Factors influencing implicit discourse relations in Czech

Šárka Zikánová

Charles University
Faculty of Mathematics and Physics
Czech Republic, Prague

Introduction

The explicitness and implicitness of discourse relations occur in a language in interplay with many other phenomena.

Previous research: Zikánová et al., TSD 2019

- Implicitness and negation (the more negation, the higher explicitness)

- Implicitness and text genres (some text genres are more implicit than the others)

- Implicitness and semantics of discourse relations (some senses are more likely to be implicit, differences even between close semantic types - synchrony, asynchrony)

Explicit and implicit coherence relations: Different, but how exactly?
Berlin, Humboldt University, January 17-18, 2020

Research questions

Current analysis – work in progress: Berlin 2020

- Implicitness and frequency of discourse relations

- Implicitness and specificity of the semantics of discourse relations

- Other signals of the discourse relations

Research questions

Discourse relations:

TEMPORAL
- Synchrony
- Asynchrony

COMPARISON
- Concession
- Confrontation (typical DC: *whereas;* within the arguments, different features are ascribed to entities which are expressed contrastively)
    A is big, whereas B is small..

Hypotheses and questions

1. Relations with a lower frequency have a higher measure of explicitness. These relations are not usual; therefore they need to be expressed explicitly with a discourse connective, in order to be understood.

2. Relations with a marked, specific semantics have a higher measure of explicitness than relations with more vague, unspecified semantics.

Higher specificity	Lower specificity
Rather explicit?	**Rather implicit?**
Comparison: concession	Comparison: confrontation
Temporal: synchrony	Temporal: asynchrony
COMPARISON as a group	TEMPORAL as a group

3. Implicit discourse relations (without a DC): are there other signals of the discourse relation present in the structure?

Enriched Discourse Annotation of Prague Discourse Treebank Subset (PDiT-EDA 1.0)
- Explicit discourse relations (primary and secondary DCs)
- Implicit discourse relations (without DCs)
- Entity-based relations
- Question-answer relations and isolated questions
- Macrostructure
- Other phenomena (attribution)

Annotation scenario: slightly modified PDTB approach
Based on the Prague Dependency Treebank, other types of annotations:
Syntax, syntactic semantics (tectogrammatics), information structure, coreference and bridging anaphora .

Extent of the corpus:
2592 sentences (41877 tokens)
100 written Czech journalistic texts
15 text genres (as an interview, a weather forecast, a news report etc.)
The size of the corpus allows applying χ^2 test reliably.

Explicit and implicit coherence relations: Different, but how exactly?
Berlin, Humboldt University, January 17-18, 2020

Results

	Explicit			Implicit			TOTAL
	Intra-S	Inter-S	Total	Intra-S	Inter-S	Total	
TEMPORAL	26	33	59	4	56	60	119
Synchrony	8	4	12	1	4	5	17
Asynchrony	18	29	47	3	52	55	102
COMPARISON	45	30	75	21	56	77	152
Confrontation	12	19	31	20	46	66	97
Concession	33	11	44	1	10	11	55
TOTAL	71	63	134	25	112	137	271

Table: Distribution of explicit and implicit discourse relations of synchrony, asynchrony, confrontation and concession in the PDiT-EDA 1.0 (instances)

Explicit and implicit coherence relations: Different, but how exactly?
Berlin, Humboldt University, January 17-18, 2020

Implicitness and frequency of discourse relations

Hypothesis:
Relations with a lower frequency have a higher measrure of explicitness.

Proved on our data (χ2 test): the rarer relations are rather expressed explicitly.

Higher occurrence:
Asynchrony	54% implicit
Confrontation	68% implicit

Lower occurrence:
Synchrony	29% implicit
Concession	20% implicit

Implicitness and frequency of discourse relations

	Explicit			Implicit			TOTAL
	Intra-S	Inter-S	Total	Intra-S	Inter-S	Total	
TEMPORAL	26	33	59	4	56	60	119
Synchrony	8	4	**12**	1	4	**5**	**17**
Asynchrony	18	29	**47**	3	52	**55**	**102**
COMPARISON	45	30	75	21	56	77	152
Confrontation	12	19	**31**	20	46	**66**	**97**
Concession	33	11	**44**	1	10	**11**	**55**
TOTAL	71	63	134	25	112	137	271

Table: Distribution of explicit and implicit discourse relations of synchrony, asynchrony, confrontation and concession in the PDiT-EDA 1.0 (instances)

Implicitness and semantics of discourse relations

Hypothesis:
Relations with a marked, specific semantics have a higher measure of explicitness than relations with vague, unspecified semantics.

Higher specificity	Lower specificity	
Rather explicit?	**Rather implicit?**	
Comparison: concession **80%**	Comparison: confrontation **32%**	Tendency: yes χ2 test: significant
Temporal: synchrony **71%**	Temporal: asynchrony **46%**	Tendency: yes χ2 test: not significant (low number of occurrences of synchrony)
COMPARISON as a group **56%**	TEMPORAL as a group **44%**	Tendency: yes χ2 test: not significant

Implicitness and semantics of discourse relations

	Explicit			Implicit			TOTAL
	Intra-S	Inter-S	Total	Intra-S	Inter-S	Total	
TEMPORAL	26	33	**59**	4	56	**60**	**119**
Synchrony	8	4	**12**	1	4	**5**	**17**
Asynchrony	18	29	**47**	3	52	**55**	**102**
COMPARISON	45	30	**75**	21	56	**77**	**152**
Confrontation	12	19	**31**	20	46	**66**	**97**
Concession	33	11	**44**	1	10	**11**	**55**
TOTAL	71	63	134	25	112	137	271

Table: Distribution of explicit and implicit discourse relations of synchrony, asynchrony, confrontation and concession in the PDiT-EDA 1.0 (instances)

11

Other means of signalling semantics of implicit discourse relations in Czech

TEMPORAL RELATIONS:

Asynchrony:

- Verbal aspect: perfective
- Explicit temporal settings (world knowledge): *last year – this year*
  *I was lucky to have survived. **Upon my return,** I learned that Karel Poláček also went to Auschwitz.*
- Shift of the actors – reaction of another person in the sequence of events
  *From the burning of stolen wallets, a penetrant smell streaked around their cottage. **The security councilor** tried to fine the Dunka family at least for the smell.*

Synchrony:

Verbal aspect: imperfective

Temporal relations: verbal aspect

Synchrony: imperfective, asynchrony: perfective

- IMPF – IMPF: *Školník opravoval světlo na chodbě. V aule zněl zpěv.*
The school janitor was repairing the light in the corridor. There was singing in the auditorium./ The singing sounded in the auditorium.
synchrony

- IMPF – PF: *opravoval – zazněl*
synchrony, the second event happened suddenly

- PF – PF: *opravil – zazněl*
asynchrony

- PF – IMPF: *opravil – zněl*
unclear; the light was already repaired when there was singing (?) Not enough infomation, confusing.

Temporal relations: verbal aspect

BUT:

IMPF – IMPF – IMPF – IMPF – IMPF
Děti se rodily, rostly, vdávaly a ženily a měly svoje děti.
Children were born, grew, married and had their children.
Asynchrony, habitual

PF – PF – PF
Já ti tám! Já ti naučím! Já ti tfakrát oslatím!
I will show you! I will teach you (how to live)! I will sweeten your life twice!
equivalence

Other means of signalling semantics of implicit discourse relations in Czech

COMPARISON RELATIONS:

Confrontation:
- Word order: contrastive topic units on the left periphery
- Intonation: contrastive stress
- Pronouns: strong forms

Concession:
- Intontation: highly marked (emphasis, astonishment)

Confrontation: word order

Contrastive topic units on the left periphery
Contrastive stress

Jakubovi *odkážu dům,* ***Anežce*** *dám pole.*
Lit.: ***To_Jakub*** *I_will_bequeath the_house,* ***to_Anežka*** *I_will_give the_field.*
As for Jakub, I will give him the house, (whereas) Anežka will get the field.

Standard neutral word order: the contrastivity gets lost.
*? Odkážu Jakubovi dům, dám Anežce pole.

Confrontation: strong forms of pronouns

Strong form: *tobě - to you,* dativ. (Weak, enclitic form: *ti)*

Jakubovi *odkážu dům,* **tobě** *nedám nic.*
Lit: To_Jakub I_will_bequeath the_house, to_you I_will_not _give anything.
As for Jakub, I will give him the house, (whereas) you will get nothing.

***Jakubovi** odkážu dům, nedám **ti** nic.*
*weak form in the confrontation

17

Concession: intonation

1 Honza byl na sebe vždycky hrozně opatrný.
2 Odjakživa měl potravinovou alergii.
3 Na vernisáži jedl lískové oříšky.
4 Teď opuchl a proklíná umění.

1 Honza was always very careful about himself. - context
2 He had always had a food allergy.
3 He ate hazelnuts at the vernissage.
4 Now he swelled and curses art. - context

Implicit concession needs marked intonation expressing astonishment
(formal feature) or explanation in the context.

Conclusions

Implicitness is inter-related with the frequency of discourse relations.

Implicitness is inter-related with the specificity of discourse relations:
-   proved for concession and confrontation
-   not proved for synchrony and asynchrony

Other signals:
-   clearer for confrontation (strong pronoun forms, typical development of information structure)
-   weaker for temporal relations (verbal aspect)
-   almost not found for concession.

Future work

Psycholinguistic experiment:

Implicit and explicit synchrony/asynchrony and concession/confrontation
Self-paced reading
Basic hypothesis: explicit discourse relations are easier to process (faster reading)

Zikánová 2020, Societas Linguistica Europaea, under review:

Implicitness and intra-/inter-sententiality
        Related to different semantic types of discourse relations

Implicitness and signalling of the text macrostructure

References

- Asr F. T. and Demberg V. 2012. Implicitness of discourse relations. 24th International Conference on Computational Linguistics (Coling), Bombay, India.

- Jínová P., Poláková L. and Mírovský J. 2014. Sentence Structure and Discourse Structure (Possible Parallels). In K. Gerdes, E. Hajičová and L. Wanner (eds.), Dependency Linguistics. Recent Advances in Linguistic Theory Using Dependency Structures. Amsterdam: Benjamins, 53–74.

- Louis A., Joshi A., Prasad R., Nenkova A.. 2010. Using entity features to classify implicit discourse relations. Proceedings of the 11th Annual Meeting of the Special Interest Group on Discourse and Dialogue. 978-1-932432-85-5. Tokyo, Japan, pp. 59-62, Association for Computational Linguistics

- Poláková L., Mírovský J., Synková P.. 2017. Signalling Implicit Relations: A PDTB - RST Comparison. Dialogue and Discourse, 8:2, pp. 225-248.

- Sanders, T., Demberg, V., Hoek, J., et al. (2018). Unifying dimensions in coherence relations: How various annotation frameworks are related . Corpus Linguistics and Linguistic Theory, 0(0), pp. -. Retrieved 2 Jan. 2020, from doi:10.1515/cllt-2016-0078

21

References

- Zeyrek D. and Kurfalı M. 2018. An Assessment of Explicit Inter- and Intra-sentential Discourse Connectives in Turkish Discourse Bank. In Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018), Miyazaki, Japan. European Language Resources Association (ELRA). https://www.aclweb.org/anthology/L18-1634

- Zikánová Š. et al. 2015. Discourse and Coherence. From the Sentence Structure to Relations in Text. Prague: Institute of Formal and Applied Linguistics.

- Zikánová Š., Mírovský J., Synková P. 2019. Explicit and Implicit Discourse Relations in the Prague Discourse Treebank. In: Lecture Notes in Computer Science, Vol. 11697, Proceedings of the 22nd International Conference on Text, Speech and Dialogue - TSD 2019, Springer International Publishing, Cham / Heidelberg / New York / Dordrecht / London, ISBN 978-3-030-27946-2, ISSN 0302-9743, pp. 236-248.

- Zikánová Š., Synková P., Mírovský J. 2018. Enriched Discourse Annotation of PDiT Subset 1.0 (PDiT-EDA 1.0). Data/software, Charles University, Prague, Czech Republic, http://hdl.handle.net/11234/1-2906.

Acknowledgement

Explicit and implicit coherence relations: Different, but how exactly?
Berlin, Humboldt University, January 17-18, 2020

Special thanks

To my colleague Jiří Mírovský for the great cooperation.

To Bonnie Webber and Jet Hoek for the inspiration.

To Debopam Das for organizing this workshop.

To the reviewers of the abstract for their important comments.

Explicit and implicit coherence relations: Different, but how exactly?
Berlin, Humboldt University, January 17-18, 2020