

On Definition of Discourse Connectives: Primary vs. Secondary Connectives (Based on a Corpus Probe)

Magdaléna Rysová, Kateřina Rysová

magdalena.rysova@ufal.mff.cuni.cz, rysova@ufal.mff.cuni.cz

Charles University in Prague, Faculty of Mathematics and Physics, Institute of Formal and Applied Linguistics

Aim of Research

The aim of our presentation is to contribute to the general discussion on discourse connectives, especially on their definition and principles we may hold as boundaries surrounding this class of expressions.

Methods and Material

The analysis was based on the data from the *Prague Dependency Treebank* (PDT; almost 50 thousand sentences).

Results of Research

Definition of Discourse Connectives: Universality Principle

1. According to their function in the text, discourse connectives serve as indicators of discourse relations;
2. Concerning their semantic nature – the status of discourse connectives must be universal (cf. universal *this is the reason why* vs. non-universal *this increase is the reason why*).

Discourse Connectives: Primary vs. Secondary

Primary connectives	Secondary connectives
synsemantics	structures with autosemantic basis
lexically frozen (grammaticalized)	open or fixed collocations (non-grammaticalized)
non-modifiable (with exceptions)	modifiable (with exceptions)
mainly one-word	mainly multiword
universal	universal
not sentence elements	sentence elements, clause modifiers or separate sentences
	convey anaphoric reference to the 1st argument
	Uniqueness of some structures:
	a) syntactically higher than the 2nd argument
	b) form of a separate sentence
	c) nominalization of the 2nd argument

1) Synsemantics (*and, but, or...*; mainly conjunctions and structuring particles) vs. connecting structures with autosemantic components (*the reason is, this means...*).

2) Primary connectives are lexically frozen (i.e. grammaticalized), secondary connectives form open or fixed collocations (that are not grammaticalized).

3) Secondary connectives mostly allow a free modification (*the main/only/first/important reason is etc. vs. *generally and*); primary connectives are modifiable only exceptionally (*simply because*).

4) Primary connectives are mainly one-word expressions, secondary connectives form mostly multiword connecting structures.

5) Primary connectives (as synsemantics like conjunctions *and, but* etc. or structuring particles like *too, only* etc.) are not sentence elements, secondary connectives are sentence elements (like *because of this*), sentence modifiers (like *simply speaking*) or form a separate sentences (*The reason is simple*).

Uniqueness of Some Structures with Secondary Connectives

1) Secondary connectives may form a separate sentence: *I will not go to school tomorrow. The reason is easy (= SC). I am ill.*

2) Secondary connectives may be syntactically higher than the second argument of the relation (i.e. the second argument is syntactically dependent on the secondary connective expressed by the main clause) – *I will not go to school tomorrow. The reason is (= SC, main clause) **that I am ill** (= the second argument, nominal content subordinate clause).*

3) Some secondary connectives allow nominalization of the second discourse argument: *The management of the club dismissed the coach. The reason are disappointing results (= NP).*

Borderlines within Connecting Expressions

grammaticalization

Primary Con. ← Secondary Con. ← Non-Connectives
universality

Therefore ← because of this ← because of his late arrival

Origin of Primary Connectives through Grammaticalization of Secondary Connectives

E.g. *because* arose from *bi cause* „by cause“, originally a phrase often followed by a subordinate clause, one word probably from around 1400.

Data sources

Bejček, E. et al. (2012). Prague Dependency Treebank 2.5 – extended version of PDT 2.0. In: *Proceedings of the 24th International Conference on Computational Linguistics (Coling 2012)*, pp. 231–246. Mumbai, India.