not be restricted to cases with focalizers, as we have seen with (4)(a), and as (23) confirms (with (23') as its simplified underlying representation).

(23) (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, SLOVAK is spoken.

(23')   Česky.*t* se-mluví.*t* v-Česku.*f*, na-Slovensku.*c* se-mluví.*t* slovensky.*f.*

Let us note that (23) is a compound sentence, in which each of the coordinated clauses exhibits its own TFA; *Česky*, as a part of the topic of the first conjunct, is contrasted as being chosen from the set of the two languages, and, in the second conjunct, *na-Slovensku* is contrasted with the focus part of the first conjunct.

It is possible to find two contrastive parts in T of a sentence:

(24)   (Rodiče odjeli na dovolenou a děti svěřili příbuzným.) *Syna*.c ve *čtvrtek*.c zavezli.f DO
         PARDUBIC.f, a *dceru*.c *v sobotu*.c DO HRADCE.f
  Lit.: (The parents left for a vacation and entrusted their children to their relatives.) Their
         son on *Thursday* they-brought to P. and their *daughter* on *Saturday* to H.

(25)   (Včerejší utkání bylo dramatické.) Po *gólu*.c SIEGLA.f se *Sparta*.c už.f ve třetí.f minutě.f
         dostala.f hladce.f DO VEDENÍ.f.
  Lit.: (Yesterday's match was dramatic.) After *goal* SIEGL'S *Sparta* already in the third minute
         got smoothly in the LEAD.

If the **degrees** of intensity of contrast are examined, it is possible to see that different dimensions are to be distinguished:

A. The **narrower** F is, the stronger the contrast:

The highest degree of contrast can be seen in the cases of **correction**, cf. Steube (2002). It may be asked whether correction differs from "second instance," as discussed for a long time especially in Czech linguistics. Another question is whether there are sentences occurring only as corrections; as was discussed already in Sgall et al. (1973, 36f), this may concern sentences with stressed items that cannot bear regular stress. However, at least in Czech, two degrees have to be distinguished:

(i) endings or affixes can only bear stress in corrections:

(26)   He carried out the analySES (not just one analysis).

(27)   Er sagt er hätte die Nachricht ERfasst, nicht VERfasst.

(ii) function words may bear stress not only in corrections, but also as constituting a narrow focus; thus, in (28), the tense value is in F, everything else belongs to T:

(28)   Er WIRD das tun.
       Lit.: He WILL do it.

Such an extremely **narrow F** can be seen as bearing a strong contrast.

B. The **set of alternatives** underlying the contrast in T is conditioned by various factors, both cotextual and contextual (situational). A scale (or a partial ordering) concerning the **explicitness** of the set of alternatives underlying the contrast, which appears to be weaker in case the set is not delimited explicitly.
     The set of alternatives is **explicitly** enumerated:

(29)   (Jakým jazykem se mluví v Rakousku a jakým na Moravě?) *V Rakousku*.c se mluví.t
       NĚMECKY.f, *na Moravě*.c se mluví.t ČESKY.f.
       (Which language is spoken in Austria and which in Moravia?)
       In Austria, GERMAN is spoken, *in Moravia*, CZECH is spoken.

The following relevant examples have been found in a set of sentence sequences from running texts in the Prague Dependency Treebank (PDT, see Sect. 1.2 above), examined by K. Veselá:

(30)   (Včera se hrál zápas mezi Brnem a Ostravou.) Ze začátku se dařilo.f DOMÁCÍM.f.
       Hosté.c se začali.f prosazovat.f až.f ve DRUHÉM.f POLOČASE.f. – (Yesterday, the match
       between B. and O. took place.) At the beginning, the LOCAL TEAM was successful. The
       guests started to succeed only in the SECOND HALF.

In (30), the set of alternatives is clearly delimited. However, the contrast is felt even stronger in case the sentences (clauses) are structured as parallel:

(31)   *Domácím*.c se dařilo.f ZE ZAČÁTKU.f. *Hostům*.c se povedl.f až.f DRUHÝ.f. POLOČAS.f.
       – *The local team* was successful AT THE BEGINNING. The *guests* succeeded only in the
       SECOND HALF.

In other cases, the alternatives are not quite clearly delimited, being determined just by the **set** being referred to:

(32)   (Terry has many friends.) My.t *brother*.c is.f one.f of his.t closest.f SCHOOLMATES.f.

There is also a possibility for the alternatives to stay **implicit**, just inferred from context. This can be illustrated by two subsequent sentences from PDT (from a newspaper article describing the feelings of the journalists when they saw and tested a new type of Toyota car, which was supposed to be suitable both for driving on roads and on the terrain):

(33) Už.f   první.c pohled.t na atypickou.f karosérii.f potvrzuje.f, že se   jim.t jejich.t
  Lit. Already first   glance   at atypical   body   confirms   that Refl. them their
      záměr.t   podařilo.f naplnit.f.
      intention managed to-accomplish

Already the first glance at the atypical body confirms that they managed to accomplish their intention.

(34)   Pro pohon.c byl zvolen.f dvoulitrový.f motor.f osvědčený.f v   Toyotě.f. Cari.f. E.f ...
  Lit. For drive   was chosen two-liter   engine well-tried   with Toyota Cari E ...

As for the drive, a two-liter engine was chosen well-tried with Toyota Cari E. In (33), *první* 'first' is chosen among different possible steps of observation (more or less thorough); in (34), *pohon* 'drive' is chosen among the attributes of the car.

C. A third dimension concerns the **range** of the set: it may be a (i) **pair** – cf. (4), (23), (29), (30), (31) above, – which is connected with a relatively strong contrast, or (ii) **larger** – cf. (32) above, – with a weaker contrast, or (iii) it has a **single** member, as is the case in the following cases:

(i) the contrastive item is coreferential with an item of the F of the preceding utterance, rather than with one of its T, i.e. a "new T" is present; it seems that the opposition between "new" and "old" T (i.e. between a (part of the) T that has not occurred in the T in the preceding co-text and a (part of the) T that has, respectively, comes close to the opposition of T proper and temporal or local setting):

(35)   (Kde se mluví česky?) Česky.c se mluví.t v Česku.f.
      (Where is Czech spoken?) Czech is spoken in Czechia.

(ii) cases with a focusing particle in T, as in (21) above,

(iii) the contrast is being newly established, as in (36), in which *já* 'T' is presented as being in contrast to other individuals.

(36) Přiznám.f se, že já.c osobně.f to.t dost.f prožívám.f.
  Lit.: I-admit   that I personally it   quite live-through.
      I admit that I personally live through this quite intensely.

To be more exact, we should note that in example (36) the contrastive item is not in T, but, rather, it is a **CB item in F**. Typically, CB items stand in T while NB ones are in F; however, elements deeply embedded (i.e. dependent on an

item that differs from the main verb) may occur as NB (contrastive or not) items within T, or as CB items in F of the whole sentence. In (36), the subject of the main clause, having a zero form, is CB and constitutes the T (the values of its grammatemes are expressed, on the morphemic level, by the agreeing personal ending of the verb). The verb together with the embedded clause constitute the F. The subject of this clause, expressed by the pronoun in its strong form, is a contrastive CB item, and together with the CB pronoun *to* 'it' it belongs to the F, since both the pronouns depend on an item in Focus different from the main verb (namely on the embedded verb).

If the patterning of a discourse is examined taking into account the TFA of the subsequent sentences, then the **prototypical** case may be found in those sequences of two sentences S_0 and S_1 in which the T of S_1 is referentially identical to T of S_0, and the F of S_1 is chosen among the alternatives of what can be asserted about T; cf. Weil's (1844) *la marche parallèle,* and Daneš's (1974) first alternative of 'thematic progression', $T_2 = T_1$.

In **marked** cases, there are the following possibilities for the choice of T_1 :

(a) associative relations with accommodation are present, rather than the referential identity of T_0 and T_1,

(b) T_2 is coreferential with F1, rather than to T1 (Weil's *progression,* Daneš's **$T_2 = F_1$,**

(c) T_1 is chosen from another part of the set of established items than from those referred to in S_0 or from those associated with these referents.

While case (a) is directly related to the protoypical situation, in cases (b) and (c) T_1 is **chosen** from a set of alternatives, i.e. a may be seen as a contrastive item. Since T may include more than one item, it is more precise to speak of **CB items.**

Thus, often also (a part of the) topic can be considered as a choice from a set of alternatives (cf. Steedman's 2002 'theme alternative set', with 'theme' marked by the L+H* pitch accent, cf. R. Jackendoff's 'B contour'). It should be noted that in a compound sentence the focus stress (intonation centre) in the non-final coordinated clause(s) primarily is rising, rather than falling. A similar kind of pitch probably can be observed in the middle of longer sentences of other types, without expressing focus or even contrast, see Sect. 1.2.(ii) above.

Other interesting examples, known from older discussions without such an interpretation, were analyzed as containing a contrastive (part of) T by Hajičová et al. (1998, 155-157):

(37)   Farmers.t that.t grow.c rice.t often.t only.f eat.f rice.c.

Here the focusing particle *only* is connected with the CB occurrence of *rice* at the end of the sentence, and a hat contour (rising pitch on *grow*) is present.

(38)   (Niemand liest Goethes Gedichte heute.)

    Sogar.f Peter.f kennt.t nur.f einen Roman.c von Goethe.t.

    Lit.: (Nobody reads Goethe's poems nowadays.) Even Peter knows only a novel by Goethe.

6 Conclusion

The view presented and illustrated in the present paper makes it possible to analyze the information structure of sentences with the use of a **single opposition** of T and F, if also the difference between contextually bound and non-bound items is observed, as well as that between contrastive and non-contrastive items in T. Thus the discrepancy between the single relationship of **aboutness** and two dichotomies often assumed to constitute the information structure can be avoided and the T-F articulation of the sentence can be assigned a specific position within the system of language, namely that of one of the basic aspects of the underlying, **tectogrammatical** representations of sentences. No separate level of information structure is needed.

7 References

Adamíková, M. and D. Fehrmann (2001): Prosodische Differenzierung von Kontrast- und Korrekturkonstruktionen im Slowakischen und Polnischen. Linguistische Arbeitsberichte 77, Leipzig, 81-112.

Daneš, F. (1974): Functional sentence perspective and the organization of the text. In: F. Daneš (ed): Papers on Functional sentence perspective. Prague: Academia (1974), 106-128.

Hajičová, E. (2000): Teorie optimality a aktuální členění věty [Optimality theory and the topic-focus articulation]. Slovo a slovesnost 61, 161-169.

Hajičová, E., B. H. Partee, and P. Sgall (1998): Topic-focus articulation, tripartite structures, and semantic content. Dordrecht: Kluwer.

Hye-Won, C. (1996): Optimizing structure in context: Scrambling and information structure. PhD Thesis. Stanford University.

Jacobs, J. (1997): I-Topikalisierung. Linguistische Berichte 168, 91-133.

Junghanns, U. and G. Zybatow (1997): Syntax and information structure of Russian clauses. In: E. W. Browne, E. Dornisch, N. Kondrashova, and D. Zec (eds): Annual Workshop on Formal Approaches to Slavic Linguistics. The Cornell Meeting 1995. Ann Arbor, Mi : Michigan Slavic Publications.

Koktová, E. (1999): Word-order based grammar. Berlin/New York: Mouton de Gruyter. Pražský závislostní korpus [Prague Dependency Treebank] (2001): http://ufal.mff.cuni.cz/pdt .

Rooth, M. (1985): Association with focus. PhD. dissertation. University of Massachusetts, Amherst.

Sgall, P., E. Hajičová, and E. Benešová (1973): Topic, focus and Generative semantics. Kronberg/Taunus:Scriptor.

Sgall, P., E. Hajičová, and J. Panevová (1986): In: J. L. Mey (ed): The Meaning of the Sentence in its Semantic and Pragmatic Aspects, Dordrecht: Reidel - Prague: Academia.

Steedman, M. (2000): Information structure and the syntax-phonology interface. Linguistic Inquiry 31, 649-689.

Steedman, M. (2002): The surface grammar of intonation and information structure. Paper delivered at the workshop on Contrast and Information Structure, Leipzig, February 2002.

Steube, A. (2001): Grammatik und Pragmatik von Hutkonturen. Linguistische Arbeitsberichte 77, Leipzig, 7-30.

Steube, A. (2002): Correction by contrastive focus. In: Theoretical Linguistics 27, 2-3, 215-49.

Umbach, C. (2001): Restriktion der Alternativen. Linguistische Arbeitsberichte 77, Leipzig, 165-198.

Weil H. (1844): De l'ordre des mots dans les langues anciennes comparées aux langues modernes. Paris. Translated as The order of words in the ancient languages compared with that of the modern languages. Boston 1887, reedited in Amsterdam: Benjamins 1978.

Anita Steube, Kai Alter, and Andreas Späth (Leipzig)

Information Structure and Modular Grammar

1 Information Structure and its pragmatic categories

Isolated grammatically correct sentences do not fulfill the pragmatic functions that sentences in coherent texts of certain types and subtypes do. It is information structure that selects the grammatical forms and constructions that embed sentences in a text. As far as known nowadays this is achieved by two pairs of pragmatic categories: the background–focus~ and the topic–comment partitions. By means of the background–focus partition, a sequence of sentences becomes coherent by the phenomenon that the subsequent sentences are anaphorically enchained with the ones preceding them. By means of the topic–comment partition, longer texts are subdivided into paragraphs each dealing with one object or event. At the same time, the realization of topics can e.g. characterize the subtype of a narrative text, such as a tale versus a news report.

All four pragmatic categories must be realized by grammatical means in order to get expressed. However, languages also differ in the grammatical level and structure of expression as a function of language specificity. It is thus not enough to give exclusively pragmatic definitions of the information structural categories. We must also know how the corresponding language maps them on its grammatical forms and their combinations. Hypothesizing that pragmatic definitions are universally comparable, their mappings on the grammatical forms and structures will not only differ from language to language but will even depend on the theory of grammar used.

Background Constituents express familiar information belonging to the common ground of both the speaker and the hearer. It has either been verbalized immediately before, can be derived from the communicative situation, is part of the common knowledge of the communication partners, or can be inferred from different kinds of holistic linguistic and/or non-linguistic knowledge. Therefore, Background Constituents can be said to be anchored in context. They are represented by the marking [-F(okus)]. One sentence may contain one or more Background Constituents. But no pragmatically expressive sentence may consist of Background Constituents only.

(1) *[Es war einmal ein alter KÖnig]*ₐF

    There once was an old king.

    *Der [hatte drei SÖHne]*ₐF

    [-F]

    He had three sons.

    *Ihre Burgen [standen entlang einer wichtigen HANdelsstrasse]*ₐF[1]

      [-F]

    Their castles stood along an important trade route.

Focus Constituents express non-familiar information which, correspondingly, is new information for the hearer, has not been verbalized before in the communicative situation, and is not givenby the context. Thus, when it is expressed, it is communicatively new or important.

Focus constituents are represented by the marking [+F]. A sentence may contain one or more focus constituents, or it may consist of focused information only. Sentences of the latter type are called fully focused sentences. Short fully focused sentences of mostly one constituent besides the finite verb are called thetic sentences. The focus constituents form the focus domain which, correspondingly, can be a minimal domain, a middle wide, or a maximal domain.

Focus constituents can be subdivided into Representative Focus (also called New Information Focus) and Contrastive Focus. Contrastive Focus partly shares the pragmatic characteristics of Representational Focus, but is used in sentences with additional communicative functions: in Correction Sentences and in the so called Bridge Contours (cf. § 4). We will represent contrastive focus by [CF]. The focus domain of contrastive focus may have all the degrees of width that New Information Focus has, but it tends to have narrow focus.[2]

Fully Focused Sentences are normally introductory sentences (cf. the first sentence in (1)). They are used when no common ground has yet been established between the communicative partners. The partition between Background and Focus in a sentence can be decided upon by means of context only. In German, the default sequence of Information Structure is: Background Constituents before Focus Constituents.

In the literature, we very often find the name "Background Constituent" replaced by Topic. Other authors differentiate between Familiarity Topics (designating our Background Constituents) and Aboutness Topics (or Topics proper). In this paper, only Aboutness Topics are called Topics. One sentence thus has got one topic. This topic expresses what the rest of the sentence (i.e.

---

[1] The pronoun *die* and the DP *ihre Burgen* are Background Constituents. Whereas the pronoun is lexically anchored, the DP must be inferred by means of world knowledge.

[2] The focus domain is bracketed by indexed angled brackets [...]ₐF. The position of the focus accent in the focus domain is represented by CAPITALS.

the Comment) is about. In a descriptive text, the chain of Topics is built by the elements of the set constituting the Theme (i.e. Topic of the text), (cf. (2) with its Theme *Thüringen*). Topics were thought to always be a subset of Background Constituents for a long time. (In (2), the area, the number of inhabitants, and the capital belong to the political elements constituting a country. These elements belong to the common ground.) But when texts were inspected more closely with respect to their substructure, focused and even contrastively focused constituents had to be analyzed as Topics as well (cf. (3)).

(2) $\underline{\text{Thüringen}}$ ist_1 [seit 1990 ein neues Bundesland der Bundesrepublik DEUTSCHland $t_1]_F$

$\underline{\text{Die Bodenfläche des Freistaates}}$ beträgt$_1$ [16 251 QuadRATkilometer $t_1]_F$

$\underline{\text{Seine Bevölkerungszahl}}$ ist t_1 [2,7 Millionen EINwohner $t_1]_F$

$\underline{\text{Die Landeshauptstadt}}$ ist$_1$ [Erfurt $t_1]_F$ [3]

Thuringia has been a new federal state of the Federal Republic of Germany since 1990.

The area of the Freestate is 16 251 million square kilometres.

Its population is 2,7 million inhabitants.

The federal capital is Erfurt.

(3) *Aldi unterm Zeltdach*

a) Vor zwei Monaten$_3$ begann$_1$ Aldi$_2$ [t_3 t_2 seine Verhandlungen mit dem Rat der Stadt LEIPzig $t_1]_F$ [-F]

b) [*Ein GeLÄNde*]$_F$ wollte der Konzern in Leipzig pachten.

c) [*Ein GROSSzelt*]$_F$ wurde auf dem Gelände errichtet.

d) [$\underline{Voll\ von\ WAren}$]$_F$ ist der Supermarkt, aber [sehr primitiv EINgerichtet]$_F$ ist er.

Aldi under a tent

Two months ago Aldi began negotiations with the Council of The Town of Leipzig.

It was an area which the trust wanted to rent.

A big tent was erected on the lot.

Full of goods is the super-market is but it is very primitively furnished.

(3) is a news report. *Aldi unterm Zeltdach* constitutes the Theme. In all sentences of the text the Focus is placed before the Background. This is untypical for tales but typical for news stories which tend to mention what is new as soon as possible. Nevertheless, sentence b) speaks about an area, sentence c) about a tent, and sentence d) speaks about the make-up of the supermarket twice. It is the focused Topics here which constitute the internal structure of the text. Extending the notion of Topic to focused constituents, we can include even the so called I-Topics of Bridge Contours in the notion of Topic. These are also used to promote text (3):

---

[3] The Topics are underlined.