Testing the Continuity Hypothesis: A decompositional approach

Debopam Das Åbo Akademi University Tehtaankatu 2 20500 Turku, Finland debopam.das@abo.fi

Abstract

The Continuity Hypothesis (CH) predicts that discontinuous discourse relations are harder to process and therefore more marked than continuous ones. To investigate this hypothesis, we annotated a corpus of discourse relations for Givón's (1993) seven continuity dimensions and also for discourse signalling, widening the perspective to discourse signals in general. Our results show that discourse relations often are simultaneously continuous and discontinuous on different continuity dimensions, and that continuity dimensions behave very differently with respect to discourse marking: Only the temporal dimension (partially) confirms the CH while the perspective dimension provides counter-evidence to the CH. Also, contrary to Givón's expectation, local discontinuity introduces more marking than global discontinuity.

1 Introduction

The signalling of discourse relations varies in kind and degree (Das, 2014; Crible, 2020). Different relation types employ different kinds of signalling; e.g., in English, CONDITION relations are mostly signalled by subordinating conjunctions like *if* or *when*, while PURPOSE relations are predominantly marked by the syntactic signal *infinitival clause*. Also, some relations are more marked than others; e.g., CONCESSION relations in comparison to HYPOTHETICAL relations.

The variation in relation signalling is often explained in terms of the Continuity Hypothesis (CH). (Murray, 1997). The CH presumes that discourse comprehension is greatly shaped by expectation, i.e., language users, while processing a text, have default assumptions about the upcoming discourse segment¹. In particular, readers have a preference

Markus Egg Humboldt-Universität zu Berlin Unter den Linden 6 10099 Berlin, Germany markus.egg@hu-berlin.de

for interpreting sequences of sentences in a continuous manner. Continuity ensues when the sentences maintain deictic dimensions such as time, reference, or perspective. Discontinuity, in contrast, arises when inter-sentential transitions are marked by deictic shifts along these dimensions. The CH predicts that discontinuous transitions between sentences are harder to process that continuous ones, and such transitions are therefore explicated more often in terms of suitable markers than continuous ones; e.g., the CONCESSION relations in (1) and (2) both convey discontinuity, but (1) is easier to understand than (2) due to the connective *even though* (examples from Zufferey and Gygax 2016, p. 533).

- (1) Peter married Jane even though he didn't love her.
- (2) Peter married Jane. He didn't love her.

Evidence for the CH mainly comes from psycholinguistic studies. Segal et al. (1991) observe that readers, when given a task to identify the relation types between successive sentences, most often chose causal or additive relations instead of contrastive relations. Murray (1997) shows that signals of discontinuity (i.e., adversative connectives like *but*) have a greater impact on on-line processing than signals of continuity. Further support for the CH comes from corpus data: Asr and Demberg (2012) observe that discontinuous relations display more explicitness than continuous ones.

In this paper, we argue that discourse relations can be simultaneously continuous and discontinuous on different continuity dimensions (*time*, *reference*, or *perspective*). We accordingly examine the CH directly on those dimensions, rather than on relation types as being categorically continuous or discontinuous. Also, unlike previous studies, we focus not only on discourse connectives (DCs), but also on non-DC signals such as lexical relations (e.g., antonymy) and syntactic structures (e.g., parallel syntactic constructions). We examine a corpus

¹Comprehension based on the notion of expectedness is also accounted for by the 'causality-by-default' hypothesis (Sanders, 2005) and the Uniform Information Density (UID) hypothesis (Frank and Jaeger, 2008). For an overview of these hypotheses, see Asr and Demberg (2012).

of about 1,000 relations from five major relation types (CAUSAL, CONDITIONAL, CONTRASTIVE, ELABORATION, and TEMPORAL) that we first annotate with respect to Givón's (1993) seven continuity dimensions (*time*, space, reference, action, perspective, modality, and speech act). We then test the CH, examining the signalling of those relations for individual continuity dimensions.

This paper is structured as follows: Section 2 outlines previous work on continuity (dimensions) in discourse relations. In Section 3, we describe the methodology adapted for the CH analysis. Section 4 presents the results and discussion. We conclude the paper with an outlook on the future work.

2 Background

2.1 Continuity and discourse relations

Previous studies on the CH generally consider continuity as a binary feature, classifying discourse relations categorically as either continuous or discontinuous. For instance, Murray (1997) considers CAUSAL relations continuous, and Zufferey and Gygax (2016) regard CONTRASTIVE relations as discontinuous. Asr and Demberg (2012) group the PDTB relations (Prasad et al., 2008) like RESULT, INSTANTIATION, and LIST as continuous and relations like PRAGMATIC CONTRAST, CONTRA-EXPECTATION, or TEMPORAL relations as discontinuous, whereas they leave CONDITIONAL relations underspecified with respect to continuity.

However, corpus evidence shows that discourse relations can be continuous on some continuity dimensions but at the same time discontinuous on other dimensions. For instance, CAUSAL relations, generally deemed continuous, can simultaneously exhibit continuity for the temporal dimension, but discontinuity for the reference dimension, as in (3).

(3) [As some securities mature and the proceeds are reinvested,] [the problems ought to ease.]

Similarly, CONTRAST relations, usually regarded as discontinuous, can show the same configuration (continuity for time, discontinuity for reference):

(4) [The gasoline picture may improve this quarter,] [but chemicals are likely to remain weak.]

Having noted these incongruities, we first set out to re-examine the relationship between continuity and discourse relations. To do so, we adopted a finegrained approach, decomposing continuity into different continuity dimensions, following Givón's framework (1993), as outlined below.

2.2 Givón's continuity dimensions

Givón defines continuity in terms of thematic coherence, which distinguishes seven continuity dimensions or 'coherence strands'. Maintaining or shifting deictic centres on these dimensions between discourse segments determines the extent of thematic coherence (continuity) or disruption (discontinuity). The seven dimensions are *time*, *space*, *reference*, *action*, *perspective*, *modality*, and *speech act*. The first four are more concrete and local, the others, more abstract and global:

	time
local	space
	reference
	action
	perspective
global	modality
	speech act

Table 1: Givón's coherence strands

The grouping of dimensions is based on effect; consider (5)-(6) from Givón (1993, p. 319, 321). In (5), a change in the temporal continuity across the two clauses causes a local break, but does not necessarily terminate a larger coherent sequence of clauses in the text. In contrast, a change in one of the global dimensions amounts to a stronger break, which can terminate such a sequence of clauses. There is such a break in (6), because it exhibits discontinuity in perspective between the two sentences (viewpoint of the author vs. the one of the protagonist).

- (5) She flew in at midnight and left the next day.
- (6) She came in and sat on the bed. She was tired, she thought.

2.3 Operationalisation of dimensions

We operationalised Givón's seven continuity dimensions in terms of distinctive features. As an example, consider the operationalisation of the *perspective* dimension². We distinguish three types of perspective (Pander Maat, 1998): *objective, author* (in the form of comments), and *other* (quotations). We consider a discourse relation continuous on the perspective dimension if its segments share the same perspective, as in (7), otherwise, as discontinuous, as in (8) (both are CONTRAST relations):

²The operationalisation of the seven dimensions is documented in detail in our previous work (Das and Egg, 2023).

- (7) ["Climate varies drastically due to natural causes," said Mr. Thompson.] [But he said ice samples from Peru, Greenland and Antarctica all show substantial signs of warming.]
- (8) ["The earnings were fine and above expectations," said Michael W. Blumstein, an analyst at First Boston Corp.] [Nevertheless, Salomon's stock fell \$1.125 yesterday to close at \$23.25 a share in New York Stock Exchange composite trading.]

2.4 Continuity annotation on relations

In order to investigate how continuity interacts with discourse relations, we annotated over 1,000 tokens of discourse relations with respect to all seven continuity dimensions. The relations constitute a subset of the RST Discourse Treebank (Carlson et al., 2002), representing five major relation types: CAUSAL, CONTRASTIVE, CONDITIONAL, ELABO-RATION, and TEMPORAL. This selection is motivated by previous classifications, which categorise, e.g., CAUSAL and ELABORATION relations as continuous (Murray, 1997), CONTRASTIVE relations as discontinuous (Zufferey and Gygax, 2016), TEM-PORAL relations as one or the other (Hopper, 1979), and CONDITIONAL relations as underspecified with respect to continuity (Asr and Demberg, 2012).

relation type	predicted continuity
CAUSAL	continuous
CONTRASTIVE	discontinuous
CONDITIONAL	not specified
ELABORATION	continuous
TEMPORAL	(dis)continuous

Table 2: Relation types and their features

We examined 1,009 relations with 276 CAUSAL, 156 CONTRASTIVE, 172 CONDITIONAL, 179 ELABORATION, and 226 TEMPORAL relations. Each relation was independently annotated by two annotators (the authors) for the seven continuity dimensions. We tested the inter-annotator agreement on 240 additional relations. Agreement was substantial according to Cohen's kappa (Landis and Koch, 1977) for the four dimensions *time*, *reference*, *perspective*, and *modality*, as shown in Table 3. For the remaining dimensions, we also agreed, rather overwhelmingly, and no meaningful κ -values could be computed due to prevalence³.

time	reference	perspective	modality
0.72	0.69	0.70	0.76

Table 3: Inter-annotator agreement on four dimensions

3 Testing CH on continuity dimensions

3.1 Results on continuity and relations

The results from our corpus analysis, as summarised in Table 4^4 , show that continuity dimensions interact with discourse relations in varying ways. In particular, some continuity dimensions show uniformity across relation types. All relation types are found to be overwhelmingly continuous (>98%) for the dimensions *space* and *speech act*, and almost never continuous (< 2%) for action. In contrast, the dimensions time, reference, perspective, and modality yield considerable differences amongst the relation types. For these dimensions, the types are not homogeneously continuous or discontinuous, but they can be simultaneously more continuous for some dimensions but less continuous or even predominantly discontinuous for other dimensions. For example, CONTRASTIVE relations are the least continuous for reference and perspective, but highly continuous for time. Furthermore, continuity is not found to be uniform even for a single dimension of one of these relations; e.g., only 82.61% (and not 100%) of the CAUSAL relations are continuous for time.

We measured the significance of the results statistically with a chi-square test, for interdependence between relation types and continuity along a specific dimension. We found that continuity correlates with relation types very significantly for *time*, *perspective*, and *modality* (p < 0.00001). The correlation is significant for *reference* (p < 0.05) and action (p < 0.001), too; but for action, low counts (< 5) reduce the validity of the test. No significant correlation was found between relation types and space or speech act. These findings imply that continuity and discontinuity systematically coexist in relations on the time, reference, perspective, and modality dimensions; consequently, relations are not fully continuous or discontinuous, neither on the level of the entire relation nor for any of these particular dimensions.

Since every relation type exhibits continuity and discontinuity in different continuity dimensions simultaneously, it seems incongruous to test the CH on the level of relation types. Therefore, we test

³The agreement scores for these dimensions were 97.07% for *space*, 95.82% for *action*, and 98.74% for *speech act*.

⁴The highest/lowest scores for a dimension are in bold font.

relation type	time	reference	perspective	modality	space	action	speech act
CAUSAL	82.61%	30.79%	85.87%	80.79%	97.46%	2.54%	99.64%
CONDITIONAL	81.98%	35.47%	93.61%	61.63%	98.84%	5.81%	98.26%
CONTRASTIVE	91.67%	23.72%	67.31%	77.56%	98.08%	0.00%	100%
ELABORATION	93.85%	34.64%	78.21%	85.47%	100%	0.56%	99.44%
TEMPORAL	74.34%	38.50%	90.27%	92.92%	97.35%	0.88%	98.67%
mean	84.04%	32.90%	83.94%	80.57%	98.23%	1.98%	99.21%

 Table 4: Continuity scores across relation types

the validity of the CH on the level of individual continuity dimensions, that is, we examine the signalling of a relation type when it is continuous for a particular dimension as opposed to when it is discontinuous for that dimension. In our analysis, we focus only on the four dimensions, *time*, *reference*, *perspective*, and *modality*, which were distinctive for continuity and discontinuity on relations.⁵

We use the RST Signalling Corpus (RST-SC, Das et al., 2015) to examine the signals of the relations chosen for our continuity analysis. The RST-SC provides the signalling information for the discourse relations in the RST Discourse Treebank (Carlson et al., 2002), where our 1,009 relations come from. The relational signals in the RST-SC include different textual devices such as reference, lexical, syntactic, semantic, and graphical features, in addition to discourse connectives (DCs). Example (9) illustrates an RST-SC signalling annotation:

(9) [Since Mexican President Carlos Salinas de Gortari took office last December,] [special agents have arrested more than 6,000 federal employees on charges ranging from extortion to tax evasion.]

The CIRCUMSTANCE relation is marked by the connective *since* as well as by the change of tense between two clauses (from simple past to present perfect), and also by the indicative phrase *last December*. We examine both DCs and all other signals in our examination of the CH.

4 Results and discussion

We gauge the impact of the four distinctive continuity dimensions (*time*, *reference*, *perspective*, and *modality*) on signalling in three ways. First, we compare the signalling of continuous and discontinuous tokens for each relation type for every continuity dimension. I.e., we examine how frequently a relation type is signalled (by a DC or/and by a non-DC signal) when it is continuous and when it is discontinuous for a particular continuity dimension. The results are summarised in Table 5.

The data show that, along the *time* dimension, relation types on average and a majority of the individual subtypes (except CONTRASTIVE and ELAB-ORATION) are marked more frequently in the absence of temporal continuity than in its presence. For *reference*, the average signalling scores do not vary much between the continuous and discontinuous relations (89.76% vs. 90.39%); still, marking in the absence of referential continuity is higher for CAUSAL and ELABORATION relations but lower for CONTRASTIVE relations. These results are not statistically significant, however.⁶

A different picture emerges for *perspective* and *modality*: Relations, when discontinuous on these dimensions, are less marked on average than the continuous ones (92.09% vs. 80.25% and 90.99% vs. 86.87%), and so are most individual relation subtypes (except CONTRASTIVE for perspective and TEMPORAL for modality continuity). In particular, the results for perspective continuity (except for CONTRASTIVE relations) provide counterevidence against the CH. The numbers are significant here for the average (p < .0001) as well as for CAUSAL and CONDITIONAL relations (p < .01 and p < .0001, respectively).

We also conducted a similar analysis for DCs only, following the spirit of previous work on the CH. The results (in Table 6) for the overall distribution of the DC-only signalling were in line with the previous analysis on general signalling (in Table 5): Again, discontinuous relations tend to be more marked for *time*, but this time the positive evidence of the temporal dimension for the CH was more pronounced (significant for the average at p < .0001 and for CAUSAL and CONDITIONAL relations at p < .05 and p < .0001). The *reference* dimension once again does not offer evidence

⁵For *space*, *action*, and *speech act*, relation types are found to be either almost continuous or discontinuous as a whole.

⁶Lack of significance in Table 5 sometimes results from data sparsity (e.g., there is only one referentially continuous unsignalled CONTRASTIVE relation or only two relations for CONDITIONAL and ELABORATION that are temporally discontinuous and unsignalled).

relation type		time			reference		perspective			modality		
relation type		cont	discont	1	cont	discont	1	cont	discont	1	cont	discont
CAUSAL		89.04%	89.58%		85.88%	90.58%		91.56%	74.36%		89.19%	88.89%
CONDITIONAL		85.11%	93.55%	1	87.30%	86.24%	1	91.30%	18.18%		89.62%	81.82%
CONTRASTIVE		90.14%	85.71%	1	97.37%	87.29%	1	89.52%	90.19%	1	90.91%	85.71%
ELABORATION		95.21%	83.33%	1	91.53%	95.83%	1	96.43%	87.18%	1	95.39%	88.89%
TEMPORAL		88.69%	98.28%	1	90.80%	91.37%	1	91.67%	86.36%	1	90.48%	100%
mean		89.71%	92.64%	1	89.76%	90.39%	1	92.09%	80.25%	1	90.99%	86.87%

Table 5: Distribution of marked relations for continuity dimensions

relation type	time				perspective			modality		
relation type	cont	discont		cont	discont	cont	discont	1	cont	discont
CAUSAL	49.56%	68.75%		57.65%	50.78%	55.69%	35.89%		50.90%	61.11%
CONDITIONAL	78.01%	87.50%		80.95%	79.82%	84.47%	18.18%	1	83.02%	75.76%
CONTRASTIVE	80.28%	78.57%	1	89.47%	77.12%	81.90%	76.47%	1	81.82%	74.29%
ELABORATION	8.38%	8.33%	1	5.08%	10.00%	10.00%	2.56%	1	7.89%	11.11%
TEMPORAL	70.24%	77.59%		70.11%	73.38%	74.02%	54.55%	1	71.43%	81.25%
mean	55.44%	72.39%		59.64%	57.46%	61.28%	41.98%		56.97%	63.13%

Table 6: Distribution of relations with DCs for continuity dimensions

relation type	discontinuous for	
relation type	local global	
CAUSAL	94.44% 83.33%	
CONDITIONAL	90.48% 22.22%	
CONTRASTIVE	80.00% 87.50%	
ELABORATION	90.00% 81.81%	
TEMPORAL	97.62% 100%	
mean	93.28% 75.00%	

Table 7: Signalling for local and global discontinuity

for or against the CH, and the *perspective* dimension clearly goes against the predictions of the CH (significant for the average at p < .0001 and for CAUSAL and CONDITIONAL relations at p < .05 and p < .0001). For *modality*, unlike what we found for general signalling (Table 5), discontinuous relations are marked more frequently by DCs than continuous relations.

Next, we compared relations that are discontinuous on the local dimensions (*time* and *reference*) to those discontinuous on the global dimensions (*perspective* and *modality*). The results (in Table 7) indicate that the first group on average shows more marking than the second one. As a break in global coherence has more impact in Givón's theory, one would have expected a higher need for signalling for the second group, i.e., the reverse result.

As a third measure for the impact of continuity on marking, we attempted to gauge the effect of continuity in general (i.e., irrespective of a particular dimension) on marking. To this end, we examined the distributions of signalled and unsignalled relations for relations that are continuous on 0-4 of the four relevant dimensions. The results (in Table 8) show that, contrary to what one would expect in the light of the CH, more continuous dimensions actually lead to an increase in marking.

We then compared the distributions of marked and unmarked signals across the five groups in terms of relative entropy S(q, p) (also known as Kullback-Leibler divergence), where both p and q are distributions over signalled relations which differ in the number of continuity dimensions. In our case, S(q, p) measures the influence of an additional continuous dimension on the distribution of signalled signals.

dimensions	0 vs. 1	1 vs. 2	2 vs. 3	3 vs. 4
entropy	.01285	.00210	.00001	.00005

Table 9: Relative entropy and continuous dimensions

As shown in Table 9, the impact of additional continuous dimension tends to be greater for smaller numbers of dimensions. This result once again suggests that the degree of continuity for a relation is correlated positively with discourse marking, because it can be interpreted in terms of diminishing marginal utility, e.g., the difference in marking between relations with three and four continuous dimensions is smaller than the one between relations with one and two.

5 Conclusions and outlook

We have argued that continuity functions as a multidimensional phenomenon in discourse relations. We have supported the claim by validating a decompositional approach of annotating relations with respect to different continuity dimensions. We have applied this decompositional approach for testing

23
20
XC

relation type	zero dim.	one dim.	two dim.	three dim.	four dim.
CAUSAL	0%	82.35%	88.10%	90.24%	90.38%
CONDITIONAL	100%	78.57%	83.02%	90.14%	90.91%
CONTRASTIVE	66.67%	84.21%	88.89%	88.41%	100%
ELABORATION	100%	80.00%	94.59%	95.92%	94.44%
TEMPORAL	66.67%	90.91%	98.31%	87.91 %	89.06%
mean	66.67%	83.33%	90.28%	90.71%	91.79%

Table 8: Scores for marked relations for different numbers of continuous dimensions

the Continuity Hypothesis for all relational signals including discourse connectives.

The results from our corpus provided no conclusive evidence for or against the CH on the level of individual continuity dimensions: Temporal continuity is found to (partially) corroborate the CH while continuity along perspective contradicts it. Furthermore, contrary to Givón's line of reasoning, global discontinuity is found to decrease the amount of discourse marking. Finally, continuity, when the specificity of its dimensions is not taken into account, correlates with discourse signalling positively, hence going counter to the CH.

We would, however, like to point out that our results on continuity and the CH are based on the newspaper genre of the corpus (RST-DT). Continuity might function differently in other genres, e.g., fiction (as in Givón's framework), and also across languages, as shown by Mendes et al. (2023).

In future work, we will incorporate more data (in terms of additional relation types and also corpus size) in the evaluation of the CH. We will also investigate whether relation types and their marking are differently susceptible to the impact of continuity. Furthermore, our results motivate searching for other potential factors for the data to explain why they do not fit in with the predictions of the CH.

References

- Fatemeh Torabi Asr and Vera Demberg. 2012. Implicitness of Discourse Relations. In *Proceedings of COLING 2012*, page 2669 –2684.
- Lynn Carlson, Daniel Marcu, and Mary Ellen Okurowski. 2002. RST Discourse Treebank, LDC2002T07.
- Ludivine Crible. 2020. Weak and strong discourse markers in speech, chat and writing: Do signals compensate for ambiguity in explicit relations? *Discourse Processes*, 57:793–807.
- Debopam Das. 2014. *Signalling of Coherence Relations in Discourse*. PhD dissertation, Simon Fraser University, Canada.

- Debopam Das and Markus Egg. 2023. Continuity in discourse relations. *Functions of Language*, 30:41– 66.
- Debopam Das, Maite Taboada, and Paul McFetridge. 2015. RST Signalling Corpus, LDC2015T10.
- Austin Frank and Florian Jaeger. 2008. Speaking rationally: Uniform information density as an optimal strategy for language production. In *Proceedings of the 30th Annual Meeting of the Cognitive Science Society*, page 933–938.
- Talmy Givón. 1993. English Grammar: A functionbased introduction, volume 2. John Benjamins.
- Paul Hopper. 1979. Aspect and foregrounding in discourse. In Talmy Givón, editor, Syntax and semantics, Vol. 12. Discourse and syntax, page 213–241. Academic Press, New York.
- Richard Landis and Gary Koch. 1977. The measurement of observer agreement for categorical data. *Biometrics*, 33:159–174.
- Amália Mendes, Deniz Zeyrek, and Giedre Oleškevicčiene. 2023. Explicitness and implicitness of discourse relations in a multilingual discourse bank. *Functions of Language*, 30(1):67–91.
- John Murray. 1997. Connectives and narrative text: The role of continuity. *Memory and Cognition*, 25:227–236.
- Henk Pander Maat. 1998. Classifying negative coherence relations on the basis of linguistic evidence. *Journal of Pragmatics*, 30:177–204.
- Rashmi Prasad, Nikhil Dinesh, Alan Lee, Eleni Miltsakaki, Livio Robaldo, Aravind Joshi, and Bonnie Webber. 2008. The Penn Discourse Treebank 2.0. In *Proceedings of LREC 2008*, pages 2961–2968.
- Ted Sanders. 2005. Coherence, causality and cognitive complexity in discourse. In *Proceedings/Actes SEM-05, First International Symposium on the Exploration and Modelling of Meaning*, page 105–114.
- Erwin Segal, Judith Duchan, and Paula Scott. 1991. The role of interclausal connectives in narrative structuring. *Discourse Processes*, 14:27–54.
- Sandrine Zufferey and Pascal Gygax. 2016. The role of perspective shifts for processing and translating discourse relations. *Discourse Processes*, 53:532–555.