Information structure: the approach of Nomi-Erteschik-Shir

Tatiana Philippova

Seminar of the Laboratory of Formal Models in Linguistics

May 24, 2019

Outline

- The basics of the system: topics, foci and their varieties
- The notion of canonical information structure
- Scope calculation
- Island constraints as IS constraints
- Architecture of grammar

Key features

- Only two primitives – Topic and Focus
- Subtypes of topic and focus (restrictive, contrastive) are combinatorially derived from the primitives
- There may be multiple topics and foci in a sentence (provided by main and subordinate focus-structures)
- There are no designated functional heads like Top and Foc: [top] and [foc] are features optionally assigned in the syntax (as part of lexical selection)

Topic

- Traditional definition: given, old, what the sentence is about

- Erteschik-Shir: the topic is the subject of predication = the pivot for truth value assessment
  (following Strawson 1964, Reinhart 1981)

Topic as the pivot of T-value assessment

- The king of France is bald.

Possible answer to the following questions:

a. What is the King of France like? → Truth value gap

b. What bald notables are there? → FALSE

- The exhibition was visited by the King of France. → FALSE

- The topic will have wider scope than other elements in the sentence because the predicate of the sentence is evaluated with respect to the topic.

Topic

The topic is the subject of predication = the pivot for truth value assessment

→ *Every sentence must have a topic*

- All-focus sentences have implicit "stage" topics indicating the spatio-temporal parameters of the sentence = the here-and-now of the discourse (Erteschik-Shir 1997, following Gundel 1974)

Possible topics in English

Subject Topic

Q: Why does John look so pleased with himself?

A: *He* washed the dishes.

Object Topic

Q: What happened to the dishes?

A: John washed *them*.

Stage Topic (Out-of-the-blue)

Q: What happened?/Why do you look so pleased?

A: John washed the dishes!

Overt stage topics in English

Temporal and spatial locations may play the role of topic:

(5) a. There was a nametag near every plate.
b. A flag was hanging *in front of every window*.
c. A ghost appeared *at midnight*..
d. A bus to Beer Sheva leaves *every 20 minutes/regularly*.

(Erteschik-Shir 2014)

Topics are given, not old

Permanently available topics

- Speaker, Hearer, the here-and-now (implicit stage topic), "permanent and temporary fixtures of our world": *the moon*, *the president*…

Temporarily available topics:

- "referents that exist or appear on the current scene and may be introduced deictically"

$$[\textit{That/That chair}]_{\text{top}}\ \textit{is ugly}$$

can be uttered out of the blue if accompanied by pointing

- A topic can be derived from a previously mentioned referent by world knowledge

John heard a beautiful concert. *The composer* directed it.

(Erteschik-Shir 2007: 17-19)

Topics are marked by

- Morphology (e.g. Japanese *wa*)
- Dislocation (topicalization, scrambling)
- Pronominalization (use of a weak or a clitic pronoun)
- Intonation (including destressing)
- Argument drop (radical form of destressing?)

(adapted from Erteschik-Shir, Ibnbari, Taube 2013)

Focus

- Traditional characterization: answer to *wh*-question, stressed
- The Focus of a sentence S = the (intension of a) constituent c of S which the speaker intends to direct the attention of his/her hearer(s) to, by uttering S. (Erteschik-Shir 1973, Erteschik-Shir and Lappin 1979)
- Sentential stress rule: Place primary stress on the focus of the sentence

Focus and stress

- The stress rule applies to *every* focus in sentences with complex f-structures.

  Q: What happened?

  A: $_{s}\text{TOP}_{t}$[the children ate the candy]$_{foc}$

  H* H* H*

- A "rhythm rule" operates in fast speech which is responsible for lowering some of the pitch accents assigned by the stress rule when fast speech prevents them from receiving their full stress:

  a. the CHILdren ATE the CANdy

  b. the CHILdren ate the CANdy

  c. the children ate the CANdy

- Topics contained in focus constituents are excluded by the stress rule.

Topic-Focus interactions

- The common ground is represented by a set of file cards. Each file card represents a discourse referent (Reinhart 1981).

- The cards are organized so that the most recently activated cards are to be found on top of the stack of cards. These are the discourse referents which provide potential topics in the discourse.

How do cards get to be on top of the file?

Topic-Focus interactions

How do cards get to be on top of the file?

- If the attention of the hearer is drawn to (the referent of) X, then the hearer (metaphorically) selects the card for X and puts it in a place of prominence, namely on top of his stack of file cards.
- The card is selected from among the already existing file cards if it is definite and therefore represents an existing referent.
- The hearer is required to make out a new card for an indefinite. This card is again positioned on top of the stack.
- The file system involves locating cards on top of a stack (topics) or positioning them there (foci).

Illustration

A says:

(1) I $_{top}$[have a dog]$_{foc}$. [It]$_{top}$ [is brown]$_{foc}$.

B's update:

1. Pick the card for A_1 from the top of the file. (TOPIC rule)
2. Enter "e has a dog" on A's card. (UPDATE)
4. Open a new card, label it dog_2. Put it on top of the file (FOCUS rule i.)
5. Enter "A_1 has e" on this card. (UPDATE)

The following cards are now on top of the hearer's file and are available as future topics:

A_1
e_1 has dog_2

dog_2
A has e_2

Pronouns are Topics

2nd sentence in (1); B's update:

1. Pick card 2 from the top of the file. (TOPIC rule)
2. Enter "it is brown" on card 2. (UPDATE)

dog_2
A has e_2
it_2 is brown

- Pronouns are interpreted according to the heading of a topic card.
  - pronouns are necessarily topics.

(Erteschik-Shir 2014)

Main and subordinate IS

Q: What did John do with the dishes?

A: [he$_{\textbf{top}}$ [washed them$_{\text{top}}$]$_{\textbf{foc}}$]$_{\text{foc}}$

Each topic must have a focus associated with it.

Each top-foc pair form a f(ocus)-structure.

Main f-structure – the one for which the truth value is calculated. The remaining one(s) – *subordinate* .

Contrastive top & foc

- Contrast is contextually constrained to occur only if a contrast (2-member) set is available. Either focus or topic can be contrasted:

Q: Which laundry did John wash, the white, or the colored?
A: *He washed* the WHITE laundry.

B: Tell me about your brothers John and Bill.
A: JOHN is the smart one.

$[\{\text{John}_{foc}, \sout{\text{Bill}}\}_{top}]_{top}$ $[\text{is the smart one}]_{foc}$

- The contrastive element has a subordinate f-structure which includes both a topic and a focus. This is why contrastive elements can function both as topics and as foci.

Restrictive top & foc

- Restrictive foci, like contrastive ones, require a context-specified set (not restricted to 2 members):

  Q: Which one of his friends wants to meet John?

  B: JANET wants to meet John.

$[\{\text{Janet}_{foc}, \text{Mary, Bill}\ldots\}_{top}]_{foc}$ [wants to meet John]$_{foc}$

- The set of John's friends provides the topic of a (subordinate) f-structure in which "Janet" is the focus. Restrictive elements combine topic properties (range over a contextually specified set) and focus properties (one element of this set is focused).
- This is why they too can function both as topics and as foci.

Canonical information structure

- Canonical IS = the language-particular alignment of topic and focus with the linear array of PF. (Pertains to root clauses)
- Dependencies such as *wh*-movement, pairing in multiple *wh*-questions, negation and its scope are restricted to canonical IS. (Erteschik-Shir 1997)
- Languages differ with respect to their canonical IS and ways of topic-marking making different predictions for the availability of dependencies.

English canonical IS

- "Objects are harder to interpret as topics than subjects (in languages with fixed word order and no morphological marking of top/foc)"

Tell me about John.

a. He is in love with Mary.

b. ??Mary is in love with him

(E-S 2007: 166)

In languages like English there is preference for aligning f-structure with syntactic structure

English canonical IS

- The unmarked position of topics in English is the subject/stage and the unmarked location of the focus is (in) the VP domain.

  $\text{SUBJECT}_{\textbf{top}}\ [_{\text{VP}}\ldots\ \text{X}\ \ldots]_{\textbf{foc}}$

  $\text{sTOP}_{\text{t}}\ [\ldots\text{X}\ldots]_{\text{foc}}$

German canonical IS

According to Fanselow et al. 2008, object-initial sentences(b) are less acceptable than subject-initial ones (a):

- a Ein Schüler hat das Buch gefunden.
  A.nom pupil has the book found.
  'A pupil has found the book.'
  b Das Buch hat ein Schüler gefunden.

However, in context, when the initial object is interpreted as a topic, they are equally acceptable → both orders are canonical Iss

- German canonical IS: X_{top} V [... Y ...]$_{foc}$

Hebrew canonical IS

- Hebrew canonical IS: Topic<<Focus
- The focus can be the VP or an element following the verb

SVO

a Yossi$_{\textbf{top}}$ [$_{\textbf{VP}}$kara et hasefer haze]$_{\textbf{foc}}$ (What did Yossi do?)

b Yossi$_{\textbf{top}}$ [$_{\textbf{VP}}$kara [et hasefer haze]$_{\textbf{foc}}$] (What did Yossi read?)

Yossi read OM the-book this

'Yossi read this book'

OVS

[Et hasefer haze]$_{\textbf{top}}$ kara [YOSSI]$_{\textbf{foc}}$ (Who read this book?)

Topic marking

- English – in situ
- German (and Danish) – by topicalization to the single preverbal position.
- Hebrew – preverbal position, not limited to one.

I(dentificational)-dependencies

I-Dependencies such as *wh*-movement, pairing in multiple *wh*-questions, negation and its scope are restricted to canonical IS (Erteschik-Shir 1997)

Wh-trace and *operator-variable* dependencies

The dependent is identified in the construction either by its antecedent or by an operator.

- The processing load of a marked IS together with the processing of the dependency leads to processing overload. This predicts many well-known constraints on dependencies and their context sensitivity.

Constraints on extraction

Extraction from topics is blocked

Erteschik-Shir 1973: only focus domains are transparent for purposes of extraction [intuition: focus domains are processed differently from non-focus domains in terms of the visibility of gaps].

(10) a Who did John say that he had seen?
b ?Who did John mumble that he had seen?
c *Who did John lisp that he had seen?

(10b) is improved in a context in which "mumbling" has been mentioned (e.g., following "At our meetings everyone always mumbles"). (10c) is acceptable in a context in which it is known that John lisps. This is because such a context enables the main verb to be defocused and consequently enables the subordinate *that*-clause to be focused.

Variable acceptability

→Whether a particular subordinate clause provides a focus and therefore allows extraction depends, among other factors, on the clause under which it is embedded. It also depends on the context in which the sentence is uttered

→variable judgments for such sentences out of context

Extraction out of sentential subjects

- Always ungrammatical and can't be contextually ameliorated.

Why?

- Recall that I-dependencies can only be processed in canonical IS. Thus, the sentential subject should be topical

*Who is [that John likes t]$_{top}$ [interesting]$_{foc}$

- However, dependents are only visible in the focus domain and can't be identified within topics, resulting in ungrammaticality

Extraction out of sentential subjects

- Restrictive (D-linked) wh-phrase

*[Which girl]$_{top}$ [is [that John likes e]$_{top'}$ [interesting]$_{foc'}$]$_{foc}$

- Although reconstruction is not required in this case, extraction leaves the topic incomplete and disqualifies it as such.
- However, dependents are only visible in the focus domain and can't be identified within topics, resulting in ungrammaticality

Reminder

- The topic has wider scope than other elements in the sentence because the predicate of the sentence is evaluated with respect to the topic.

- Since the topic is contextually determined, scopal relations also depend on context. Scopal ambiguity is eliminated once a sentence is contextualized.

Scope calculation

- *Every man loves some woman*

a. $[\text{Every man}]_{\text{TOP}}$ $[\text{loves some woman}]_{\text{FOC}}$

$\forall x, \exists y$ (x loves y)

b. TOP_i [every man loves $[\text{some woman}]_i]_{\text{FOC}}$

$\exists y, \forall x$ (x loves y)

The reading in a. is much easier to get than the one in b.

This is because the reading in which the object is the topic is marked (does not adhere to the canonical IS)

TOPICS AND SCOPE

Since the topic is contextually determined, the scopal relations will also depend on context. Scopal ambiguity is therefore eliminated once a sentence is contextualized.

a. $[Q_1]_{TOP}$ $[V\ Q_2]_{FOC}$ $Q_1 > Q_2$

b. TOP_2 $[Q_1\ V\ Q_2]_{FOC}$ $Q_2 > Q_1$

c. $sTOP_t$ $[Q_1\ V\ Q_2]_{FOC}$ unscoped

Negation

- Pragmatic scope of negation = association with focus
- Sentential or VP focus → Negation has wide scope
- Focusing a constituent within S, VP→ narrow scope

Architecture of grammar

- IS is signaled by prosody, morphology and word order. It makes sense for all of these to be the responsibility of the externalization system. IS is a part of the phonological computation (E-S 2007:214)

- The externalization system PF/IS is responsible for linear (topological) order (Erteschik-Shir, various)

- EPP effects may follow from principles of syntax/IS alignment = canonical IS specified for each language.

  English: subject, stage=topic; predicate=focus

  Scandinavian: prefield = topic, predicate=focus

- Important: no need for LF → possible that the output of phonology that includes IS-processing is visible to both the conceptual-intentional and the articulatory-perceptual interfaces, capturing the fact that IS contributes both to interpretation and pronunciation.

Architecture of grammar

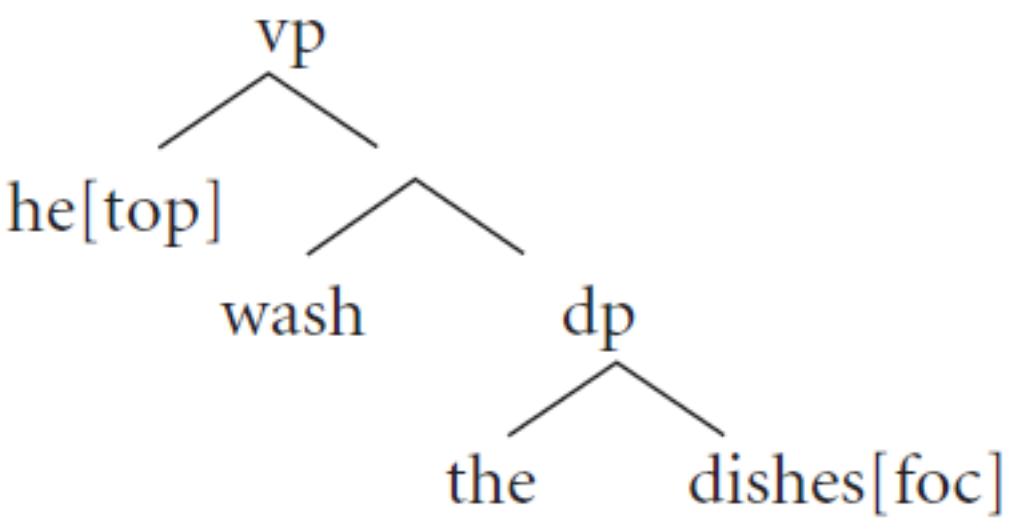
- "F-structure [Information structure] is a structural description, annotated for topic and focus which interfaces with syntax and both semantics and intonation. Due to topic marking this structural description is scopally unambiguous making LF redundant". (Erteschik-Shir 2007: 43)

- [top] and [foc] are features optionally assigned in the syntax (as part of lexical selection). Like φ-features, they may percolate to the maximal projection of the lexical item they are assigned to.

(97) a Q: What did John wash?
A: *He washed the dishes..*

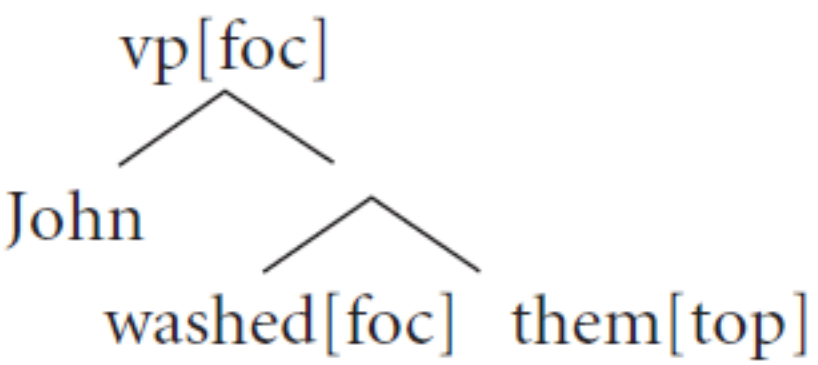
b select "dishes" → assign [foc]
select "the" → no assignment
select "wash" → no assignment
select "he" → assign [top]

c



(98) a What happened to the dishes?
John washed them.

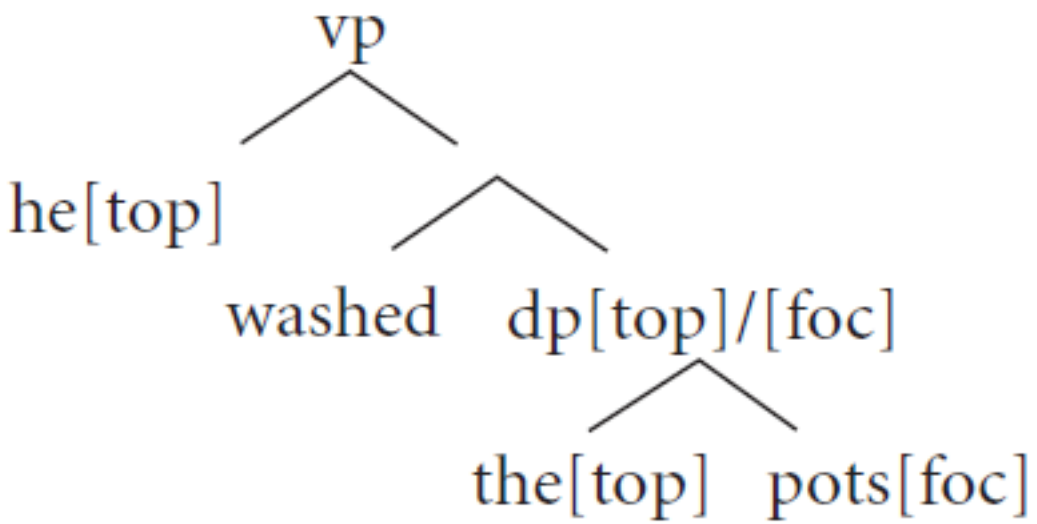
b select "them" → assign [top]
select "washed" → assign [foc]
select "John" → no assignment

c 

(99) a Q: What did John wash, the pots or the pans?
A: *He washed the pots (not the pans).*

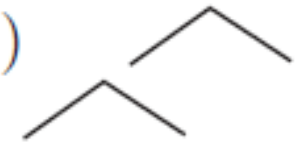
b select "pots" → assign [foc]
select "the" → assign [top]
select "washed" → no assignment
select "he" → assign [top]

c



Architecture schematically

Initial merge (incl. top/foc)



→ θ-assignments

→ Interpret

↓

→ dislocation

PF computation

→ morphology

→ phonology

References

- Erteschik-Shir, Nomi (1973). 'On the nature of island constraints'. Ph.D. M.I.T.
- --- (1997). *The Dynamics of Focus Structure.* Cambridge: Cambridge University Press.
- --- (2007). *Information Structure: The Syntax-Discourse Interface.* ed. Robert Van Valin. Syntax and Morphology; Oxford: OUP.
- Erteschik-Shir, Nomi and Lappin, Shalom (1979). 'Dominance and the functional explanation of island phenomena', *Theoretical Linguistics.* 6, 41-85.
- Gundel, Jeanette K. (1974). 'The role of topic and comment in linguistic theory'. Ph.D. dissertation (University of Texas).
- Reinhart, Tanya (1981). 'Pragmatics and linguistics: an analysis of sentence topics', *Philosophica.* 27, 53-94.
- Strawson, P. F. (1964). 'Identifying reference and truth-values', *Theoria.* 30, 86-99.

Nature of null topics

- Overt and null pronouns are computed somewhat differently: Whereas a fully specified feature bundle is computed syntactically in the case of overt pronouns, null topics are merged as a set of unvalued ϕ- features

Nature of null topics

Null arguments enter the computation as a bundle of unvalued ϕ-features

α	person
β	number
γ	gender

The PF interface accesses the discourse file and valuation of the feature bundle occurs by matching with a (topic) card from the top of the file, thereby also providing a reference. The assumption that the features are unvalued, forces a process of valuation where the only source of such valuation is the discourse. It is therefore not necessary to mark missing arguments with a topic feature. Topichood follows from the need for valuation.

(E-S, Ibnbari, Taube 2013: 148)

Overt pronouns

- Overt pronouns are merged as a bundle of valued ϕ-features:
  - ± Person
  - ± Number
  - ± Gender

In view of the fact that their features are valued, they can receive their reference from an argument with matching features in the sentence as well as from an available topic card. In fact, it is well known that unless the context forces it, pronouns necessarily find their antecedents within the sentence.

(E-S, Ibnbari, Taube 2013: 149)