

Information- and discourse-structure analysis with questions under discussion

Session 4: Topics

Kordula De Kuthy, Eberhard-Karls-Universität Tübingen

Arndt Riester, Universität Bielefeld

September 2022

Summer School *Corpus Annotation and Data Analysis (CANDA)*

Georg-August-Universität Göttingen

- Topic notions
- Centering Theory, types of topic transitions (Daneš 1974, Grosz et al. 1995) from one discourse unit to the next
- Corpus of Spanish dialogue on resolving a murder case (www.sgscorpus.com, Adli 2011) [joint work with Amalia Canes Nápoles, University of Cologne]
- How to extend QUD-tree analyses with topical information
 - Sentence topics (Reinhart 1981, Lambrecht 1994)
 - Discourse topics / topics under discussion (Van Dijk 1977, van Kuppevelt 1995, Asher 2004)
 - Towards capturing topic transitions

Topic transitions and Centering

Phenomenon I – Marked topic shift

- (1) [Speaker A] Q₃₈: ¿y Ø_T era más o menos de su misma edad?
and was:3SG more or less of POSS same age
'and was he more or less of the same age as him?'
- [Speaker B] A₃₈: Ø_T un poquito más joven.
a little more young
'he was slightly younger'
- [Speaker A] Q₄₀: ¿y a qué se dedicaba [este señor]?_T
and to what 3SG.REFL dedicate this man
'and what did this man do for a living?'
- [Speaker B] A₄₁: [este señor]_T es el regidor de cultura y
this man is the councillor for culture and
urbanismo de Barcelona.
urban planning of Barcelona
'this man is the councillor for culture and urban planning
of Barcelona'

(sgs Spanish 52/74-78)

(2) [Speaker A] Q₁₂: ¿y tenían hijos o?
'And they had children, right?'

[Speaker B] A₁₂: sí
'yes'

A₁₃: ∅_T tenían dos hijos_F.
'they had two children'

[Speaker A] Q₁₅: ¿∅_T mayores?
grown.up
'Are they grown up?'

(sgs Spanish 52/24-26)

While it is generally agreed that **topics** are what sentences (discourses, books, movies, ...) are **about** (Hockett 1958, Strawson 1964, Reinhart 1981, Lambrecht 1994, Krifka 2008), many sub-concepts have been suggested:

- sentence topic (Reinhart 1981)
- discourse topic (e.g. Asher 2004)
- familiar topic (Frascarelli and Hinterhölzl 2007)
- contrastive topic (e.g. Büring 2003)
- shifted topic (e.g. Givón 1976)
- new topic (Reinhart 1981)
- frame-setting topic (Chafe 1976)
- relevance topic (Repp 2011)
- backward-looking center (Grosz et al. 1995)
- topic time (Klein 1992)
- ...

- It is not entirely clear how all these notions relate to one another.
- Previous attempts at annotating topics (e.g. Cook and Bildhauer 2013) experienced difficulties:

- It is not entirely clear how all these notions relate to one another.
- Previous attempts at annotating topics (e.g. Cook and Bildhauer 2013) experienced difficulties:
 - partly due to conflicting features: morphosyntactic (e.g. subjecthood, other topic positions), semantic (e.g. animacy) and pragmatic (e.g. givenness);
 - partly due to lack of discourse structure

Grosz et al. (1983, 1995), Grosz and Sidner (1986), Brennan et al. (1987), and Walker et al. (1998)

- Framework for the description of the “focus of attention” (topic, center), the choice of referring expressions, and the coherence of discourse
- Part of a comprehensive and influential theory of discourse structure

1. LINGUISTIC STRUCTURE: Discourse segments and their relations
2. INTENTIONAL STRUCTURE: Communicative goal (and sub-goals) of a discourse (semantic structure)
3. ATTENTIONAL STATE: Dynamic status of what counts as salient at a certain point in the discourse

Intentional structure

- Every discourse has a goal (**intention**), called **discourse purpose (DP)**.

Intentional structure

- Every discourse has a goal (**intention**), called **discourse purpose (DP)**.
- Every discourse segment has a goal as well, the **discourse segment purpose (DSP)**.

Intentional structure

- Every discourse has a goal (**intention**), called **discourse purpose (DP)**.
- Every discourse segment has a goal as well, the **discourse segment purpose (DSP)**.
- In multi-speaker discourses (conversations), intentions belong to the different speakers.
- There is a close correspondence between these DSPs and our **questions under discussion**

Intentional structure

- Every discourse has a goal (**intention**), called **discourse purpose (DP)**.
- Every discourse segment has a goal as well, the **discourse segment purpose (DSP)**.
- In multi-speaker discourses (conversations), intentions belong to the different speakers.
- There is a close correspondence between these DSPs and our **questions under discussion**
- Intentional structure can be modelled as a **stack**. Intentions are dynamically pushed onto and removed from the stack, once they have been fulfilled.

Intentional structure

- Every discourse has a goal (**intention**), called **discourse purpose (DP)**.
- Every discourse segment has a goal as well, the **discourse segment purpose (DSP)**.
- In multi-speaker discourses (conversations), intentions belong to the different speakers.
- There is a close correspondence between these DSPs and our **questions under discussion**
- Intentional structure can be modelled as a **stack**. Intentions are dynamically pushed onto and removed from the stack, once they have been fulfilled.
- Intentional structure can be evaluated in terms of **global coherence**.

- Centering is a model that also keeps track of **local coherence** and the **attentional state**, in connection with the use of different referring expressions.
- Different types of anaphors (pronouns, definites, demonstratives, ellipses) give rise to different inferences.
- For each referring expression, antecedents have a different likelihood.

- (3)
- a. John went to his favorite music store to buy a piano.
 - b. He had frequented the store for many years.
 - c. He was excited that he could finally buy a piano.
 - d. He arrived just as the store was closing for the day.
- (4)
- a. John went to his favorite music store to buy a piano.
 - b. It was a store that John had frequented for many years.
 - c. He was excited that he could finally buy a piano.
 - d. It was closing just as John arrived.

Center: John

- (5)
- a. John went to his favorite music store to buy a piano.
 - b. He had frequented the store for many years.
 - c. He was excited that he could finally buy a piano.
 - d. He arrived just as the store was closing for the day.

No clear center, “aboutness” is constantly changing.

- (6)
- a. John went to his favorite music store to buy a piano.
 - b. It was a store that John had frequented for many years.
 - c. He was excited that he could finally buy a piano.
 - d. ?It was closing just as John arrived.

- (7)
- a. Terry really goofs sometimes
 - b. Yesterday was a beautiful day and he was excited about trying out his new sailboat.
 - c. He wanted Tony to join him on a sailing expedition.
 - d. He called him at 6 a.m.
 - e. He was sick and furious at being woken up so early.

- (8)
- a. Terry really goofs sometimes
 - b. Yesterday was a beautiful day and he was excited about trying out his new sailboat.
 - c. He wanted Tony to join him on a sailing expedition.
 - d. He called him at 6 a.m.
 - e. **He** was sick and furious at being woken up so early.

- (9)
- a. Terry really goofs sometimes
 - b. Yesterday was a beautiful day and he was excited about trying out his new sailboat.
 - c. He wanted Tony to join him on a sailing expedition.
 - d. He called him at 6 a.m.
 - e. **Tony** was sick and furious at being woken up so early.
 - f. He told Terry to get lost and hung up.
 - g. Of course, he hadn't intended to upset Tony.

- (10)
- a. Terry really goofs sometimes
 - b. Yesterday was a beautiful day and he was excited about trying out his new sailboat.
 - c. He wanted Tony to join him on a sailing expedition.
 - d. He called him at 6 a.m.
 - e. Tony was sick and furious at being woken up so early.
 - f. He told Terry to get lost and hung up.
 - g. Of course, **he** hadn't intended to upset Tony.

- (11)
- a. Terry really goofs sometimes
 - b. Yesterday was a beautiful day and he was excited about trying out his new sailboat.
 - c. He wanted Tony to join him on a sailing expedition.
 - d. He called him at 6 A.M.
 - e. Tony was sick and furious at being woken up so early.
 - f. He told Terry to get lost and hung up.
 - g. Of course, Terry hadn't intended to upset Tony.

- Each utterance U is assigned a set of **forward-looking centers** $C_f(U)$ and a single **backward-looking center** $C_b(U)$ (\approx sentence topic).
- The elements of $C_f(U)$ are partially ordered according to their prominence.
- $C_b(U_{n+1})$ connects with the most highly ranked element of $C_f(U)$ that is realized in U_{n+1} .
- Different connection types

1. CENTER CONTINUATION: $C_b(U_{n+1}) = C_b(U_n)$, and this entity is the most highly ranked element of $C_f(U_{n+1})$

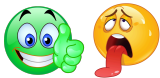


Basic transition relations

1. CENTER CONTINUATION: $C_b(U_{n+1}) = C_b(U_n)$, and this entity is the most highly ranked element of $C_f(U_{n+1})$



2. CENTER RETAINING: $C_b(U_{n+1}) = C_b(U_n)$, but this entity is not the most highly ranked element in $C_f(U_{n+1})$



Basic transition relations

1. CENTER CONTINUATION: $C_b(U_{n+1}) = C_b(U_n)$, and this entity is the most highly ranked element of $C_f(U_{n+1})$



2. CENTER RETAINING: $C_b(U_{n+1}) = C_b(U_n)$, but this entity is not the most highly ranked element in $C_f(U_{n+1})$



3. CENTER SHIFTING: $C_b(U_{n+1}) \neq C_b(U_n)$



Claims of original Centering Theory

- There is a unique C_b .
- Elements of C_f are partially ordered according to a number of factors, in particular grammatical role (SUBJ > OBJ > OTHER).
- Centering constrains realization possibilities.
- There is a preference among transitions.

Which transition relations apply?

- (12)
 - a. Susan gave Betsy a pet hamster.
 - b. She reminded her that such hamster were quite shy.
 - c. She asked Betsy whether she liked the gift.
- (13)
 - a. Susan gave betsy a pet hamster.
 - b. She reminded her that such hamsters were quite shy.
 - c. Betsy told her that she really liked the gift.
- (14)
 - a. Susan gave Betsy a pet hamster.
 - b. She reminded her that such hamsters were quite shy.
 - c. Susan asked her whether she liked the gift.
- (15)
 - a. Susan gave Betsy a pet hamster.
 - b. She reminded her that such hamsters were quite shy.
 - c. She told Susan that she really liked the gift.

Constraints on center movement and realization

Rule 1: If any element of $C_f(U_n)$ is realized by a pronoun in U_{n+1} , then the $C_b(U_{n+1})$ must be realized by a pronoun also.

- (16) a. John has been acting quite odd. [→ CONTINUE →]
b. He called up Mike yesterday. [C_b = John]
c. John wanted to meet him urgently. [C_b = John]

Violation of Rule 1 → Incoherence

- (17) a. John has been acting quite odd.
b. He called up Mike yesterday.
c. #He was annoyed by John's call.
- (18) a. John has been acting quite odd.
b. He called up Mike yesterday. [$C_b = \text{John}$] [→ SHIFT →]
c. Mike was studying for his driver's test. [$C_b = \text{Mike}$]
[→ CONTINUE →]
d. He was annoyed by John's call.



Constraints on center movement and realization

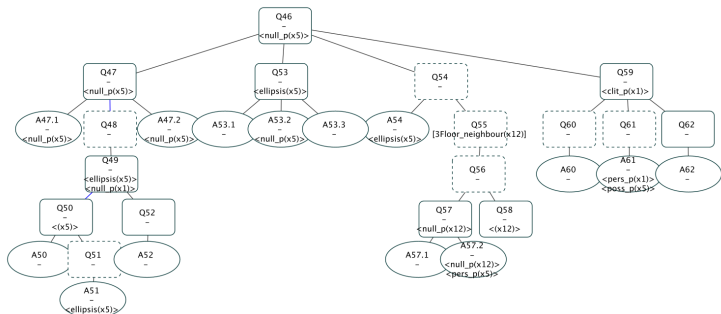
Rule 2: Preference: CONTINUE > RETAIN > SHIFT (cognitive load increases from left to right)

- (19) a. John has been having a lot of trouble arranging his vacation.
b. He cannot find anyone to take over his responsibilities.
[$C_b = \text{John}; C_f = \{\text{John}\}$]
[→ CONTINUE →]
c. He called up Mike yesterday to work out a plan.
[$C_b = \text{John}; C_f = \{\text{John}, \text{Mike}\}$]
[→ RETAIN →]
d. Mike has annoyed him a lot recently.
[$C_b = \text{John}; C_f = \{\text{Mike}, \text{John}\}$]
[→ SHIFT →]
e. He called John at 5 A.M. on Friday last week.
[$C_b = \text{Mike}; C_f = \{\text{Mike}, \text{John}\}$]

Topics and topic shifts in QUD trees

QUD trees

- Goal: identify a QUD  for each utterance / assertion / elementary discourse unit 
- QUDs reflect the **“topical organization”** of a discourse, and determine the **information structure** (background-focus) of each discourse unit.



Why a QUD analysis to approach topics?

1. Clear approach to **discourse segmentation**: scope of a topic
2. Determination of **background-focus structure**: focal elements are, by definition, no topic candidates.
3. **Hierarchical discourse structure** (in contrast to linear analysis):
 - Coordinated structures / parallelism indicate **topic continuity**
 - Availability of a **right frontier** of potential topics

- Segmentation is oriented along the lines of **information units**.
- An information unit is a segment of discourse that:
 - can function as the **answer to a question**.
 - is semantically optional / independent from the preceding or upcoming information.
- In particular, **adjuncts** and **conjuncts** typically form separate – elliptical – discourse units (cf. Brunetti et al. 2021).

(20) A₇₅: and when she arrived,
she found him like this,
A₇₆: **with a bag over his
head**

(21) A_{195.1}: he was found **naked**
A_{195.2}: **with underpants**
A_{195.3}: **tied to the bed**

Two ways of QUD-identification

(i) Givenness-based:

- Orientation towards previous discourse
- Identify **given** (\approx topic / background) vs. **new** (\approx focus) information

Q₁₁₂: And did **they_T** ever **consider** a **divorce**?

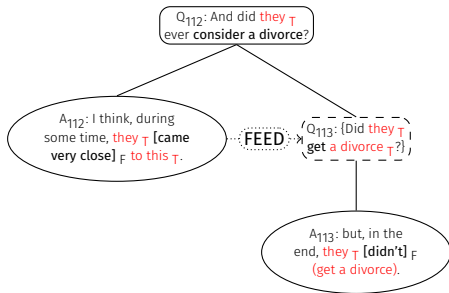
A₁₁₂: I think, during some time, **they_T** [**came very close**]_F **to this_T**.

A₁₁₃: but, in the end, **they_T** [**didn't**]_F.

Two ways of QUD-identification

(i) Givenness-based:

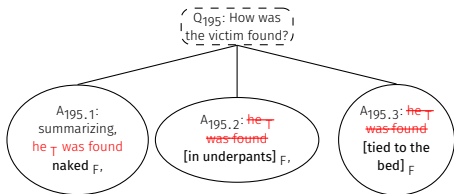
- Orientation towards previous discourse
- Identify **given** (\approx topic / background) vs. **new** (\approx focus) information
- Antecedent A_{112} **feeds** the subsequent question Q_{112} (van Kuppevelt 1995).



Claim: Given material cannot attach higher than its antecedent.

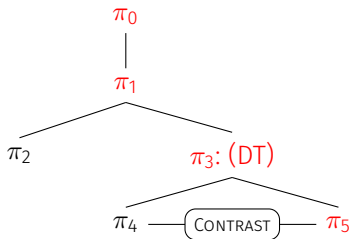
(ii) Based on parallelism:

- Orientation towards upcoming discourse
- Discourse coordination
- QUD is defined by (semantically) shared material.
- Givenness principle does not apply.



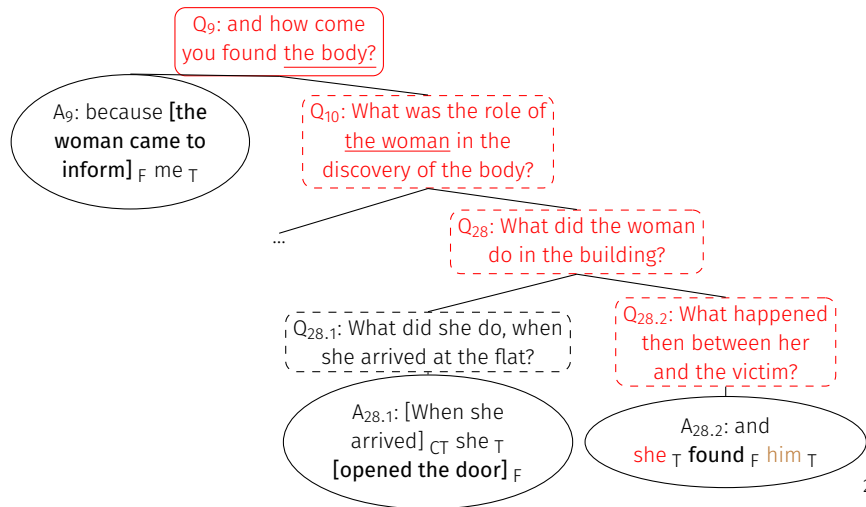
The Right Frontier

- **Right Frontier Constraint (RFC)** (Polanyi 1988, Asher 1993, Asher and Lascarides 2003, Prévot and Vieu 2008) defines the available attachment points for a new discourse unit:
 - The previous discourse unit, and
 - any unit dominating it.
- SDRT discourse units can be **simple** or **complex** (i.e. dominated by a “discourse topic” (DT), cf. Asher 2004: p. 171)



Right Frontier of QUDs

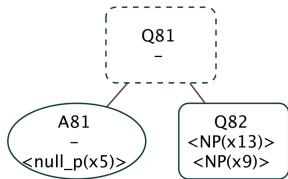
- Non-terminal nodes of QUD trees are always questions.
- Therefore, the Right Frontier also generally consists of questions (ranging over discourse topics).



Topics within QUDs

Sentence topic (Reinhart 1981)

- What a **discourse segment** is about
- One (or several) sentence topic(s) per segment
- **Semantically** present, typically mentioned in the segment (except for ellipsis and null pronouns)
- Expression is often subject to topic marking (e.g. left dislocation of *la puerta* in Q₈₂)



x5:victim, x13:door, x9:flat

(22) > Q₈₁: {In what state was the body?}

[Spk. B] >> A₈₁: Ø_T estaba atado a la cama
was tied to the bed
'He was tied to the bed.'

[Spk. A] >> Q₈₂: ¿y la puerta_T por eso estaba abierta la del
and the door for that was opened the.one of.the
piso_T?
flat

'And is that why the door of the apartment was open?'

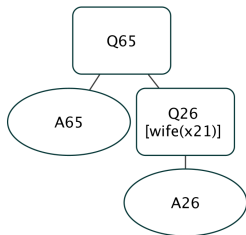
Set of topics under discussion:

- The set of all entities and concepts that are currently *topical* at a certain point in discourse
- Members of this set remain ...
 - under discussion (Van Kuppevelt's (1995) *active/open questions*)
 - cognitively *active* (Chafe 1994)
 - *at the right frontier*:
the latest topic and topic(s) in dominating node(s).
- Set grows with increasing depth of the current discourse tree
- NB: There is not only one such discourse topic, but many!
- Typically, **each one has scope over an entire discourse section, and over other topics under discussion.**

Two interpretations of “topic”

refers previously under discussion:

[speaker(x1), victim(x3), death(x4,x3),...]



Note that the discourse topic 'mujer' is declared in the question Q₂₆, not earlier; since in A₆₅ it is still focus.

- (23) [Spk. A] Q₆₅: ¿y quién dio el primer aviso del fallecimiento de la víctima?
and who give.PST the first notice of.the death of the victim

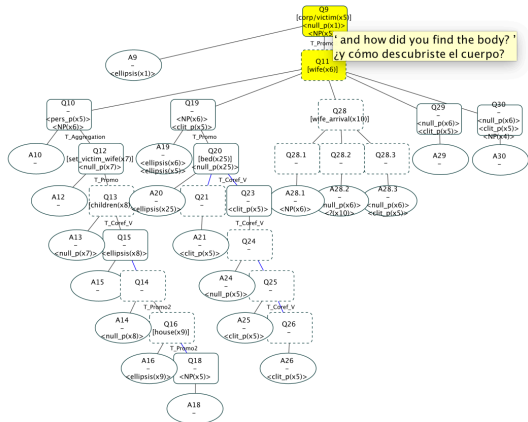
And who was the first to report the victim's death?

[Spk. B] > A₆₅: [su mujer]F.
his wife

[Spk. A] > Q₂₆: ¿o sea y su mujer estaba en casa?
so and his wife be.PST at home?

So his wife was at home?

Topics under discussion – Labelling conventions



Format: [topics under discussion]

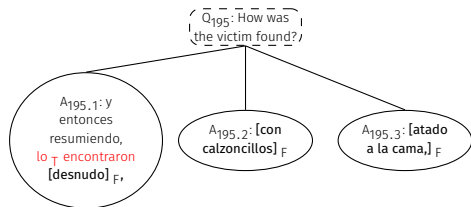
- Specified on the first use of a discourse entity as a topical entity within an implicit or explicit QUD.
- Example:
 - Q₉: [body/victim(x5)]
 - Q₁₁: [wife(x6)]



Transitions

Topic continuity

- Parallel assertions share the same QUD / topic referent
- Here: victim(x5)



	A _{195.1}	A _{195.2}	A _{195.3}
C _f :	{x5}	{x5, calz.(x101)}	{x5, cama(x102)}
C _b :	-	x5	x5

(24) Q₁₄: {What about the children?}

> A₁₄: ∅_T [ya no viven **en casa**]_F.
 already no 3SG.live in house

‘they do not live in the house any more’

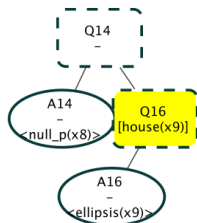
> Q₁₆: {Who lives in **the house**?}

>> A₁₆: o sea, ∅_T vivían sólomente la mujer y
 i.e live.3SG.PST only the woman and

el marido.
 the husband

‘that is, only the wife and the husband lived in the house’

	A ₁₄	A ₁₆
C _f :	{children(x8), house(x9) }	{x9, wife(x7), husbandx5}
C _b :	x8	x9



(25) [Spk. A] Q₁₂: ¿y tenían hijos o?
 'They had children, right?'

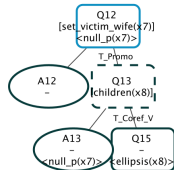
[Spk. B] > A₁₃: sí/Yes

> Q₁₃: {How many children did the victim
 and the wife have?}

[Spk. B] >> A₁₃: tenían dos hijos_F.
 'they had two children'

[Spk. A] >> Q₁₄: ¿Ø_T mayores?
 grown up

'Are the children grown up?'



	Q ₁₂	A ₁₃	Q ₁₄
C _f :	{victim+wife(x7)}	{x7, children(x8)}	{x8}
C _b :	x7	x7	x8

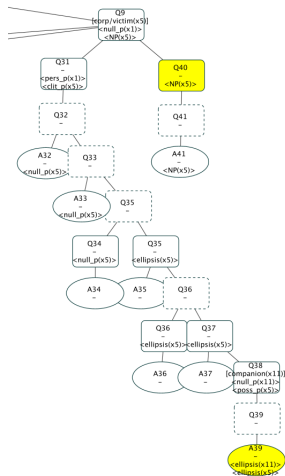
Topic shift

- (26) [Spk. A] Q₃₉: ¿y \emptyset_T era más o menos de
 and was.3SG more or less of
 su misma edad?
 POSS same age
 ‘and was he more or less of the same
 age as him?’







[Spk. B] > A₃₉: \emptyset_T un poquito más joven.
 a little more young
 ‘he was slightly younger’







- [Spk. A] Q₄₀: ¿y a qué se dedicaba [este
 and to what 3SG.REFL dedicate this
 señor?]_T
 man
 ‘and what did this man do for a living?’








	Q ₃₉	A ₃₉	Q ₄₀
C _f :	{companion(x11), victim(x5)}	{x11}	{x5}
C _b :	x11	x11	x5










References

-  Adli, A. 2011. Gradient acceptability and frequency effects in information structure: A quantitative study on Spanish, Catalan, and Persian. Habilitationsschrift. Universität Freiburg.
-  Asher, N. 1993. *Reference to abstract objects in discourse*. Dordrecht: Kluwer.
-  Asher, N. 2004. Discourse topic. *Theoretical Linguistics* 30(2-3). 163–202.
-  Asher, N. and A. Lascarides. 2003. *Logics of conversation*. Cambridge University Press.
-  Brennan, S., M. Friedman and C. Pollard. 1987. A centering approach to pronouns. In *25th annual meeting of the association for computational linguistics*, 155–162.
-  Brunetti, L., K. De Kuthy and A. Riester. 2021. The information-structural status of adjuncts: A question-under-discussion-based approach. *Discours* 28. 3–48. <https://doi.org/10.4000/discours.11454>.

-  Buring, D. 2003. On D-trees, beans, and B-accent. *Linguistics and Philosophy* 26(5). 511–545.
-  Chafe, W. 1976. Givenness, Contrastiveness, Definiteness, Subjects, Topics and Point of View. In C. N. Li (ed.), *Subject and topic*, 25–55. New York: Academic Press.
-  Chafe, W. 1994. *Discourse, consciousness, and time*. University of Chicago Press.
-  Cook, P. and F. Bildhauer. 2013. Identifying ‘aboutness topics’: Two annotation experiments. *Dialogue and Discourse* 4(2). 118–141.
-  Daneš, F. 1974. Functional Sentence Perspective and the organization of text. In F. Daneš (ed.), *Papers on Functional Sentence Perspective*, 2017–2222. Prague & The Hague: Academia, Mouton.
-  Frascarelli, M. and R. Hinterhölzl. 2007. Types of topics in German and Italian. In S. Winkler and K. Schwabe (eds.), *On information structure, meaning and form*, 87–116. Amsterdam: Benjamins.

-  Givón, T. 1976. Topic, Pronoun, and Grammatical Agreement. In C. Li (ed.), *Subject and topic*, 151–188. New York: Academic Press.
-  Grosz, B., A. Joshi and S. Weinstein. 1983. Providing a unified account of definite noun phrases in discourse. In *Proceedings of the 21st annual meeting of the association for computational linguistics*.
-  Grosz, B., A. Joshi and S. Weinstein. 1995. Centering: A framework for modelling the local coherence of discourse. *Computational Linguistics*. 203–225.
-  Grosz, B. and C. Sidner. 1986. Attention, intentions, and the structure of discourse. *Computational Linguistics* 12(3). 175–204.
-  Hockett, C. 1958. *A course in modern linguistics*. New York: Macmillan.
-  Klein, W. 1992. The present perfect puzzle. *Language*. 525–552.
-  Krifka, M. 2008. Basic notions of information structure. *Acta Linguistica Hungarica* 55(3-4). 243–276.

-  Lambrecht, K. 1994. *Information structure and sentence form*. Cambridge University Press.
-  Polanyi, L. 1988. A formal model of the structure of discourse. *Journal of Pragmatics* 12(5-6). 601–638.
-  Prévot, L. and L. Vieu. 2008. The moving right frontier. In A. Benz and P. Kühnlein (eds.), *Constraints in discourse*, 53–66. Amsterdam: Benjamins.
-  Reinhart, T. 1981. Pragmatics and linguistics: An analysis of sentence topics in pragmatics and philosophy I. *Philosophica* 27(1). 53–94.
-  Repp, S. 2011. Relevance topics. In *Proceedings of sinn und bedeutung*, vol. 15, 483–498.
-  Strawson, P. 1964. Identifying reference and truth-values. *Theoria* 30. 96–118.
-  Van Dijk, T. A. 1977. Sentence topic and discourse topic. *Papers in Slavic Philology* 1. 49–61.



van Kuppevelt, J. 1995. Discourse structure, topicality and questioning. *Journal of Linguistics* 31. 109–149.



Walker, M., A. Joshi and E. Prince. 1998. Centering in naturally occurring discourse. In *Centering theory in discourse*. Oxford University Press.