The linear structuring of complex sentences

Holger Diessel and Karsten Schmidtke
University of Jena

holger.diessel@uni-jena.de
http://www.holger-diessel.de/

Summary

	Head-initial	Head-final
N-REL	x	x
REL-N		x
V-COMP	x	x
COMP-V		x
S-ADV	x	x
ADV-S	x	

Competing motivations

- The correlation with other dependent categories can be explained by the preference for consistent branching directions (i.e. Minimize Domains).
- The asymmetries arise from competing forces.

Competing motivations

- Relative clauses are overall more frequent after the noun they modify because the pronominal element a relative clause includes is easier to process if it follows the lexical antecedent, i.e. the head.
- Complement clauses are often extraposed to the position after the matrix verb so as to reduces the dependency domain.
- The positioning of subject complement clauses seems to be determined by information structure: Initial subject complement clauses resume information from the previous discourse, whereas extraposed complement clauses typically provide new information.
- Adverbial clauses are overall more frequent in the position before the main clause, because they are commonly used to provide a thematic ground for subsequent clauses.

Adverbial clauses

- Temporal clauses
- Conditional clauses
- Causal clauses
- Purpose clauses
- Result clauses

Adverbial clauses

In languages with initial and final adverbial clauses, different semantic types of adverbial clauses tend to occur in different positions vis-à-vis the main clause.

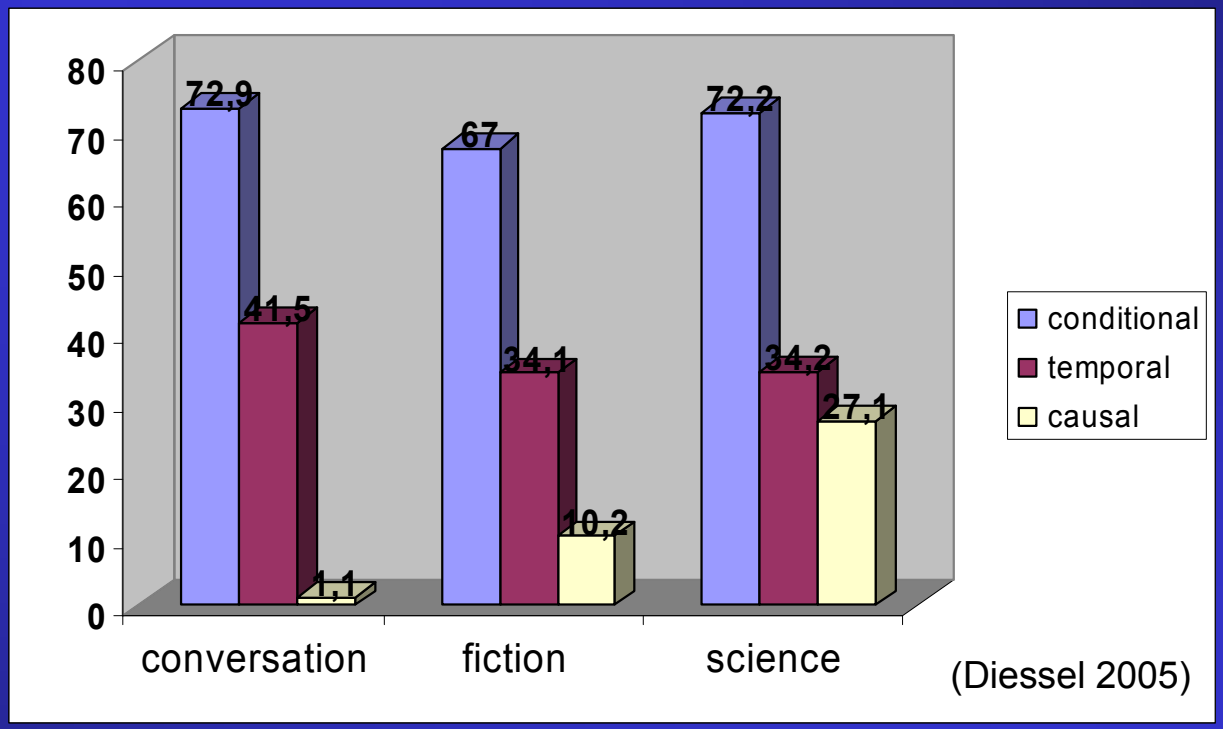
Adverbial clauses

	Conditional	Temporal	Causal	Purpose
Babungo	initial	initial+final	final	final
Supyire	initial	initial+final	final	final
Wari’	initial	initial+final	final	final
Kwami	initial	initial+(final)	final	final
Kera	initial	initial	final	final

Data

Genre	Source	Number of speakers	Total number of utterances
Conversations	Santa Barbara Corpus	15	388
Fiction	Short stories from British / American authors	15	878
Scientific writings	Academic articles from the journal *Cognition*	15	768

(Diessel 2005)



(Diessel 2005)

Friedman's ANOVA

conversations
$\chi 2 = 21.65$; $p < .001$
fiction
$\chi 2 = 26.53$; $p < .001$
scientific writing
$\chi 2 = 20.44$; $p < .001$

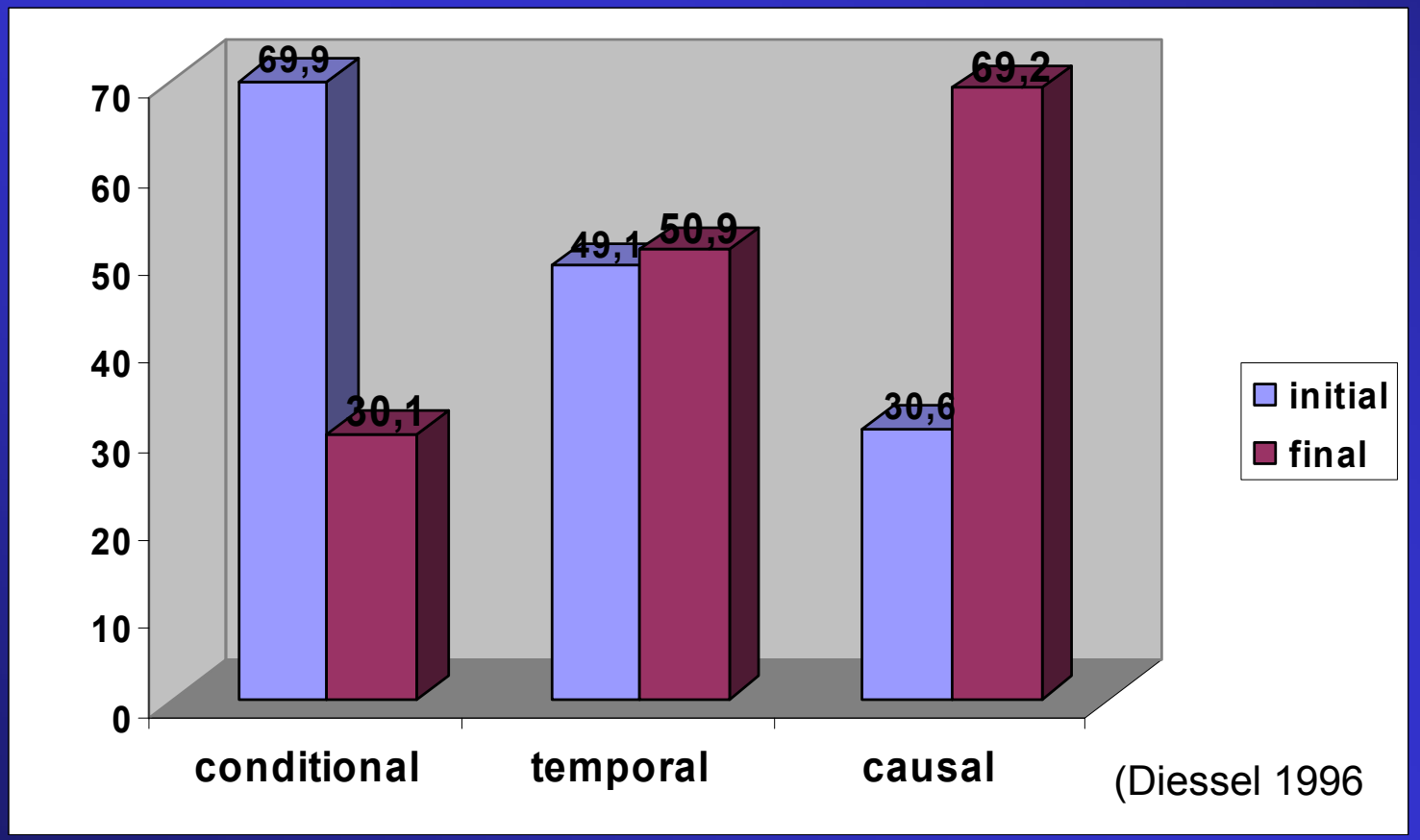
conditional & temporal

conversations	$Z = 2.897$, $p < .002$
fiction	$Z = 3.408$, $p < .001$
scientific writing	$Z = 3.351$, $p < .001$

temporal & causal

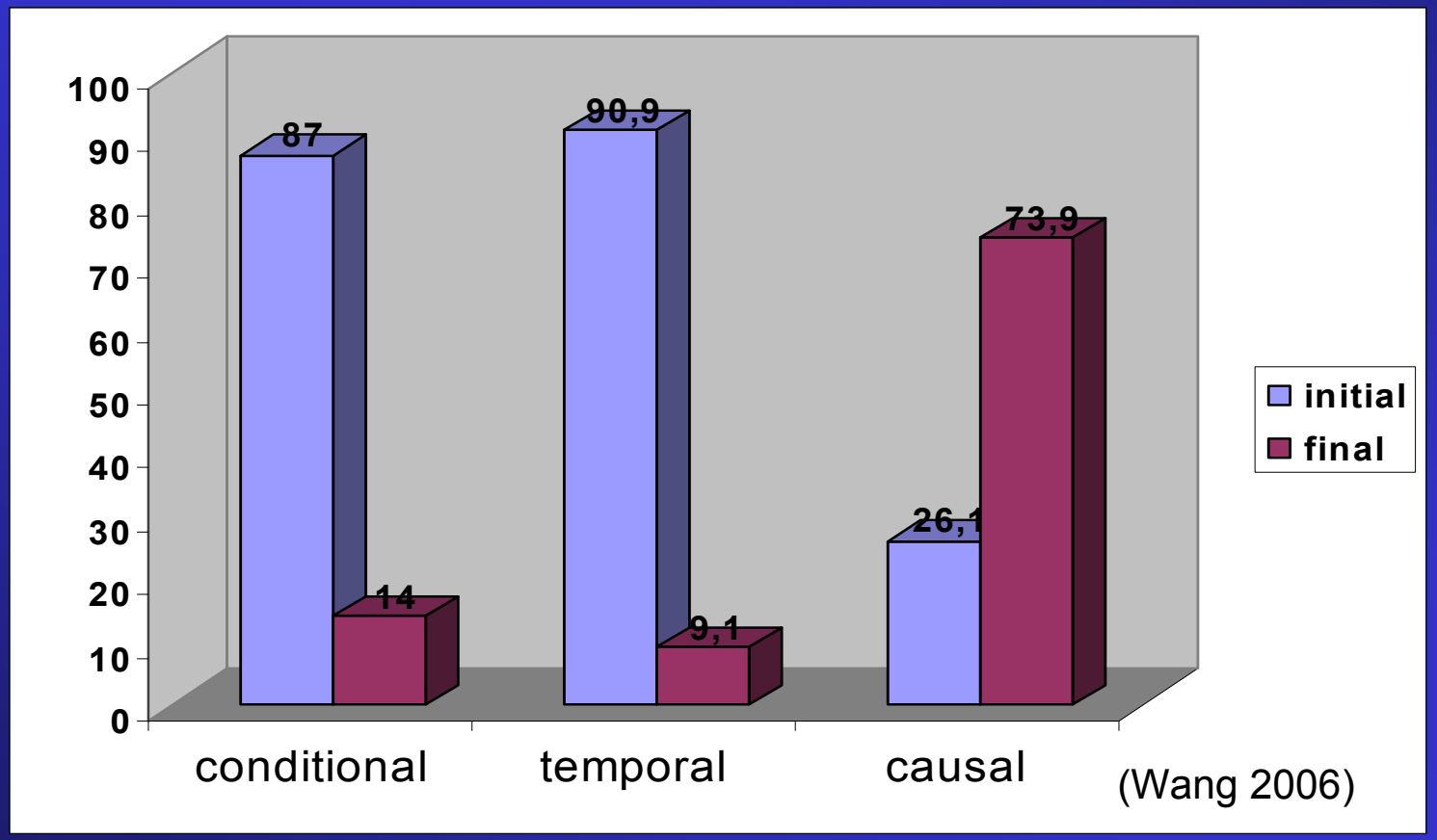
conversations	$Z = 3.111$, $p < .003$
fiction	$Z = 3.010$, $p < .003$
scientific writing	$Z = 1.250$, $p > .211$

German



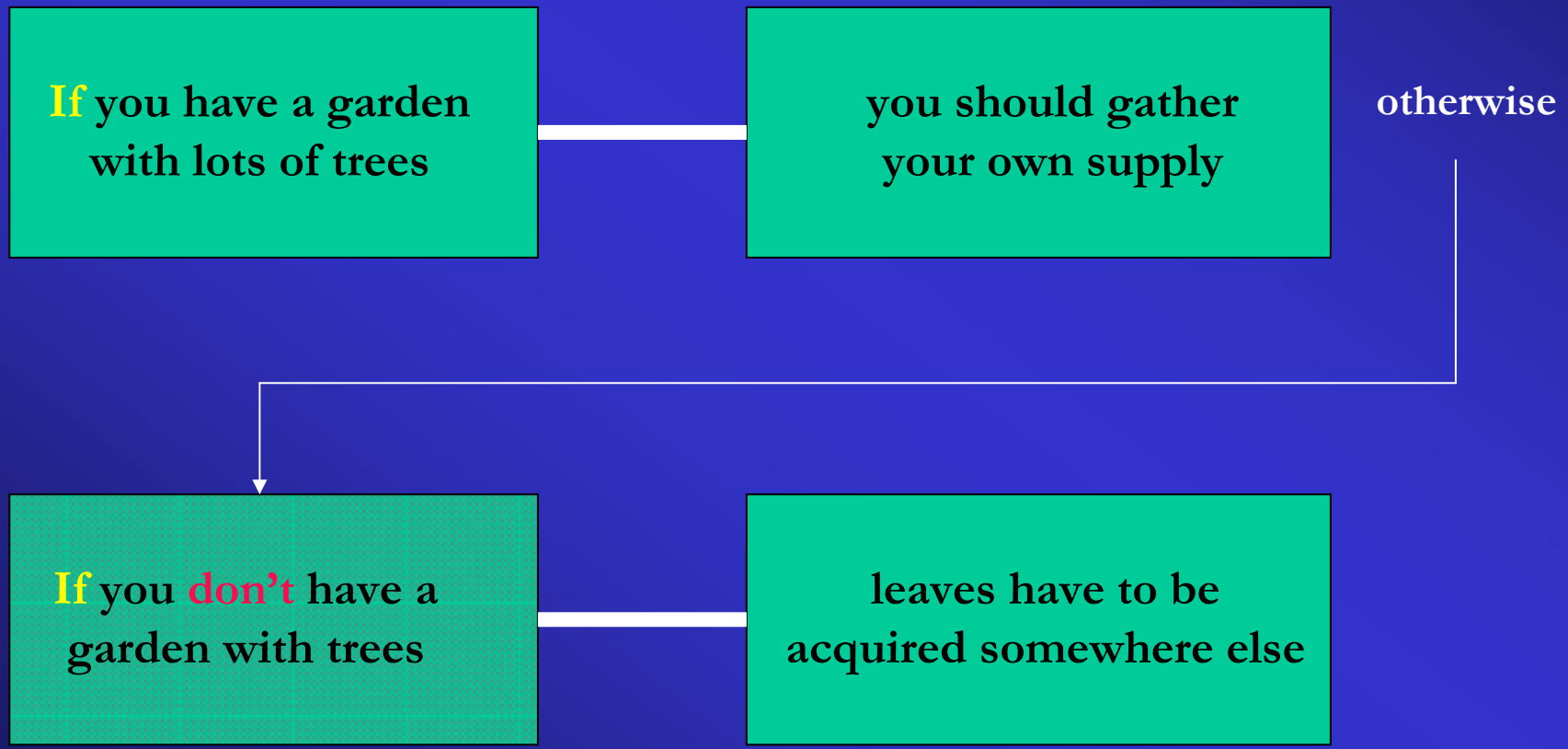
($\chi 2$=54.7, df =2, p<0.001)

Mandarin Chinese



($\chi 2$=177.43, df=2, p<0.001)

Conditional clauses



Conditional clauses

(1) The sentence can ***only*** be assigned the right truth condition, or alternatively be given the correct semantic representation, **if** the grammatical significance of 'and'… is taken into account.

(2) I **wouldn't** be sick **if** I were, excuse me, … pregnant.

(3) I will take the big one, … **if** you don't mind.

(4) I guess we ought to put those in the oven, **if** we're gonna eat them.

Causal clauses

The predominant use of causal clauses after the main clause results from a communicative (i.e. rhetorical) strategy in which causal clauses function to support a controversial statement.

Causal clauses

Because the climate is changing there will more floods in the future.

Causal clauses

(1) I .. played with them all week long,
which was really stupid,
… because they got worked up.

(2) ... And me and mom always accused her of being lazy.
... You know,
… because she was just, ... all she did was sleep.

(3) I realize it takes two to three weeks to process,
but just tell me whether it's on file.
…because if not, I want her to have another one

Causal clauses

A: We could spend a lot of our life trying to contradict that.
B: Why?
A: Well, **because** … it may be a very bad chemical bath.

Causal clauses

(1) a. I guess we should call off the picnic **because** it's raining, isn't it?
b. *I guess we should call off the picnic **if** it's raining, isn't it?

(2) a. I'm leaving **because** here comes my bus.
b. *I'm leaving **if** here comes the bus.

(3) a. The Knicks are going to win, **because** who on earth can stop Bernard?
b. *The Knicks are going to win, **if** who on earth can stop Bernard?

Causal clauses

(1) *Because it is raining, isn't it, I guess we should call off the picnic.

(2) *Because here comes the sun, I'm leaving.

(3) *Because who on earth can stop Bill, the Knicks are going to win.

Causal clauses in German

Der hat sicher wieder gsoffen;
He has probably again drunken

(0.3) **weil** … sie läuft total deprimiert durch die Gegend
(0.3) because she run totally depressed through the area

'He is probably drunk again, because she is running around totally depressed.'

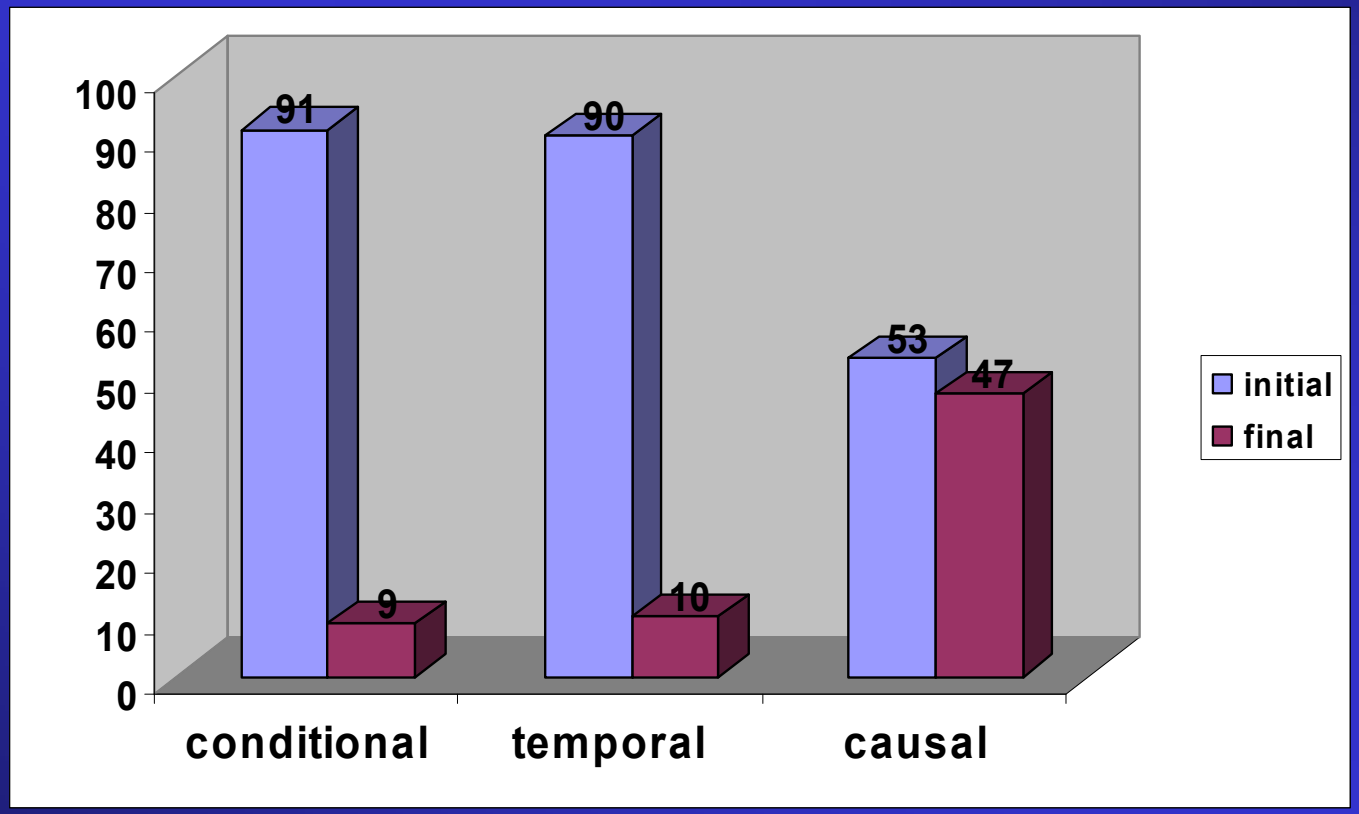
(Keller 1993; Günther 1996; Pasch 1997; Gohl and Günther 1999; Uhmann 1998)

Causal clauses in German

(1) Ich mache das nicht, **weil** dazu habe ich einfach keine Lust.

(2) *__Weil__ dazu habe ich einfach keine Lust, mache ich das nicht.

Causal clauses in Japanese



(Ford and Mori 2005)

Causal clauses

Causal relationships can also be expressed by coordinate sentences. In fact, in some languages causal clauses can only be realized by coordinate constructions, i.e. these languages do not have causal adverbial clauses.

Multifactorial analysis

The usage-based approach is a competition model. Multifactorial statistical models can determine the relative strength and interaction between the competing force.

Temporal clauses

Iconicity of sequence: Clause order tends to mirror the order of the events they describe.

Temporal clauses

(1) We shall make up our mind *when* the IMF has reported. [prior]

(2) They had already made breaches in the defensive wall of sand […] *when* the order came. [posterior]

(3) I did cook occasionally, *when* they were out. [simultaneous]

Temporal clauses

(1) a. *After* her father died, of course, Isabel's trust fund included quite a substantial holding in the company. [prior]

b. I put Emily back in her own bed, *after* she'd fallen asleep. [prior]

(2) a. *Before* the debt crisis set in, Brazil was enjoying growth rates of 7 percent per year. [posterior]

b. The heat […] from the sun is retained by the earth for a while, *before* it's radiated away. [posterior]

Temporal clauses

(1) a. *Once* the problem became clear, policy was tightened. [prior]

b. We'll be pretty busy *once* our course gets back into full swing. [prior]

(2) a. *Until* I'd spoken to William Davis I'd no idea that the monarchy was the only bright spot on our horizon. [posterior]

b. There should be no further cuts in interest rates, *until* the underlying rate of inflation begins to tumble. [posterior]

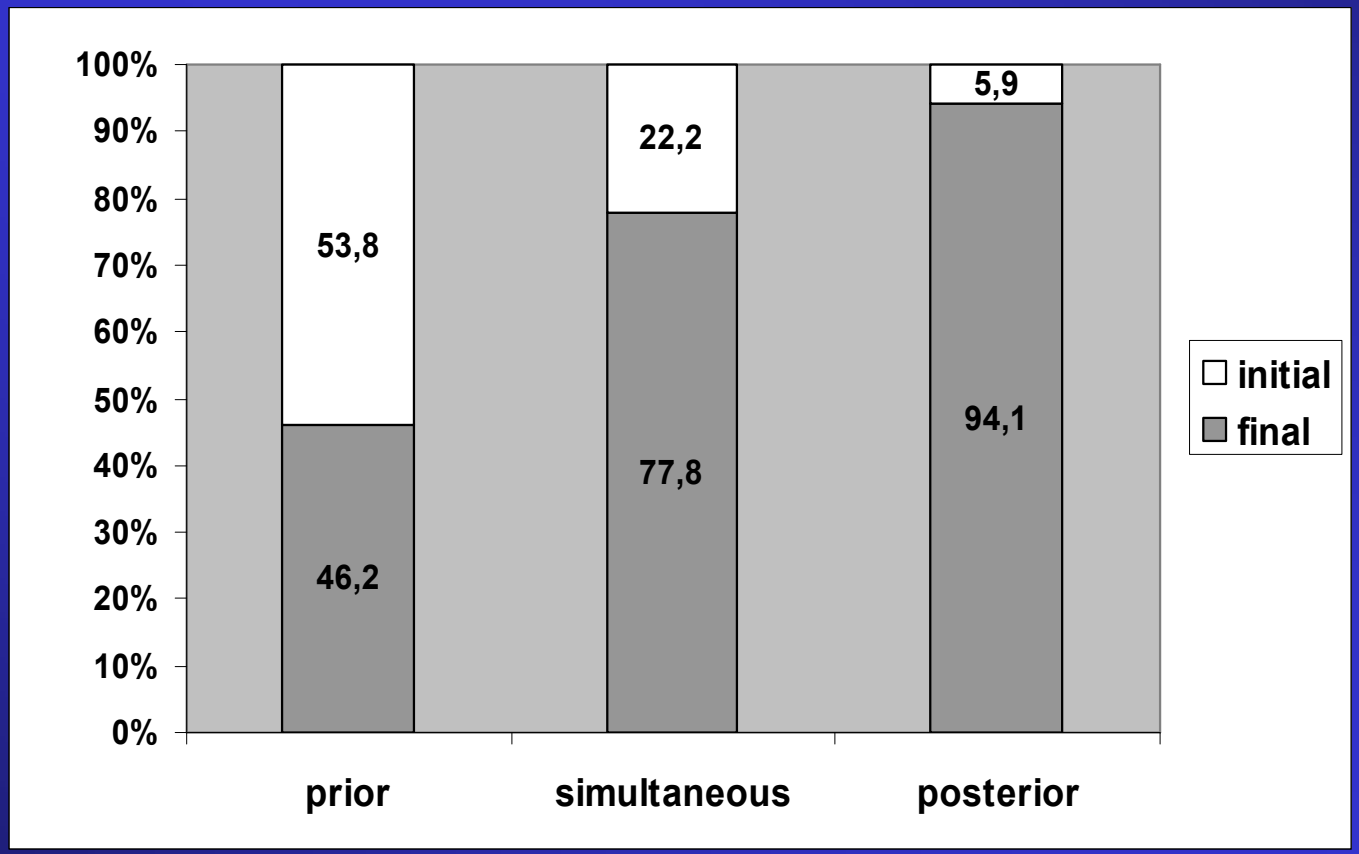
Temporal clauses

- *When*-clauses may have a conditional interpretation
- *After*-clauses are sometimes interpreted with a causal connotation
- *Before*-clauses can express a purpose or goal
- *Once*-clauses are often conditional
- *Until*-clauses may express a combination of time, purpose, and result.

Data (ICE-GB)

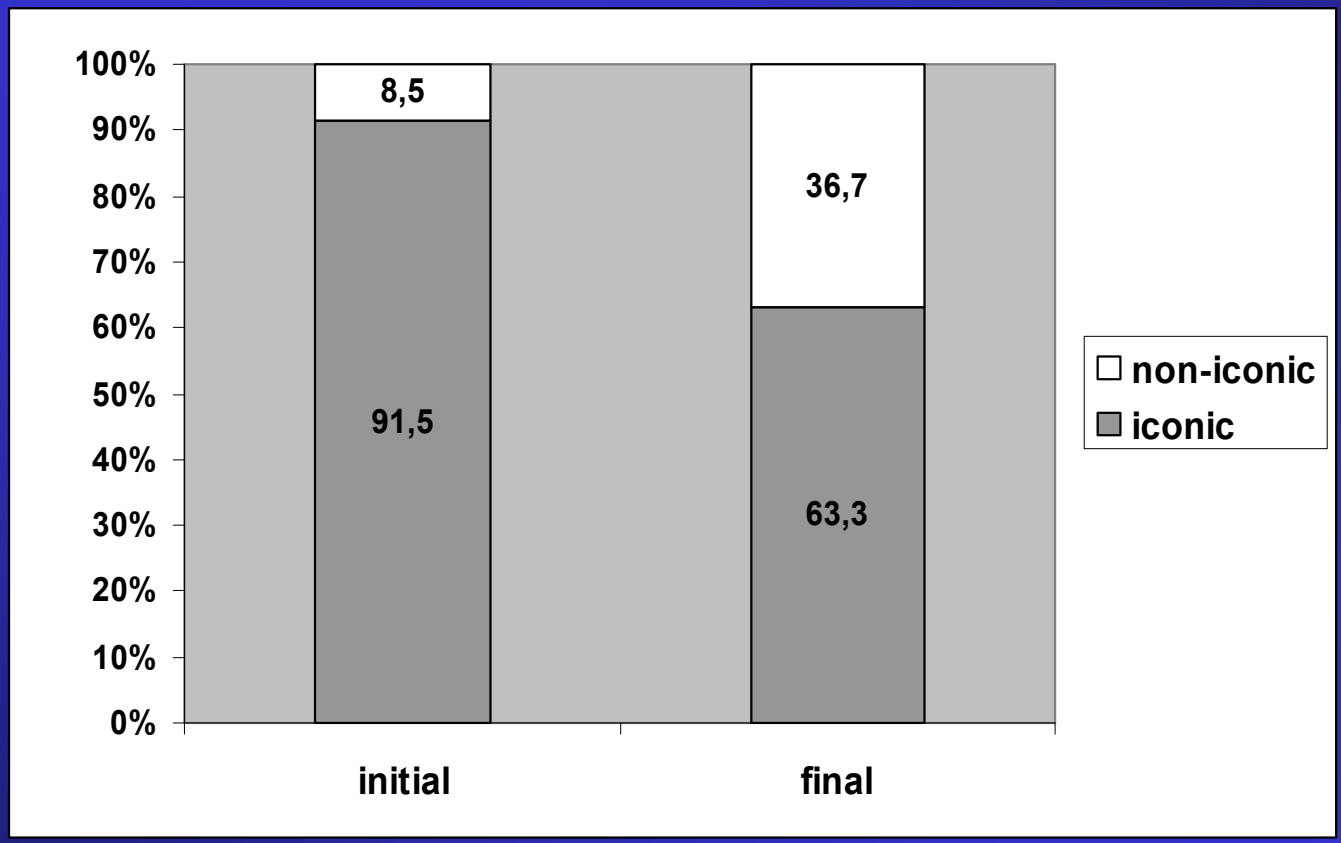
	Spoken	Written	Total
When	94	95	189
After	47	50	97
Before	41	46	87
Once	48	50	98
until	49	50	99
Total	297	291	570

Monofactorial analysis



($\chi 2 = 185.13$, $df = 2$, $p < 0.001$)

Monofactorial analysis



($\chi 2 = 35.25$, $df = 1$, $p < 0.001$)

Multifactorial analysis

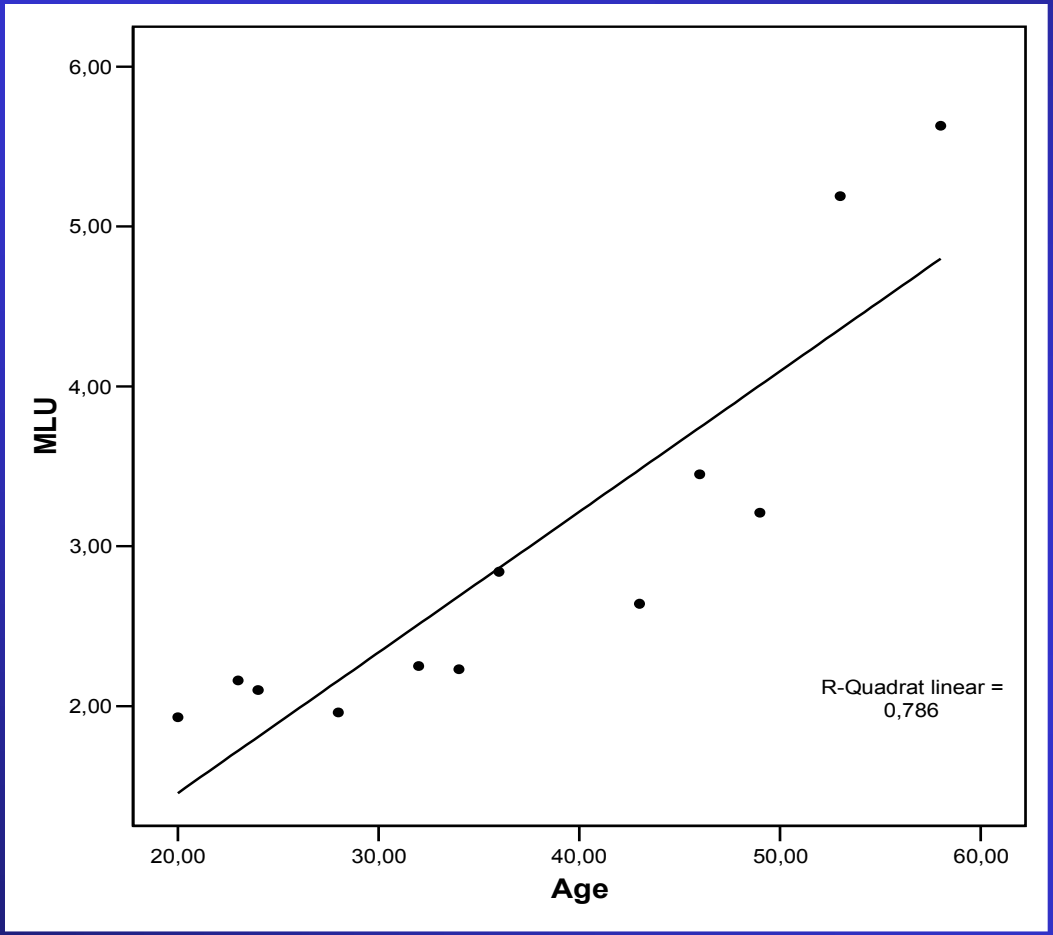
- Conceptual order, i.e. the iconicity of sequence
- The precise meaning of the adverbial clause
- The length of the adverbial clause
- The complexity of the adverbial clause

Multifactorial analysis

Length: Heavy constituents tend to occur at the end of a sentence (in right-branching languages). Two possible reasons:

- Information structure
- Parsing (i.e. Minimize Domains)

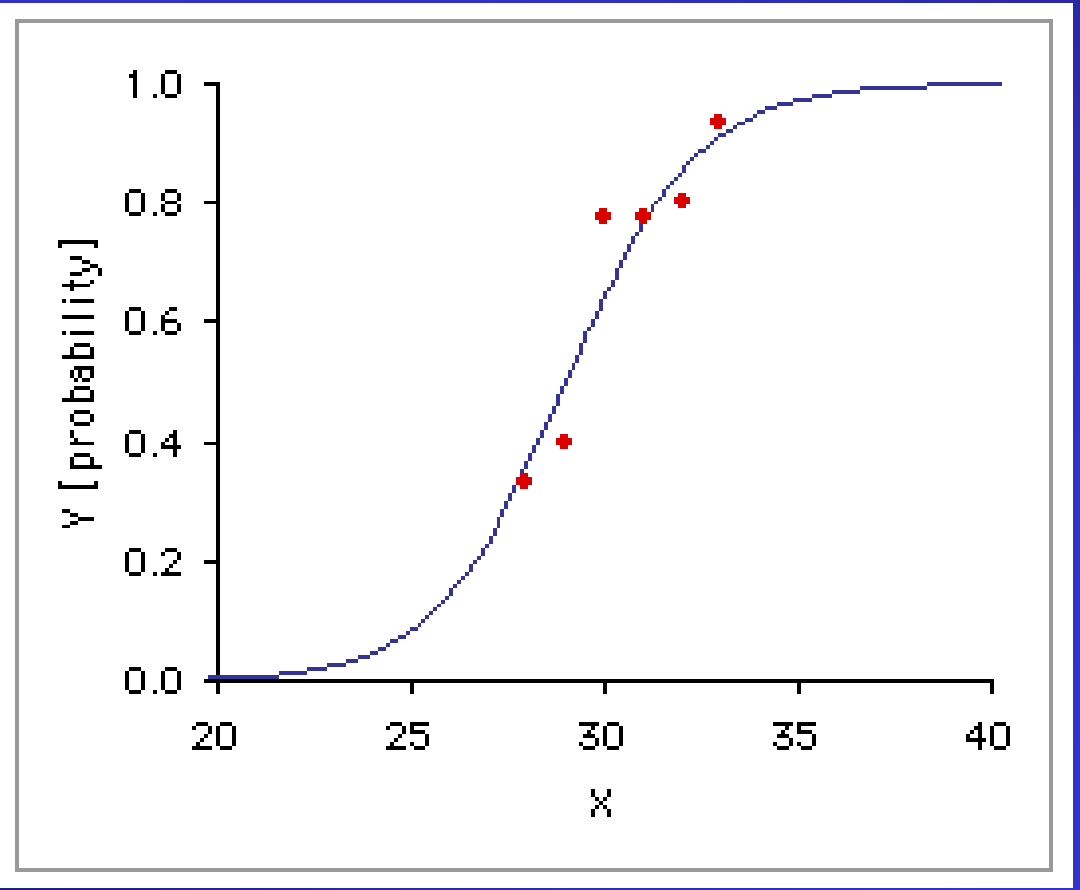
Regression analysis



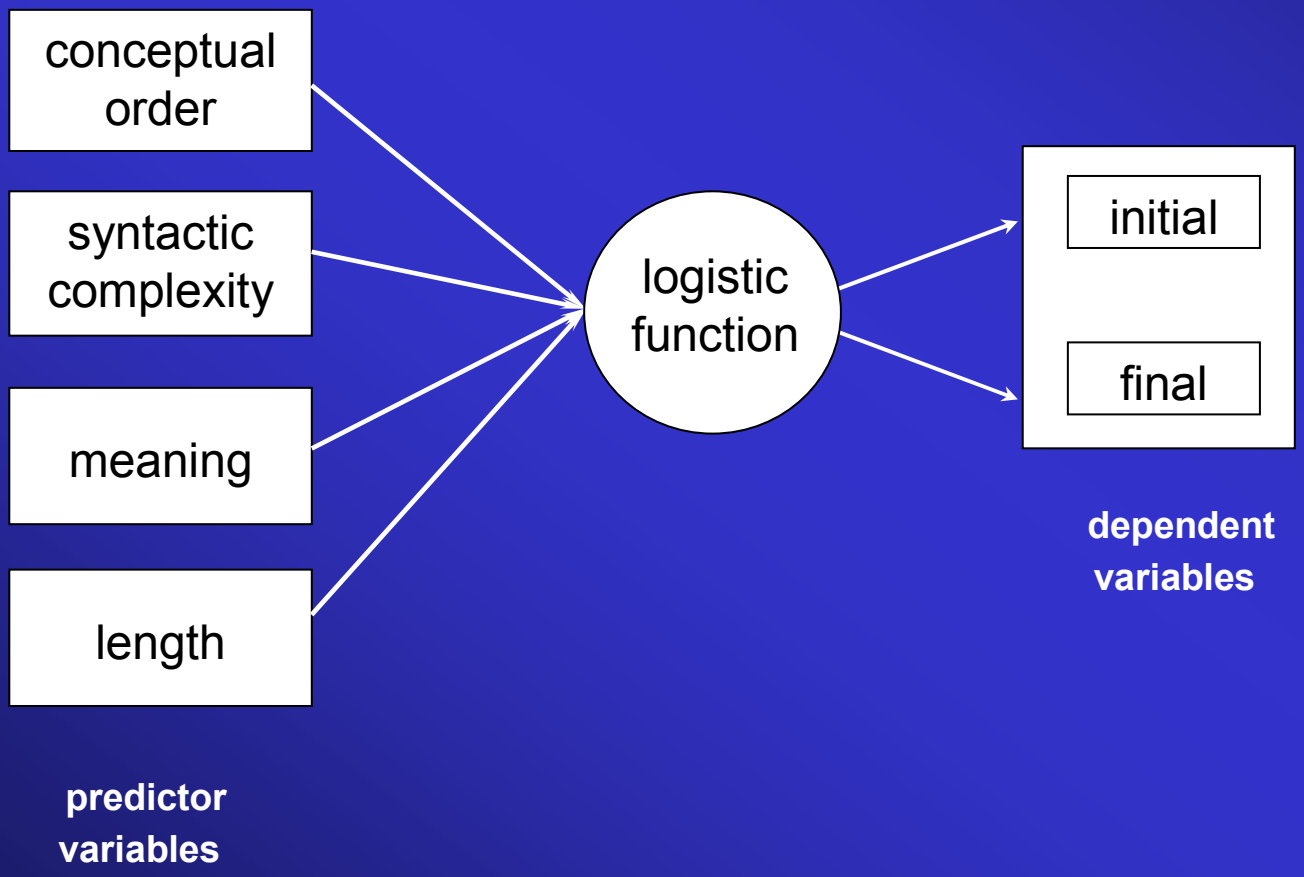
Logistic regression

Logistic regression predict the value of a categorical dependent variable from one or more predictor variables that can be continuous, categorical, or a mix of them.

Logistic regression



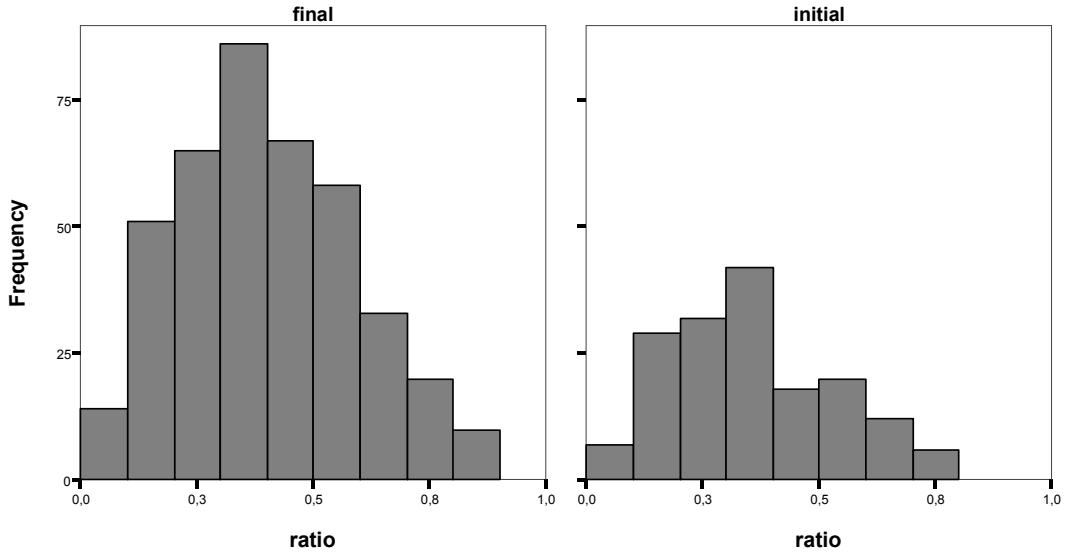
Research design



Variables

VARIABLE	LEVEL	INITIAL	FINAL	TOTAL
Conceptual order	1. posterior/simultaneous 2. prior	47 119	302 102	349 221
Complexity	1. simple 2. complex	138 28	309 95	447 123
Meaning	1. purely temporal 2. conditional 3. causal/purposive	89 76 1	299 52 53	388 128 54

Variables



N	404
SD	0.190
M	0.450

N	166
SD	0.176
M	0.405

Results

Three predictor variables are significant: (1) conceptual order, (2) meaning, and (3) length. Since complexity was not significant, it was excluded from the model.

The resulting minimally adequate model fit the data significantly better than the null model ($\chi 2 = 174.69$, $df = 4$, $p < 0.001$), i.e. the model without predictor variables, and achieved a prediction accuracy of 80.0 percent; i.e. for 80 percent of the temporal clauses the model was able to correctly predict the position from the four independent variables.

Results

Factor	reg. coef. B	Wald χ^2	df	p	odds ratio	lower IC	upper IC
Conceptual order	1.902	73.69	1	0.001	6.70	4.34	10.35
Meaning	-2.775	41.07	2	0.001	0.06	0.01	0.469
(a)causal/purpose	1.364	7.27	1	0.007	3.91	2.42	6.31
(b)conditional		31.20	1	0.001			
Length	-1.343	7.39	1	0.001	0.19	0.06	0.63

Discussion

In the literature, iconicity of sequence is commonly described as a semantic principle, but in the usage-based approach it can be seen as a processing principle that contributes to the overall processing load of a complex sentence because a non-iconic clause order is difficult to plan and to interpret.

Assuming that non-iconic orders are difficult to plan and to interpret, it makes sense that complex sentences tend to be iconic because speakers prefer linguistic structures that are easy to process.

Like iconicity, meaning and length can be interpreted as processing factors.

Discussion

Incremental model of sentence processing (MacDonald et al. 1994)

Sentence processing is influenced by syntactic, semantic, and other processing constraints that together determine the overall processing load of a construction.

Speakers tend to avoid structures whose processing load exceeds a certain level.

Discussion

There is a particularly strong motivation to use iconic clause order in sentences with initial adverbial clauses because the combined effect of the initial position (which is difficult to parse) and the occurrence of a non-iconic clause order (which is difficult to conceptualize) can raise the overall processing load to a very high level.

If the adverbial clause follows the main clause there is less processing pressure to use an iconic clause order because complex sentences with final adverbial clauses are easier to parse; i.e. there is more tolerance in complex sentences with final adverbial clauses for the increased processing load that arises from the violation of the iconicity principle.

References

Holger Diessel and Katja Hetterle. to appear. Cross-linguistic aspects of causal clauses.

Holger Diessel. to appear.
Iconicity of sequence. A corpus-based analysis of the positioning of temporal adverbial clauses in English. *Cognitive Linguistics*.

Holger Diessel. 2005.
Competing motivations for the ordering of main and adverbial clauses. *Linguistics* 43: 449-470.

Holger Diessel. 2001.
The ordering distribution of main and adverbial clauses a typological study. *Language* 77: 345-365.