

# Information Structure Across Frameworks

**Part I: Problems defining IS notions.  
Major generative IS theories**

Natalia Slioussar

HSE, Moscow, and St.Petersburg State University

[slioussar@gmail.com](mailto:slioussar@gmail.com)

# Some examples to start with

- (1) a. *Mal'chik el kashu.*  
boy.NOM ate porridge.ACC
- b. *Mal'chik kashu el.*
- c. *Kashu mal'chik el.*
- d. *Kashu el mal'chik.*
- e. *El mal'chik kashu.*
- f. *El kashu mal'chik.*

# Some examples to start with

(1) a. *Mal'chik el kashu.*  
boy.NOM ate porridge.ACC

b. *Mal'chik kashu el.*

c. *Kashu mal'chik el.*

d. *Kashu el mal'chik.*

e. *El mal'chik kashu.*

f. *El kashu mal'chik.*

(2) a. *Mal'chik el KASHU.*

b. *MAL'CHIK el kashu.*

# Some examples to start with

(1) a. *Mal'chik el kashu.*  
boy.NOM ate porridge.ACC

b. *Mal'chik kashu el.*

c. *Kashu mal'chik el.*

d. *Kashu el mal'chik.*

e. *El mal'chik kashu.*

f. *El kashu mal'chik.*

(2) a. *Mal'chik el KASHU.*

b. *MAL'CHIK el kashu.*

Also IS-related  
morphemes and  
constructions and  
many other things.

# General idea of these lectures

- Present several major IS theories and various problems faced by them
- Show how by looking at IS phenomena we can address several 'big questions':
  - what is the nature of syntactic movement?
  - what is the architecture of the grammar?
  - the problem of grammatically encoded meanings
  - the problem of testing linguistic theories in psycholinguistic experiments

# Some history

Prague School: the first IS studies and still a productive tradition (Mathesius 1932, Hajičová and Sgall 1988; Hajičová, Partee and Sgall 1998).

Also many works by Russian linguists (e.g. Kovtunova 1976; Krylova 1992; Krylova and Khavronina 1976; Lapteva 1976; Paducheva 1984; Shvedova, ed. 1980; Sirotinina 1965, 1974; Zemskaya 1979, ed. 1973; Zemskaya, Kitaygorodskaya, and Shiryaev 1981; Zolotova 1982; Yanko, 2001).

My lectures: focusing mainly on formal models, specifically on generative ones, with their characteristic questions.

# IS notions

Mathesius: theme = the starting point of the utterance,  
rheme = what we want to say about the theme.

Also topic and focus.

Many authors introduce different types of topics and foci  
(e.g. new information vs. contrastive focus) or add other  
notions (e.g. given, contrast).

Major problem: common terminology masking very different  
views.

# Focus at first glance

Mathesius: theme = the starting point of the utterance, rheme = what we want to say about the theme.

Focus = new (Halliday (1966); many authors incorporate in their definitions, e.g. Selkirk (1984) or use as a focus test).  
Rochemont (1986): previously not activated information.

Focus = not presupposed information (e.g. Lambrecht (1994)).

Question-answer test (Hatcher 1956). Widely used, incorporated in many definitions (e.g. Selkirk 1984; Rooth 1985, 1992). QUD, alternative semantics (see below).

# Focus: alternative semantics

Rooth (1992): “Informally, the focus semantic value for a phrase of category  $S$  is the set of propositions obtainable from the ordinary semantic value by making a substitution in the position corresponding to the focused phrase. For instance, the focus semantic value for the sentence  $[_S[\text{Mary}]_F \text{ likes Sue}]$  is the set of propositions of the form ‘ $x$  likes Sue’” (p. 76). In the set abstraction notation this can be stated as “ $\llbracket [_S[\text{Mary}]_F \text{ likes Sue}] \rrbracket^f = \{\text{like}(x, s) \mid x \in E\}$ , where  $E$  is the domain of individuals” (p. 76, n. 2a).

# Focus: where are you?

In practice: most authors rely on the question-answer test, and, in most European languages, even more often just check where the main stress is.

This does not follow from any definition and, as we will see, the results sometimes contradicts other tests. Another problem: the main stress does not define the borders of the focused constituent.

As a result...

# Focus: problems

Major problem: impossibility to pinpoint a single continuous constituent as the focus of the sentence

- 'Nested foci':

- (2) a. *When I came home, rather than doing his homework, [<sub>IP</sub> Johnny was [<sub>VP</sub> reading [<sub>DP</sub> SUPERMAN] to some kid]]* (Neeleman & Szendrői, 2004).
- b. *[An [AMERICAN] farmer was talking to a [CANADIAN] farmer]* (Rooth, 1992).
- c. *[People who [GROW] rice only[EAT] rice]* (Rooth, 1992).

# Focus: problems

- Discontinuous foci:

(3) Q: *Did John give a book to Bill?*

A: *No, this time he [grew a pot of NARCISSUS for him]*  
(Selkirk, 1984)

(4) *John only [turned it OFF]* (Krifka, 1991)

# Topic at first glance

Lambrecht (1994):

Following Strawson (1964), Kuno (1972), Gundel (1974), Chomsky (1976), Dik (1978), Reinhart (1982) and many other linguists: “The topic of a sentence is the thing which the proposition expressed by the sentence is about” (p. 118).

To clarify, Lambrecht cites Strawson (1964): “Statements, or the pieces of discourse to which they belong, have subjects, not only in the relatively precise sense of logic and grammar, but in a vaguer sense with which I shall associate the words ‘topic’ and ‘about’ [...]. We [...] intend in general to give or add information about what is a matter of standing current interest or concern. There is great variety of possible types of answer to the question what the topic of a statement is, what a statement is ‘about’ [...] and not every such answer excludes every other in a given case” (p. 97).

# Topic at first glance

Lambrecht (1994) notes that Strawson's remark "stresses the inherently vague character of the notions of aboutness and relevance. This inherent vagueness has consequences for the grammatical coding of topics in sentences. If the topic is seen as the matter of current interest which a statement is about and with respect to which a proposition is to be interpreted as relevant, it is clear that one cannot always point to a particular element in a proposition, let alone to a particular constituent of a sentence, and determine that this element and nothing else is the topic of the sentence. As there are degrees of relevance, there are degrees to which elements of propositions qualify as topics. It is this fact, I believe, which accounts for the absence of unambiguous formal marking of the topic relation in many languages. And, as a corollary, it accounts for the fact that in those languages which do have formal topic marking this marking reflects only imperfectly the relative degrees of topicality of given referents" (p. 119).

# Topic: examples

An example: Reinhart (1982) and Lambrecht (1994) on topics in the following passage from (Ochs, 1979, p. 63):

(5) *And it's uh got ta good taste, it's good. And the cereal – grandma e don't like cereal but she finished to the last (dish) and I enjoy – I like it too. It's tasty! And I uh (1.2) He didn't want the cereal, doesn't eat. I said, "Todd, it wouldn't kill ya, taste it!"...*

Reinhart: in all sentences before the pause, the topic is the cereal and after the pause the topic is the grandson.

Lambrecht: also *it* in *it wouldn't kill ya, grandma* and *I* in *grandma don't like cereal but she finished and I like it too...*

# Topic: problems

Major problem: extremely diverse means of encoding

Classical cases:

- different positions for subject and object topics

(5) a. Knigu Ivan pročel.  
book.ACC Ivan.NOM read  
'The book, Ivan read.'

b. Ivan pročel knigu.  
Ivan.NOM read book.ACC  
'Ivan read the book.'

# Topic: problems

- pronominalization on a par with syntactic movement.

(6) Q: *Čto slučilos' s Petrom?*  
What happened to Petr?

A1: *Ivan ego UDARIL.*  
Ivan.NOM he.ACC hit  
'Ivan HIT him.'

A2: *Ivan UDARIL ego.*  
Ivan.NOM hit he.ACC  
'Ivan HIT him.'

# Other IS notions

Given, D-linked = linked to an accessible discourse entity (Pesetsky 1987; Reinhart 1995, 2006 etc.), anaphoric (e.g. Williams 1997; Schwarzschild 1999).

- (7) a. *Has John read 'Slaughterhouse five'?* (Ladd 1980)  
b. *No, he does not READ books.*

Accessibility hierarchies (Chafe 1976; Givón 1983; Ariel 1990).  
“Cognitive” definitions, encoding problem.

Background. Topic, focus and tail (e.g. Vallduví and Vilkuna 1998).

Choi (1999), LFG framework: [+/-new], [+/-prom].

# Other IS notions

Contrast, contrastive vs. non-contrastive topics and foci (e.g. Rochemont 1986; Vallduví and Vilkuna 1998; Zubizarreta and Vergnaud 2005).

Terminology: new information / presentational vs. contrastive foci, aboutness vs. contrastive topics.

Problems: Kiss vs. Szendroi on Hungarian (Kiss 1998; Szendroi 2001 etc.), Brunetti (2004) on Italian, an experiment on English by Ladd (1980).

# Same notions, different views

Does every sentence have a topic? And a focus?

Thetic sentences. Can an SVO sentence be thetic?

Some approaches are hardly incompatible with multiple foci.

# Major generative approaches to IS

## 1. IS features similar to other syntactic features

(Rizzi 1997) and the majority of other works (Bródy 1990, 1995; Laka 1990; Ortiz de Urbina 1999; Ouhalla 1994; Tsimpli 1995; Tuller 1992; Vilkuna 1995 etc.)

## 2. F feature different from other syntactic features

(Jackendoff 1972) and many other works (e.g. (Rooth 1985, 1992), we will look at (Selkirk 1984, 1995; Büring 2006))

## 3. Configurations

(Reinhart 1995, 2006; Neeleman and Reinhart 1998; Neeleman and van de Koot 2008; Szendrői 2001 etc.)

Also a **separate IS level...**

# Feature-based theories. Rizzi (1997)

[<sub>ForceP</sub> Force [<sub>TopP</sub> Top [<sub>FocP</sub> Foc [<sub>TopP</sub> Top [<sub>FinP</sub> Fin [<sub>IP</sub> ...]]]]]]

Arguments for *Force*, *Fin* and IS projections in between:

- (7) a. Credo che il tuo libro, loro lo apprezzerebbero molto  
“I believe that your book, they would appreciate it a lot”  
b.\*Credo, il tuo libro, che loro lo apprezzerebbero molto  
“I believe, your book, that they would appreciate it a lot”
- (8) a.\*Credo di il tuo libro, apprezzarlo molto  
“I believe ‘of’ your book to appreciate it a lot”  
b. Credo, il tuo libro, di apprezzarlo molto  
“I believe, your book, ‘of’ to appreciate it a lot”

# Feature-based theories. Rizzi (1997)

- (9) a. Un uomo a cui, il premio Nobel, lo daranno senz'altro  
"A man to whom, the Nobel Prize, they will give it undoubtedly"
- b.\*Un uomo, il premio Nobel, a cui lo daranno senz'altro  
"A man, the Nobel Prize, to whom they will give it undoubtedly"
- (10) a.\* A chi, il premio Nobel, lo daranno?  
"To whom, the Nobel prize, will they give it?"
- b. Il premio Nobel, a chi lo daranno? INT DIS (TRANS)  
"The Nobel prize, to whom will they give it?"
- (11) a. Mi domando, il premio Nobel, a chi lo potrebbero dare  
"I wonder, the Nobel Prize, to whom they could give it"
- b.? Mi domando a chi, il premio Nobel, lo potrebbero dare INT  
"I wonder to whom, the Nobel Prize, they could give it"

# Feature-based theories. Rizzi (1997)

Multiple TopPs around a single FocP:

- (12) **Il libro, a Gianni, domani, glielo darò senz'altro**  
“The book, to John, tomorrow, I'll give it to him for sure”
- (13) \***A GIANNI IL LIBRO darò (non a Piero, l'articolo)**  
“TO JOHN THE BOOK I'll give, not to Piero, the article”
- (14) **A Gianni, QUESTO, domani, gli dovrete dire**  
“To Gianni, THIS, tomorrow, you should tell him”

# Feature-based theories. Rizzi (1997)

Some minuses:

1. Most foci are not fronted. Covert movement?..

WCO effects with overt focus fronting:

(15) **Gianni<sub>i</sub>, sua<sub>i</sub> madre lo<sub>i</sub> ha sempre apprezzato**  
“Gianni, his mother always appreciated him”

(16) ?? **GIANNI<sub>i</sub> sua<sub>i</sub> madre ha sempre apprezzato t<sub>i</sub> (non Piero)**  
“GIANNI his mother always appreciated, not Piero”

Similar picture with *in situ* foci:

- (17) a. \* *The man that she<sub>i</sub> met liked MARY<sub>i</sub>.*  
b. *The man that she<sub>i</sub> met LIKED Mary<sub>i</sub>.*

# Feature-based theories. Rizzi (1997)

But there is an alternative explanation (Rochemont 1986; Szendrői 2005):

- (18) a. *A: Sally and the woman John loves are leaving the country today.*
- b. *B: I thought that the woman he loves has BETRAYED Sally.*
- c. *A: No, the woman he<sub>i</sub> loves betrayed JOHN<sub>i</sub>.*

# Feature-based theories. Rizzi (1997)

2. Inclusiveness condition

3. Main stress is a strange morpheme

4. Languages with 'regular' IS morphemes exhibit various properties, but never what can be expected based on (Rizzi 1997)

# IS morphemes

Gungbe (Aboh 2004 etc.):

(19) a. *Súru tò àklà cè zé dù`.*  
Suru PROG cookie 1SG.POSS take eat.NR  
'Suru takes and eats my cookie.'

b. *Àklà cè wè Súru tò zízé dù`.*  
cookie 1SG.POSS FM Suru PROG take.take eat.NR  
'It is my cookie that Suru takes and eats.'

# IS morphemes

VP can also be focused (no *wè* , but movement + progressive marker changes from *tò* to *tè*):

- (19) *Àklà cè zé d̀ù` Súrù tè*  
cookie 1SG.POSS take eat.NR Suru PROG  
«Suru takes and eats my cookie».

# IS morphemes

Somali (Hyman and Watters 1984; Svolaccia, Mereu and Puglieli 1995; Lecarme 1999), data reanalysed by Szendrői (2005):

- (20) a. *Cali baa moos cunay.*  
A. FM banana ate  
'Ali ate a banana.'
- b. *Cali moos buu cunay.*  
A. banana FM ate  
'Ali ate a banana.'
- c. *Cali moos waa cunay.*  
A. banana FM ate  
'Ali ate a banana.'

# IS morphemes

(21) a. *Cali muxuu sameeyay?*

A. what did  
«What did Ali do?».

b. *Cali yuu dilay?*

A. who hit  
«Who did Ali hit?».

(22) a. *Cali Maryan buu dilay.*

A. M. FM hit  
«Ali hit Maryan».

b. *Maryan buu dilay Cali.*

M. FM hit A.  
«Ali hit Maryan».

# IS morphemes

(23) *Maxaad hacay?*  
what happened

(24) a. *Cali baa Maryan dilay.*  
A. FM M. hit  
«Ali hit Maryan».

b. ? *Maryan buu Cali dilay*  
M. FM A. hit  
«Ali hit Maryan».

# IS morphemes

Tagalog (examples from Wikipedia):

- (25) a. *Binilí ng lalaki ang saging sa tindahan para sa unggóy.*  
bought man banana at store for monkey
- b. *Bumilí ng saging ang lalaki sa tindahan para sa unggóy.*  
bought banana man at store for monkey
- c. *Binilhán ng lalaki ng saging ang tindahan.*  
bought man banana at store
- d. *Ibinilí ng lalaki ng saging ang unggóy.*  
bought man banana for monkey

# Feature-based theories.

## *F* as a special feature

Jackendoff (1972) and many other works (e.g. (Rooth 1985, 1992), we will look at (Selkirk 1984, 1995; Büring 2006))

Unlike other grammatical features, *F* does not interact with syntax (no agreement, no movement related to it). It only attracts the main stress.

I will not discuss these theories now because they obviously cannot account for IS-related word order alternations.

# Configurational theories.

## Reinhart and Neeleman's model

Reinhart and Neeleman's model  
(Reinhart 1995, 2006; Neeleman and Reinhart 1998)

Focus set: all and only the constituents containing the main stress of the sentence. The actual focus is chosen at the discourse interface.

I.e. configurations rather than IS features and direct reliance on prosody.

# Configurational theories.

## Reinhart and Neeleman's model

*My neighbor is building a DESK*: focus set {IP, VP, DP<sub>o</sub>}

- (26) a. A: *What is this noise?*  
b. B: *My neighbor is building a DESK.*
- (27) a. A: *What is your neighbor doing?*  
b. B: *My neighbor* *is building a DESK.*
- (28) a. A: *What is your neighbor building?*  
b. B: *My neighbor is building* *a DESK.*

# Configurational theories.

## Reinhart and Neeleman's model

If an XP is not in the focus set, the stress can be shifted:

- (29) a. A: *Who is building a desk?*  
b. B: *My NEIGHBOR is building a desk.*

# Configurational theories.

## Reinhart and Neeleman's model

If an XP is not in the focus set, the stress can be shifted:

(29) a. A: *Who is building a desk?*

b. B: *My NEIGHBOR is building a desk.*

This operation is subject to Economy Principle: can be used only to get a new interpretation at the discourse interface.

Therefore: no wide focus in (29b) (but consider (2a)!).

Psycholinguistic predictions...

# Configurational theories.

## Reinhart and Neeleman's model

Stress shift in (29b): stress strengthening (new element in the focus set) + destressing (D-linking)

D-linking leads to destressing:

- (30) a. A: *Has John read 'War and Peace'?*  
b. B: *No, he does not READ books.*

Destressing without stress strengthening:

- (31) a. *Max SAW her / it.*  
b. # *Max saw HER / IT.*

# Configurational theories.

## Reinhart and Neeleman's model

Reinhart and Neeleman's account of Dutch scrambling:

- (32) a. ... *dat Jan langzaam het BOEK las.*  
          that Jan slowly       the book read
- b. ... *dat Jan het boek langzaam LAS.*

# Configurational theories.

## Reinhart and Neeleman's model

Reinhart and Neeleman's account of Dutch scrambling:

- (32) a. ... *dat Jan langzaam het BOEK las.*  
          that Jan slowly       the book read
- b. ... *dat Jan het boek langzaam LAS.*

R&N use base-generation for (32a-b). Otherwise look-ahead and independent prosody...

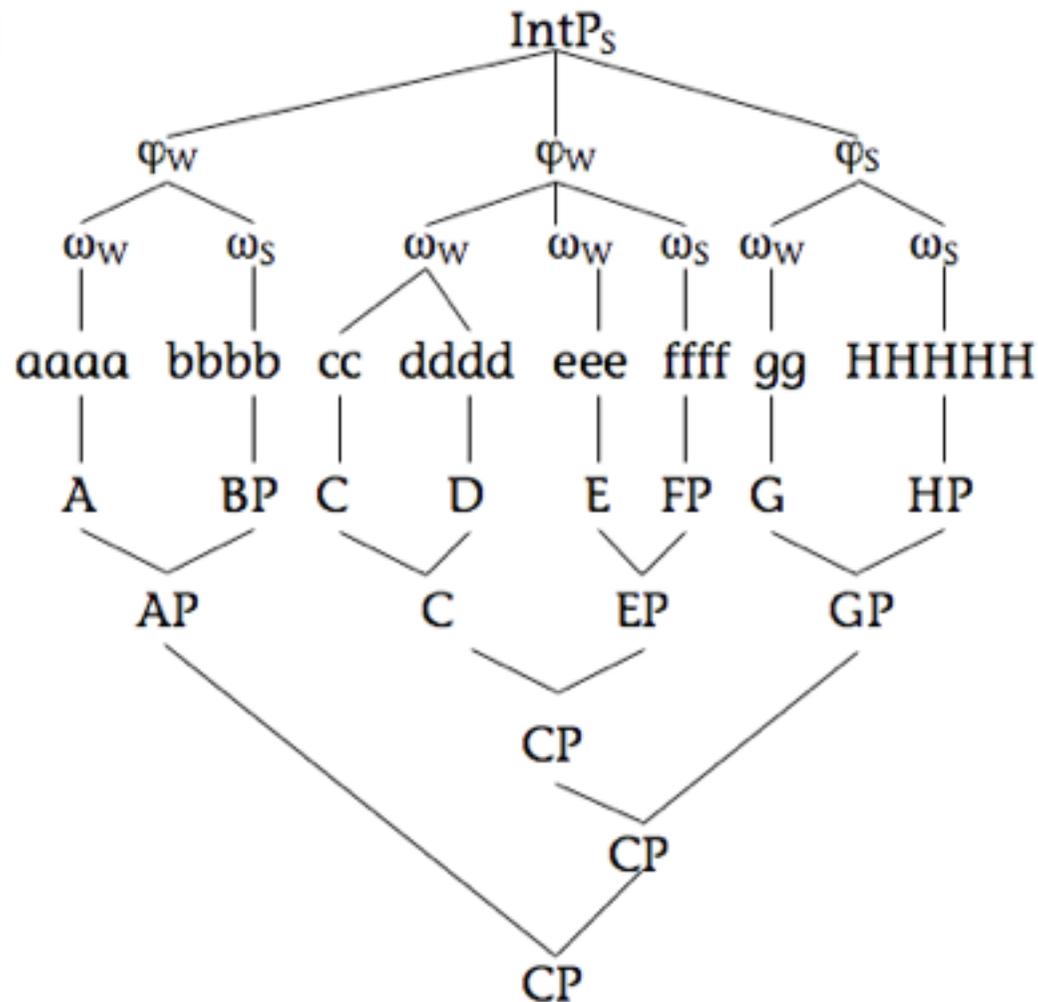
But IS-related reorderings e.g. in Italian or Russian cannot be base-generated. So Szendrői (2001, 2005) who extends the theory to Italian, Hungarian etc. opts for independent prosody.



# Configurational theories.

## Szendrői's model

Syntax and prosody:



# Configurational theories.

## Szendrői's model

Intuition: some languages move the main stress to  $X$ , the others move words so that  $X$  receives the main stress.

# Configurational theories.

## Szendrői's model

Intuition: some languages move the main stress to  $X$ , the others move words so that  $X$  receives the main stress.

Implemented using OT constraints:  
SYNTAX, PROSODY (and also MAP)

NB! Essentially non-feature-driven prosodically motivated movement.

Thus, choice of IS theory => implications for the grammar architecture and the nature of movement (these problems will be discussed in more detail).

# Configurational theories.

## Szendrői's model

English:

- (33) a. *My neighbor is building a DESK.*  
b. *My NEIGHBOR is building a desk.*

Italian:

- (34) a. *Il mio vicino sta costruendo una SCRIVANIA.*  
b. *Sta costruendo una scrivania il mio VICINO.*

English:

- (35) *Max SAW it.*

Italian:

- (36) *Lo ho VEDUTO.*

# Configurational theories. Szendrői's model

PROSODY: RAS and LAS

Used to explain e.g. the following differences between English and Hungarian:

(37) a. *the red DOOR*

b. *a PIROS ajtó*  
the red door

(38) a. *Esterhazy's most recent BOOK*

b. *ESTERHÁZYNAK a legújabb könyve*  
Esterhazy's the most recent book

# Configurational theories. Szendrői's model

These constraints are also used to explain focus movement in Hungarian (a classical case for feature-based theories):

- (39) a. *Józsi*            *MARIT*    *ismeri.*  
          Joseph.NOM    Mary.ACC    knows
- b. *MARIT*    *ismeri*    *Józsi.*
- c. *JÓZSI*    *ismeri*    *Marit.*
- d. *Józsi*    *ISMERI*    *Marit.*
- e. *Józsi*    *ISMERI*    *Marit.*
- f. *ISMERI*    *Józsi*    *Marit.*