

Basic Notions of Information Structure

ESSLLI 2014 Annotating Corpora with Information Structure

Kordula De Kuthy and Arndt Riester

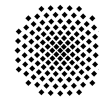
August 18, 2014



Motivation

Why do we want to annotate corpora with information structure?

- ▶ Notions like *information structure*, *focus and/or topic* have been intensively investigated in the theoretical literature.
 - ▶ These theories mostly rely on intuitions about hand-crafted sentences.
 - ▶ Furthermore, these theories get more and more complex and it is not always obvious how they can be operationalized.
- It is important to test these theories with naturally occurring data.



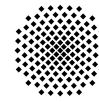
Motivation (cont.)

Information Structure in NLP Applications

- ▶ A lot of NLP applications rely on a very crude notion of discourse and anaphora resolutions.
- ▶ These applications could benefit from a more elaborate notion of discourse and how discourse relations can be identified in naturally occurring sentences.

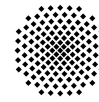
The Cross-linguistic Perspective

- ▶ Universal pragmatic categories are useful for comparing the realization of information structure across languages.
- ▶ Vice versa, the empirical investigation of discourse properties of various languages is an important step towards identifying such universal pragmatic categories.



Background on Information Structure

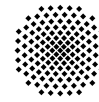
- ▶ Introduction: What is information structure and basic notions
- ▶ Historical development of information structure approaches (largely based on von Stechow 1999, ch. 3)
 - ▶ The Beginnings of Information Structure
 - ▶ The Prague School
 - ▶ Halliday and the American structuralism
 - ▶ Information Structure in Generative Grammar
- ▶ The Semantics of Information Structure
 - ▶ Structured Meaning
 - ▶ Alternative Semantics
- ▶ Intonation and Information Structure
- ▶ Word order and Information Structure



There is more than just syntax and semantics

A simple sentence (1) can be used in many different contexts (2–4), conveying different kinds of information.

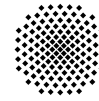
- (1) Tim bought a new car.
- (2) a. There is a brand-new Mercedes outside. Did anybody buy a new car?
b. TIM bought a new car.
- (3) a. Tim looks so happy these days. What did he do?
b. Tim bought a new CAR.
- (4) a. What did Tim do after his old car broke down? Did he lease a new car?
b. No, Tim BOUGHT a new car.



Prosodic mismatches

- (5) a. What did Fred eat?
b. Fred ate BEANS.
c. # FRED ate beans.
- (6) Eric is so stupid.
a. He doesn't even know the capital of FRANCE.
b. # HE doesn't even know the capital of FRANCE.
c. ?He DOESN'T even know the capital of France.
d. He doesn't EVEN know the capital of France.

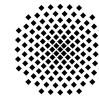
Funny situations can arise from wrong assumptions about the prosody of written language.



What is information structure?

Beaver & Clark (2008)

(7) David only wears a bowtie when teaching.

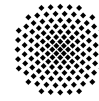


What is information structure?

Beaver & Clark (2008)

(7) David only wears a bowtie when teaching.

(8) David only wears a bowtie when TEAching. (*He takes it very seriously.*)



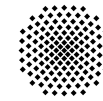
What is information structure?

Beaver & Clark (2008)

(7) David only wears a bowtie when teaching.

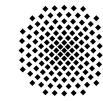
(8) David only wears a bowtie when TEAching. (*He takes it very seriously.*)

(9) David only wears a BOWtie when teaching. (*Well...*)



What is information structure?

- ▶ Very generally speaking, the information structure encodes which part of an utterance is informative in which way, in relation to a particular context.
- ▶ Information structure generally deals with the organisation (“packaging”) of propositional information, in order to make an utterance fit its current context.
- ▶ Information structure spans several linguistic areas: pragmatics, semantics, syntax, phonetics (in particular prosody)
- ▶ Some people would say that information structure is itself a linguistic domain.
- ▶ A wide range of approaches exists with respect to the question what should be regarded as the primitives of the information structure, with diverse and often confusing terminology.



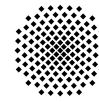
Two primitives of information structure

Many approaches include one or both of the following distinctions:

- ▶ *Givenness*: A distinction between
 - ▶ what is new information advancing the discourse (*focus*)
 - ▶ what is known, i.e., anchoring the sentence in existing (or presupposed) knowledge or discourse (*background*)
- ▶ *Aboutness*: A distinction between
 - ▶ what the utterance is about (*topic, theme*)
 - ▶ what the speaker has to say about it (*comment, rheme*).

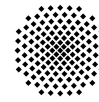
Example: (10) a. What does John drink?

	<i>background</i>	<i>focus</i>
b.	John	drinks BEER.
	<i>topic</i>	<i>comment</i>



The Focus/Background distinction

- ▶ A sentence can be structured into two units according to their informativeness, i.e., which part is informative (new) with respect to the discourse, the *focus*; and which part is uninformative (known), the *background*.
- ▶ The typical test for the focus unit of a sentence is the constituent question:
 - (11) a. Q: Who did Sue introduce to Bill?
A: Sue introduced [JOHN]_F to Bill.
 - b. Q: Who did Sue introduce to Bill?
A: Sue introduced [the woman with the red SCARF]_F to Bill.
 - c. Q: What happened?
A: [Sue introduced John to BILL]_F



The Focus/Background distinction (cont.)

- ▶ Linguistic means of marking such an information structuring are, for example, word order, morphology and prosody.
- ▶ English and German are so called intonation languages, i.e., they use pitch accents to highlight informational units of the utterance in a particular way.
- ▶ The intonationally highlighted part is associated with the most informative part, i.e., the focus, while the remainder of the sentence contains mainly background knowledge, i.e., information that is already available in the discourse.



The Focus/Background distinction: Contrast

(12) Peter hit the iceman.

Different corrections:

- (13) a. Nonsense, FRED_F hit the iceman.
b. Bullshit, Peter hit [the poLICEman]_F .
c. Rubbsish, Peter BIT_F the iceman.

Again, the focussed expressions convey new information. In addition to that, the **contrast** with their counterparts in (12).

Double correction:

- (14) No [FRED]_F hit [the poLICEman]_F .



Contrast without novelty

- (15) a. Peter hit the iceman.
b. Nonsense, [the ICEman]_F hit [PETER]_F .

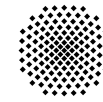
In (15b) all words are **given/old** (apart from *nonsense*). However, they contrast with their counterpart in (15a).

Selection:

- (16) a. I am very fond of my two aunts. Their names are Karolina and Elfriede.
b. ELFRIEDE_F I like BEST_F .

Best: new

Elfriede: given, but used in order to make a choice



The Topic/Comment distinction

- ▶ In the topic-comment structure, topic refers to what the utterance is about and comment what the speaker says about it.
- ▶ The topical element can be associated with the question: *What about X?*
- ▶ In English, topic is marked by a pitch accent, just like focus is, but of a different kind: The focus accent is a typical falling movement while the topic accent is realized as a fall-rise.

(17) Q: Well, what about FRED? What did HE eat?

A: FRED ate the BEANS.

topic

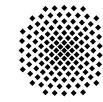
focus

(18) Q: Well, what about the BEANS? Who ate THEM?

A: FRED ate the BEANS.

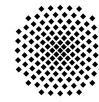
focus

topic



Semantic effects





Semantic effects

A sign in the London underground reads (Halliday 1967):

(19) Dogs must be carried.

This sentence can be read in two different ways:

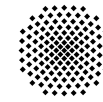
(20) a. Dogs must be CARried.

b. DOGS must be carried.

There is a difference in meaning:

(21) a. If you have a dog, you must carry it.

b. What you must do is carry a dog.



Semantic effects (cont.)

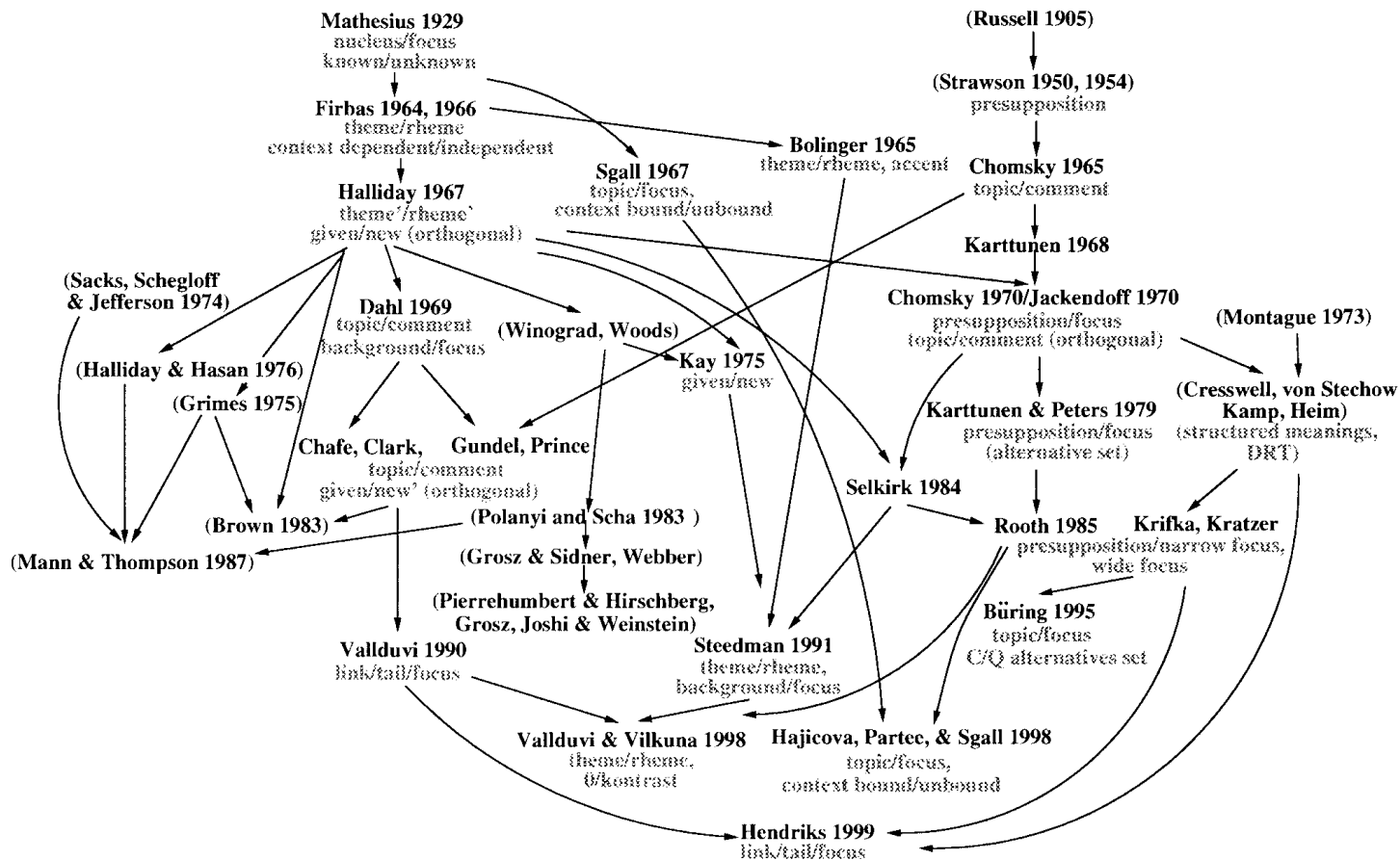
The reading in (21b) is odd , but it is the preferred one for:

(22) Shoes must be worn.

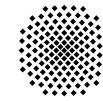




Historical development of information structure approaches

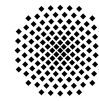


(from Kruijff-Korbayová & Steedman 2003)



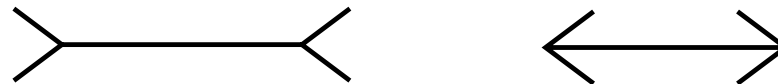
The beginnings of information structure

- ▶ In the 19th century it became obvious that the grammatical description of the sentence does not cover all aspects of sentence meaning. Differences in the presentation of the sentence content were attributed to an underlying psychological structure.
- ▶ In psychology, the so-called *Gestalt theory*, assumed that perception functions as a whole gestalt and not by constructing something out of small units.
- ▶ The gestalt perception includes two different parts: *figure* and *ground*.

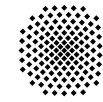


Gestalt and language

- ▶ The figure is recognized only against the ground. This is the principle behind many optical illusions, where one and the same stimulus (the line) is perceived differently depending on the ground.

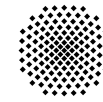


- ▶ Related to the Gestalt theory in psychology, the idea of a dichotomy of the sentence organization was developed, which inherited the terms figure and ground.
- ▶ The figure represents the prominent or highlighted part, while the ground represents the given or less informative material of the sentence.



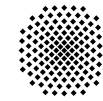
The communicative function of language

- ▶ At the beginning of the 20th century, the interest in the communicative function of language increased.
- ▶ In order to distinguish between the grammatical structure of the sentence (subject–predicate), the psychological structure of concepts or ideas, and the informational structure of a message in a communication, Ammann (1928) introduces a new pair of terms for the latter: *theme* and *rheme*.
- ▶ The Prague School integrated the distinction between theme and rheme into the grammatical system.



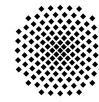
The Prague School

- ▶ The most characteristic feature of the Prague structuralists, in contrast to other structuralist schools, is the functional perspective:
 - ▶ Language is understood as a tool for communication and the information structure is important for both the system of language and the process of communication.
- ▶ Firbas (1964) argues that information structure is not a dichotomy but rather a whole scale, or hierarchy, or what he calls *communicative dynamism*.
- ▶ The newer Prague School (cf., e.g., Sgall et al. 1973, 1986) derive the *topic-focus articulation* from a notion of contextual-boundedness and make it part of the grammatical model of a sentence.



Halliday and the American structuralists

- ▶ Halliday (1967) introduced the Praguean distinction of theme and rheme into American structuralist linguistics.
- ▶ He is the first who uses the term *information structure* and establishes an independent concept of it. He assumes that an utterance is organized into “information units”, which do not correspond to constituent structure.
- ▶ Analogously, Halliday assumes two structural aspects of information structure:
 - ▶ the informational partition of the utterance, the *thematic structure* (theme-rheme), organizes the linear ordering of the informational units.
 - ▶ the internal organization of each informational unit, the *givenness*, elements are marked with respect to their discourse anchoring.
 - ▶ the center of an information unit is the *information focus*, which contains new material



Information Structure in Generative Grammar

- ▶ Chomsky (1971) assumes a *focus/presupposition* distinction.
- ▶ The function of focus is to determine the relation of the utterance to responses, to utterances to which it is a possible response, and to other sentences in the discourse.
 - ▶ *Focus* is defined as the phrase containing the intonation center.
 - ▶ *Presupposition* is described as that part of the sentence that is conveyed independently of the speech act or a negation made in the sentence.



Information Structure in Generative Grammar (cont.)

- ▶ Each sentence in (23) presupposes that John writes poetry somewhere.

- (23) a. John writes poetry in his STUdy.
b. Does John write poetry in his STUdy?
c. John doesn't write poetry in his STUdy.

Each can be followed by:

(24) No, John writes poetry in the [GARden]_{Focus}.

- ▶ On this basis, Jackendoff (1972) developed an approach which is the basis for a number of semantic theories of focus.



Semantic theories of focus

- ▶ In the wake of Jackendoff (1972), formal theories of the semantics of focus associate with each sentence a model-theoretic entity which directly reflects its focal structure. This entity is often called the *focus-induced interpretation*.
 - ▶ The value of the focus, i.e., the ordinary denotation of the focused expression, is part of the *set of alternatives*, the *p(resuppositional)-set*.
 - ▶ The rest of the sentence corresponds to a semantic structure that is called *p-skeleton*. It is formed by substituting the focused expressions with appropriate variables, for example:

- (25) a. John introduced [BILL]_F to Sue.
 p-skeleton: John introduced x to Sue.
- b. John introduced Bill to [SUE]_F .
 p-skeleton: John introduced Bill to y.



The semantics of focus: Structured meaning

- ▶ The structured meaning theory of focus was developed by Stechow (1981), Stechow & Cresswell (1983), Jacobs (1983), and Krifka (1992).
- ▶ The focus-induced interpretation of a sentence is an ordered sequence, the *structured meaning*, whose members are
 - ▶ the property obtained by λ -abstracting on the focus (or foci), and
 - ▶ the ordinary semantic interpretation of the focus (or foci).

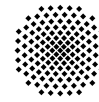
As an example, consider the structured meaning representation of the examples repeated in (25):

(25) a. John introduced [BILL]_F to Sue.

$\langle \lambda x [\textit{introduce}(\textit{john}', x, \textit{sue}')] , \textit{bill}' \rangle$

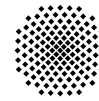
b. John introduced Bill to [SUE]_F .

$\langle \lambda y [\textit{introduce}(\textit{john}', \textit{bill}', y)] , \textit{sue}' \rangle$



The semantics of focus: Alternative semantics

- ▶ The alternative semantics theory of focus was proposed in Rooth (1985).
- ▶ Each sentence receives two distinct model-theoretic interpretations:
 - ▶ an ordinary semantic value (written as $\llbracket \cdot \rrbracket^o$), and
 - ▶ a separate focus-induced interpretation called the p-set or *focus-semantic value* (written as $\llbracket \cdot \rrbracket^f$), which is the set of all propositions obtainable by replacing each focus with an alternative of the same type.

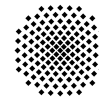


The semantics of focus: Alternative semantics (cont.)

- ▶ The focus semantic value of (25a), i.e., $\llbracket \text{John introduced [BILL]}_F \text{ to Sue.} \rrbracket^f$ is shown in (26a).
- ▶ In (26b), it is spelled out assuming that the only individuals in D are John, Bill, Sue, and Mary.

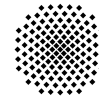
(26) a. {the proposition that John introduced d to Sue : $d \in D$ }

b. $\left\{ \begin{array}{l} \llbracket \text{John introduced John to Sue} \rrbracket^o, \\ \llbracket \text{John introduced Bill to Sue} \rrbracket^o, \\ \llbracket \text{John introduced Sue to Sue} \rrbracket^o, \\ \llbracket \text{John introduced Mary to Sue} \rrbracket^o \end{array} \right\}$



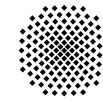
Expressing information structure

- ▶ Languages differ with respect to how the information structure of an utterance is represented.



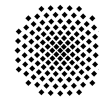
Expressing information structure

- ▶ Languages differ with respect to how the information structure of an utterance is represented.
- ▶ Linguistic means of marking information structure include:
 - ▶ word order
 - ▶ morphology
 - ▶ prosody



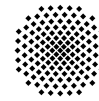
Expressing information structure

- ▶ Languages differ with respect to how the information structure of an utterance is represented.
- ▶ Linguistic means of marking information structure include:
 - ▶ word order
 - ▶ morphology
 - ▶ prosody
- ▶ English and German are so-called intonation languages
 - ▶ Information structuring is signaled by the intonation (contour) of an utterance, including pitch accents.
 - ▶ The absence or presence of an accent is an indicator of the discourse function of a constituent in a sentence.



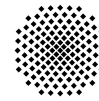
Characterizing intonation

- ▶ Intonation patterns consist of intonation features:
 - ▶ intonational contour (tune)
 - ▶ prominence (stress)
 - ▶ intonational phrasing
 - ▶ pitch range



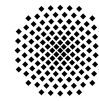
Characterizing intonation

- ▶ Intonation patterns consist of intonation features:
 - ▶ intonational contour (tune)
 - ▶ prominence (stress)
 - ▶ intonational phrasing
 - ▶ pitch range
- ▶ The contour indicates the movement of pitch.
 - ▶ For example, the intonation pattern of an assertion has a distinct contour from that of a question (cf. Ladd 2008, p. 7ff.)



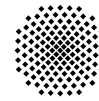
Characterizing intonation

- ▶ Intonation patterns consist of intonation features:
 - ▶ intonational contour (tune)
 - ▶ prominence (stress)
 - ▶ intonational phrasing
 - ▶ pitch range
- ▶ The contour indicates the movement of pitch.
 - ▶ For example, the intonation pattern of an assertion has a distinct contour from that of a question (cf. Ladd 2008, p. 7ff.)
- ▶ Intonational phrasing divides the sequence of words into intonational units, the intonational (prosodic) phrases.
 - ▶ Phrase boundaries are marked by pauses, boundary tones and duration patterns.



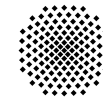
Characterizing intonation

- ▶ Intonation patterns consist of intonation features:
 - ▶ intonational contour (tune)
 - ▶ prominence (stress)
 - ▶ intonational phrasing
 - ▶ pitch range
- ▶ The contour indicates the movement of pitch.
 - ▶ For example, the intonation pattern of an assertion has a distinct contour from that of a question (cf. Ladd 2008, p. 7ff.)
- ▶ Intonational phrasing divides the sequence of words into intonational units, the intonational (prosodic) phrases.
 - ▶ Phrase boundaries are marked by pauses, boundary tones and duration patterns.
- ▶ Prominence is realized through pitch accents - defined as:
 - ▶ a local feature of a pitch contour - usually a *pitch change*, and often involving a local maximum or minimum - which signals that the syllable with which it is associated is prominent in the utterance.
(Ladd 2008, p. 48)



Autosegmental-metrical approach to intonation

- ▶ Pierrehumbert (1980) and Beckman & Pierrehumbert (1986) propose a description of intonation:
 - ▶ the grammar of phrasal tunes, consisting of L and H tones:
 - ▶ pitch accents
 - ▶ phrase accents
 - ▶ boundary tones



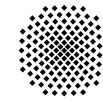
Autosegmental-metrical approach to intonation

- ▶ Pierrehumbert (1980) and Beckman & Pierrehumbert (1986) propose a description of intonation:
 - ▶ the grammar of phrasal tunes, consisting of L and H tones:
 - ▶ pitch accents
 - ▶ phrase accents
 - ▶ boundary tones
 - ▶ the metrical representation of the text
 - ▶ rules for lining up the tune with the text
- ▶ Phonological tones
 - ▶ Each phrase requires at least one pitch accent
 - ▶ English: H^*, L^* , or bitonal: H^*+L , $H+L^*$, L^*+H , $L+H^*$, H^*+H



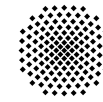
Autosegmental-metrical approach to intonation

- ▶ Pierrehumbert (1980) and Beckman & Pierrehumbert (1986) propose a description of intonation:
 - ▶ the grammar of phrasal tunes, consisting of L and H tones:
 - ▶ pitch accents
 - ▶ phrase accents
 - ▶ boundary tones
 - ▶ the metrical representation of the text
 - ▶ rules for lining up the tune with the text
- ▶ Phonological tones
 - ▶ Each phrase requires at least one pitch accent
 - ▶ English: H^*, L^* , or bitonal: H^*+L , $H+L^*$, L^*+H , $L+H^*$, H^*+H
 - ▶ Each phrase receives a phrase accent at the end of the word associated with the last pitch accent:
 - ▶ H^- , L^-
 - ▶ Each phrase ends with a boundary tone:
 - ▶ $H\%$, $L\%$



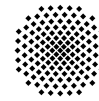
Intonational meaning

- ▶ There are two main questions with respect to intonational meaning:
 - ▶ What are the meaningful units of intonation?
 - ▶ What kind of meanings are associated with these units?
- ▶ Domains of intonational patterns: tune, phrasing, and pitch accent
- ▶ Meaning types that are associated with each of the domains:
 - ▶ Tune is often correlated with speech acts.
 - ▶ Phrasing is mostly associated with information structure.
 - ▶ The pitch accent is linked with the notion of focus.



The limits of prosody labelling for the annotation of information structure

- ▶ Based on autosegmental-metrical approach to intonation ToBI (Tones and Break Indices) a system for transcribing the intonation patterns and other aspects of the prosody of English utterances was developed.
 - ▶ However, labelling the prosody of spoken-language corpora with ToBI is a non-trivial enterprise:
 - ▶ very time-consuming
 - ▶ it is usually done manually with very low inter-annotator scores
 - ▶ systems doing automatic ToBI labelling are still in their infancy and not very reliable
- We will therefore annotate information structure without relying on the labelling of prosody.



Relating intonation and interpretation: focus projection

- ▶ The word marked by a pitch accent and the extension of the focus are related to each other by rules of *focus projection*.

(27) Mary bought a book about BATS.

(28) a. Q: What did Mary buy a book about?

A: Mary bought a book about [BATS.]_F

b. Q: What did Mary buy?

A: Mary bought [a book about BATS.]_F

c. Q: What did Mary do?

A: Mary [bought a book about BATS.]_F

d. Q: What happened?

A: [Mary bought a book about BATS.]_F



Syntactic explanation of focus projection

- ▶ F is a *syntactic feature*, which originates in a pitch accented word, and then spreads onto larger constituents.
- ▶ Focus projection rules (Selkirk, 1996):

Basic Focus Rule: An accented word is F-marked.

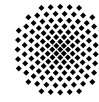
(29) BATS → BATS_F

Horizontal focus projection: An F-feature may project from an internal argument to the head of a phrase.

(30) [_{PP} [_P about] BATS_F] → [_{PP} [_P about]_F BATS_F]

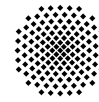
Vertical focus projection: An F-feature may project from the head of a phrase onto the phrase itself.

(31) [_{PP} [_P about]_F BATS_F] → [_{PP} [_P about]_F BATS_F]_F



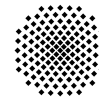
Full example

(32) [Mary [bought [a [book [about BATS]]]]]



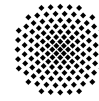
Full example

(32) [Mary [bought [a [book [about $BATS_F$]]]]]



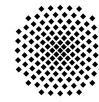
Full example

(32) [Mary [bought [a [book [about_F BATS_f]]]]]



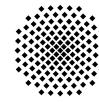
Full example

(32) [Mary [bought [a [book [about_f BATS_f]_F]]]]



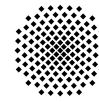
Full example

(32) [Mary [bought [a [book_F [about_f BATS_f]_f]]]]



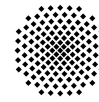
Full example

(32) [Mary [bought [a [book_f [about_f BATS_f]_f]_F]]]



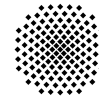
Full example

(32) [Mary [bought [a_F [book_f [about_f BATS_f]_f]]]]



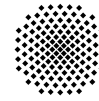
Full example

(32) [Mary [bought [a_f [book $_f$ [about $_f$ BATS $_f$] $_f$] $_f$] $_F$]



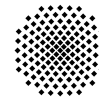
Full example

(32) [Mary [bought_F [a_f [book_f [about_f BATS_f]_f]_f]_f]]



Full example

(32) [Mary [bought_f [a_f [book_f [about_f BATS_f]_f]_f]_f]_F]



Full example

(32) [Mary [bought_f [a_f [book_f [about_f BATS_f]]]]]_F



Focus projection (cont.)

- ▶ The focus projection rules determine the focus projection potential of a pitch accent dependent on the syntactic surface structure.

For example, the pitch accents in (33) and (34) cannot project focus to larger constituents, i.e. they are not felicitous answers to the questions in (35).

(33) Q: Who bought a book about bats?

A: [MAry]_F bought a book about bats.

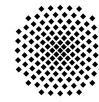
(34) Q: What related to bats did Mary buy?

A: Mary bought a [BOOK]_F about bats.

(35) a. What did Mary buy?

b. What did Mary do?

c. What happened?

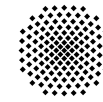


Focus projection - some “real life” examples

(36) *A reporter*: Why do you rob banks?

Bank robber Willie Sutton: Because this is where the money is.

- ▶ The reporter’s question is intended to have broad focus on the phrase *rob banks*.
- ▶ Sutton’s reply treats the question as having narrow focus on *banks*.



Focus projection - some “real life” examples

(Ladd 2008, p. 255)

Peanuts by Charles Schulz



© PEANUTS Worldwide LLC - All Rights Reserved.



Focus projection and constituency

- ▶ A rarely noted fact is that the focus resulting from one pitch accent does not always correspond to a constituent.

Höhle (1982)

(37) Was hat das Kind erlebt? / What did the child experience?

a. [Karl]_F hat dem Kind [das BUCH geschenkt]_F .

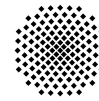
Karl has the child the book given

‘Karl gave the child the book as a present.’

Büring (2007)

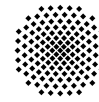
(38) (What happened to my harp?)

a. Someone_F sent_F it to NORway_F .



Word order and information structure

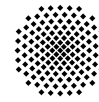
- ▶ Languages such as Russian, Hungarian, Czech, Catalan, or Turkish mark information structure through word order.
- ▶ But even intonational languages like English and German combine intonation and word order to mark certain information structurings.
- ▶ Word order phenomena used for information structuring are for example: topicalization, cleft constructions, and scrambling



Topicalization in English

- (39) Q: Who did you meet in Germany?
A: In Germany, I met a lot of old friends.
- (40) Q: You look so happy, what happened?
A: # In Germany, I met a lot of old friends.

Topicalization in English is not possible when an all-focus answer is expected.



Cleft constructions

It-cleft construction:

(41) It was a bar of chocolate that Peter bought for Mary.

Wh-cleft construction:

(42) What Peter bought for Mary was a bar of chocolate.

- ▶ The clefted element *a bar of chocolate* is the focus of the sentence.
- ▶ The cleft clause containing the gap constitutes the background information.



C'est cleft in French

(43) Marie a mangé un biscuit.

(44) a. Qui a mangé un biscuit?

b. C'est Marie, qui a mangé un biscuit.

- ▶ It is sometimes argued that French categorically bans prosodic marking from sentence-initial position (cf. Lambrecht 1994).
- ▶ In the case of narrow subject focus, as in (44b), the subject has to be moved into a dedicated focus position, the cleft.



References

- Ammann, H. (1928). *Die menschliche Rede. Sprachphilosophische Untersuchungen. 2. Teil: Der Satz*. Darmstadt: Wissenschaftliche Buchgesellschaft. Reprinted 1962, Lahr im Schwarzwald: Moritz Schauenburg.
- Beaver, D. I. & B. Z. Clark (2008). *Sense and sensitivity: How focus determines meaning*. John Wiley & Sons.
- Beckman, M. & J. Pierrehumbert (1986). Intonational Structure in Japanese and English. *Phonology Yearbook* 3, 255–309.
- Büring, D. (2007). Semantics, intonation and information structure. In *The Oxford Handbook of Linguistic Interfaces*, Oxford: Oxford University Press.
- Chomsky, N. (1971). Deep Structure, Surface Structure, and Semantic Interpretation. In D. Steinberg & L. Jakobovits (eds.), *Semantics: An Interdisciplinary Reader in Philosophy, Linguistics, and Psychology*, Cambridge, UK: Cambridge University Press.
- Firbas, J. (1964). On Defining the Theme in Functional Sentence Analysis. *Travaux Linguistique de Prague* 1, 267–280.
- Halliday, M. (1967). Notes on Transitivity and Theme in English. Part 1 and 2. *Journal of Linguistics* 3, 37–81, 199–244.
- Höhle, T. N. (1982). Explikationen für ‘normale Betonung’ und ‘normale Wortstellung’. In W. Abraham (ed.), *Satzglieder im Deutschen*, Tübingen: Gunter Narr Verlag, pp. 75–153.
- Jackendoff, R. (1972). *Semantic Interpretation in Generative Grammar*. Cambridge, MA: MIT Press.
- Jacobs, J. (1983). *Fokus und Skalen. Zur Syntax und Semantik der Gradpartikel im Deutschen*. Tübingen: Max Niemeyer Verlag.
- Krifka, M. (1992). A Compositional Semantics for Multiple Focus Constructions. In J. Jacobs (ed.), *Informationsstruktur und Grammatik*, Opladen: Westdeutscher Verlag, pp. 17–54.
- Kruijff-Korbayová, I. & M. Steedman (2003). Discourse and Information Structure. *Journal of Logic, Language and Information (Introduction to the Special Issue)*.



-
- Ladd, D. R. (2008). *Intonational Phonology*, vol. 119 of *Cambridge Studies in Linguistics*. Cambridge: Cambridge University Press.
- Lambrecht, K. (1994). *Information Structure and Sentence Form — Topic, focus, and the mental representation of discourse referents*, vol. 71 of *Cambridge Studies in Linguistics*. Cambridge, UK: Cambridge University Press.
- Pierrehumbert, J. (1980). *The Phonetics and Phonology of English Intonation*. Ph.d., MIT.
- Rooth, M. (1985). *Association with Focus*. Ph.D. thesis, University of Massachusetts, Amherst, MA.
- Sgall, P., E. Hajičová & E. Benešová (1973). *Topic, Focus and Generative Semantics*. Kronberg/Taunus: Scriptor.
- Sgall, P., E. Hajičová & J. Panevová (1986). *The meaning of the Sentence in Its Semantic and Pragmatic Aspects*. Prague and Dordrecht: Academia and Reidel.
- Stechow, A. v. (1981). Topic, Focus, and Local Relevance. In W. Klein & W. Levelt (eds.), *Crossing the Boundaries in Linguistics*, Dordrecht: Reidel, pp. 95–130.
- Stechow, A. v. & M. J. Cresswell (1983). De Re Belief Generalized. *Linguistics & Philosophy* 5, 503–535.
- von Heusinger, K. (1999). *Intonation and Information Structure. The Representation of Focus in Phonology and Semantics*. Habilitationsschrift, Universität Konstanz, Konstanz, Germany. URL <http://www.ilg.uni-stuttgart.de/vonHeusinger/publications/abstracts/habil.html>.