THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

When do we leave discourse relations underspecified?

OPEN ACCESS

When Do We Leave Discourse Relations Underspecified? The Effect of Formality and Relation Type

Ludivine Crible and Vera Demberg

OpenEdition
Journals

When Do We Leave Discourse Relations Underspecified? The Effect of Formality and Relation Type

Ludivine Crible

Université catholique de Louvain / University of Edinburgh

Vera Demberg

Universität des Saarlandes

When Do We Leave Discourse Relations Underspecified?
The Effect of Formality and Relation Type

Ludivine Crible

Université catholique de Louvain / University of Edinburgh

Vera Demberg

Université des Saarlandes

Speakers have several options when they express a discourse relation: they can leave it implicit, or make it explicit, usually through a *connective*. Although not all connectives can go with every relation, there is one that is particularly frequent and compatible with very many discourse relations, namely *and*. In this paper, we investigate the effect of discourse relation type and text genre on the production and perception of underspecified relations of contrast and consequence signalled by *and*. We combine a corpus study of spoken English, a production experiment and a perception experiment in order to test two hypotheses: (1) *and* is more compatible with relations of consequence than of contrast, due to factors of cognitive complexity and conceptual differences; (2) *and* is more compatible with informal than formal genres, because of requirements of recipient design. The three studies partially converge in identifying a stable effect of relation type and genre on the production and perception of underspecified relations of consequence and contrast marked by *and*.

Keywords: discourse relations, underspecification, production, perception, experimental

Pour exprimer une relation discursive, les locuteurs disposent de plusieurs options : ils peuvent la laisser implicite, ou la rendre explicite, généralement par le biais d'un connecteur. Bien que chaque connecteur soit spécialisé dans l'expression d'une ou de plusieurs relations, il y en a un qui est particulièrement fréquent et compatible avec de nombreuses relations discursives : and *[et]. Dans cet article, nous étudions l'effet du type de relation et du genre de texte sur la production et la perception de relations de contraste et de conséquence sous-spécifiées exprimées par* and*. Nous combinons une étude de corpus sur l'anglais oral, une expérience de production et une expérience de perception afin de tester deux hypothèses : (1)* and *est plus compatible avec la relation de conséquence que de contraste, en raison de facteurs liés à la complexité cognitive et aux différences conceptuelles ; (2)* and *est plus compatible avec les gens informels que formels, à cause des exigences de* recipient design*. Les trois études convergent partiellement et identifient un effet stable du type de relation discursive et du genre sur la production et la perception des relations de conséquence et de contraste sous-spécifiées exprimées par* and*.*

Mots clés : relations discursives, sous-spécification, production, perception, expérimental

1.   Introduction

1   It is generally assumed that coherence relies on building discourse relations across utterances. Speakers have several options when they express a discourse relation: they can leave it implicit, or make it explicit (e.g., Taboada, 2009), usually through a "connective". There are many connectives to choose from, and a lot of them

have similar meanings and can be used in different relations. Although not all connectives can go with every relation, there is one that is particularly frequent and compatible with very many discourse relations, namely *and*. Besides its core meaning of addition, *and* has been found to co-occur with relations as varied as temporal sequence, specification, consequence, even concession (Crible, 2017; Gómez Txurruka, 2003; Kitis, 2000; Luscher & Moeschler, 1990; Prasad et al., 2008; Spooren, 1997). Converging evidence suggests that *and* is not very informative in terms of the instruction it provides for discourse interpretation and processing: *and* is often omitted from translations (Crible et al., 2019); it yields longer reading times than more specific connectives (Cain & Nash, 2011); it is produced more frequently by children than by adults (Cain & Nash, 2011; Spooren, 1997). As such, *and* is not an ideal choice in all contexts, from the perspective of optimal communication.

2   The imbalance between the additive semantics of *and* and its enriched (causal, temporal, contrastive, etc.) interpretations in context can be termed "underspecification", following Spooren (1997). Underspecified uses of connectives (*and* in particular) correspond to cases where the meaning of the discourse relation is more specific than the encoded semantics of the connective. For instance, a relation of consequence is more specific than a basic additive relation in that it involves a causal inference: the two arguments are not merely conjoined but causally related. Similarly, a contrastive relation signals that there is some difference between the conjoined arguments. In this paper, we use "polyfunctionality" as a cover term for all types of variation in meaning, including underspecification as well as other phenomena such as ambiguity (e.g., *but* between contrast and concession, Asr & Demberg, 2016). While many connectives are polyfunctional (i.e., they can express multiple relations), the specificity of underspecification is that the different meanings of the connective are not equally entrenched (the underspecified use "has not been incorporated into the lexical meaning", Spooren, 1997: 153) nor equally informative. We here consider *and* to be such an underspecified connective.

3   Relations marked with an underspecified connective thus constitute an intermediate level between implicit and explicitly marked discourse relations: there is an explicit connective, but it only partially matches the target interpretation and is underinformative in this respect. The objective of this paper is to determine the contextual and linguistic conditions that favour the production of underspecified uses of *and*. In particular, we want to test whether underspecification is more frequent and more acceptable in informal, unplanned genres, and whether it equally applies to different discourse relations (consequence vs. contrast), compared to relations marked by stronger connectives such as *so* or *but*.

4   We report on the findings of a corpus study and two offline crowd-sourcing experiments. First, we discuss the distribution and functions of *and* in different genres of spoken English, in order to compare its underspecification across contexts. We then present the results of a production experiment (connective insertion task) and a perception experiment (forced choice task), testing the preference for

and over other options in two genres (informal chat conversation, more formal online comments) and in two relations (consequence and contrast). The three studies converge in identifying a strong effect of relation type (*and* is dispreferred in contrastive relations) and, to a smaller extent, a partial effect of genre (*and* is slightly dispreferred in the formal setting), thus confirming both our hypotheses. Relevant previous studies are reviewed in the next section.

2.   Previous studies

2.1.   Discourse relations from implicit to explicit

5   Discourse or coherence relations are non-compositional meanings that arise from the connection between two propositional segments, typically clauses or larger units. They include addition, contrast, cause, condition, concession, among many others (e.g., Mann & Thompson, 1988; Asher & Lascarides, 2003), and are pervasive in both spoken and written language. Such relations can be expressed more or less explicitly in language use, and this explicit marking is typically achieved through the functional category of connectives (also called "discourse markers"; e.g., Schiffrin, 1987) [1], although syntactic or semantic features also play a part (Morera et al., 2010; Taboada, 2006). Still, the marking of discourse relations is usually considered as either implicit or explicit (as in the Penn Discourse Treebank, Prasad et al., 2008) [2], although recent work is striving to refine this dichotomy (Crible, in press; Das & Taboada, 2018; Hoek et al., 2018).

6   The absence or presence of an explicit connective is not random nor does it apply equally to all discourse relations. Many studies have observed that some relations are more frequently implicit than others, either because they are more expected (Asr & Demberg, 2012), less cognitively complex (Hoek et al., 2017) or because genre conventions allow it (Taboada, 2006). In this perspective, it is useful to adopt Sanders et al.'s (1993) operational classification of discourse relations into four binary features: basic operation (additive or causal), order (basic or non-basic), polarity (positive or negative) and source of coherence (objective or subjective). With this framework, each type of discourse relation can be defined as a combination of these features. We can compare the three relations under scrutiny in these terms: the relation of addition is additive-positive, contrast is additive-negative, consequence is causal-positive. We can thus see on which specific aspect various relations differ.

---

1.   "Connectives" and "discourse markers" are sometimes used as terminological equivalents, although most authors draw a distinction and consider connectives as a subtype of discourse markers that specialise in discourse relations, as opposed to other expressions which are more specific to spoken language (*oh*, *well*, *now*), as described in Schiffrin (1987).

2.   The PDTB also identified a number of relations that were expressed by "alternative lexicalisations" such as "the reason for this is", but these cases are very rare in the corpus.

7   Sanders et al.'s (1993) framework is also beneficial because it accounts for the variation between discourse relations in terms of order of acquisition (Evers-Vermeul & Sanders, 2009), implicitation in translation (Hoek et al., 2017) or processing (Sanders & Noordman, 2000). Converging evidence indeed suggests that negative relations (such as contrast) are more complex than positive ones (e.g., addition; Murray, 1997). No such difference has been strongly established between additive (e.g., addition) and causal relations (e.g., consequence). Sanders (2005) talks of a "paradox of causality", which explains why causal inferences are automatic in human processing and therefore not costly. On this basis, contrastive relations can be expected to be more often explicit than implicit, in order to relieve the interpretation task. Inferring an implicit consequence relation, however, should not cost too much extra effort. We can wonder whether this line of reasoning for implicit relations also applies for underspecification. If that is the case, underspecified consequence relations can be expected to be more frequent and more acceptable than underspecified contrastive uses.

2.2.   Connective strength and informativeness

8   Within explicit relations, connectives vary in their degree of informativeness or "strength" in marking particular relations, depending on the number of relations they can express and their different frequencies. The polyfunctionality of connectives is well-known and concerns many items, such as *while* or *actually*. In spoken language in particular, speakers have to attend to multiple tasks at once, which may result in some connectives performing several functions simultaneously (Petukhova & Bunt, 2009). Asr and Demberg's (2012) probabilistic measure of "cue strength" can be used to quantify how specific a given connective is for a given relation, thus ranking connectives from weaker to stronger signals. By applying this measure to the PDTB 2.0 data (Prasad et al., 2008), Asr and Demberg observe that most connectives are weak (i.e., they express many relations, which are in turn expressed by many connectives).

9   The strength or informativeness of a connective can be connected to pragmatic principles of discourse production and comprehension, starting from Grice's conversational maxims, Horn's (1984) Q and R principles, to more recent accounts (Uniform Information Density hypothesis, Levy & Jaeger, 2007; Rational Speech Act theory, Frank & Goodman, 2012). They all suggest that expected or available information will not be further marked explicitly, while unavailable information requires explicit marking, in order to avoid too much effort from both speaker and listener (see Asr & Demberg, 2015 for an application to implicit discourse relations). By combining considerations of relation complexity and informativeness, it can be hypothesised that contrast will require stronger connectives than addition, given that a weak cue of contrast would result in a peak in information density in the second argument (once the contrast relation is retrieved).

10   Another factor impacting connective strength, besides the complexity of the target relation, may be related to the communicative setting. Corpus studies suggest

that connective types and discourse relations vary tremendously across contexts and modalities. Comparing various academic settings, Biber (2006) concluded that some connectives such as *however* or *therefore* mostly occur in writing, not in speech. Crible (2018: 109) further showed that some relations (e.g., consequence, concession) are highly frequent in "intermediary and formal registers but not in casual conversations". Such findings can be explained by genre conventions, which impose a certain level of formality and recipient design, and by cognitive resources, which are typically limited in spontaneous conversation, because of time and planning pressure. It is therefore interesting to investigate whether such genre variation also applies to *and*, given that it is frequently underspecified. Experimental validation of this genre effect is still lacking and is precisely the subject of the present study.

2.3. *And*: underspecification beyond addition

11   The additive conjunction *and* is a particularly weak cue for relations such as contrast or consequence, which are conceptually more specific (or more complex) than addition and are typically expressed by other connectives. Studies from various frameworks and methods (theoretical, corpus-based, experimental) converge in observing a large number of relations expressed by *and* or its crosslinguistic equivalents, including negative and causal ones (e.g., Gómez Txurruka, 2003; Kitis, 2000; Pander Maat, 1999). Consider examples [1]-[3] from the ICE-GB corpus (Nelson et al., 2002).

[1]   my name is Sam (560 ms) um *and* I run an amphibious tour operation in Plymouth

[2]   God was so horrified at the fact (340 ms) um that he kind of smited the church *and* the spire (900 ms) uh became crooked

[3]   there are the kind of people that are interested in other people (950 ms) *and* there are the kind of people that are uninterested in other people

12   In [1], *and* is used in its basic additive or continuative function in a descriptive passage, as Schiffrin (2006) already described in spoken data. In [2], there is a cause-effect relation between the two clauses connected by *and*, which is reflected in the use of the verb *became*. In [3], the speaker is contrasting two kinds of people and does so with a clear structural repetition and antonyms in the two clauses. Despite the additive semantics of *and*, these examples show that it is compatible with a broad range of discourse relations.

13   This large functional spectrum is the product of the fact that the basic, most frequent meaning of *and* is virtually empty and is therefore compatible with many contextual inferences. In addition, *and* is one of the most frequent connectives in English (Prasad et al., 2008) and thus occurs in different types of contexts. In spoken discourse, Barth-Weingarten (2014) even showed that uses of *and* in utterance-final positions are developing, in which case it has a more discourse-organisational function, rather than its additive connective use. This pervasiveness of *and* calls for a better understanding of when and how it can be used.

14   Yet, *and* has been somewhat overlooked so far, especially in experimental studies which tend to discard it as an empty connective. The underspecified uses of *and* (in its Dutch form *en*) are the focus of Spooren's (1997) production study, in which he found that the temporal and causal (consequence) uses of the conjunction are produced more frequently by children than by adults, and by non-native adults than by proficient speakers. Cain and Nash (2011) included contrastive ("adversative") relations in their study and corroborated these findings; they also reported that *and* yields longer reading times in causal or adversative relations, but that such underspecified uses are less disruptive than "wrong" uses (e.g., the adversative connective *but* in a causal relation).

15   In sum, *and* is highly polyfunctional and can occur in contrastive and consequence relations, among others, although these underspecified uses may vary with the context. We can formulate two hypotheses from the literature. Firstly, the production and perception of *and* vary with the target discourse relation, as an effect of the cognitive complexity of the relation: specifically, *and* should be more frequent and acceptable in consequence relations than in contrast, given that negative relations are more complex and therefore require stronger marking (HP1). Secondly, the production and perception of *and* vary across genres: its underspecified uses should be more frequent and more acceptable in informal genres, where cognitive resources are limited (resulting in speaker economy, i.e. using the same connective for multiple purposes) and where requirements of recipient design are less important (HP2). These hypotheses will be tested in three studies presented below.

3.   Production: corpus study

16   In this corpus study, we investigate how the underspecified uses of *and* vary across genres. Following our second hypothesis, the core meaning of the connective is expected to account for a larger proportion of all its uses in the formal genres, whereas more variation should be observed in informal language.

3.1.   Data and method

17   This study analyses spoken data sampled from the English component of the DisFrEn database where all connectives have previously been manually identified and disambiguated by the first author (Crible, 2017). The sample contains about 60,000 words of spoken English from various genres, which are grouped into three categories of formality: formal (news broadcasts, political speeches), semi-formal (interviews, classroom lessons) and informal settings (face-to-face and phone conversations). The transcripts are sound-aligned and were annotated under the EXMARaLDA software (Schmidt & Wörner, 2009). This corpus was thus selected because of the different genres it included and the rich annotations it provided.

18   A total of 2,908 connectives were manually identified. The selection followed criteria of optionality, clausal-level scope, formal fixedness (high grammaticalisation), procedural meaning, and relational function (see Crible, 2017 for details). The senses of these tokens were then disambiguated by the same annotator following a coding

scheme including discourse relations such as addition, contrast or consequence, as well as non-relational uses such as topic-shift[3]. The coding scheme provides operational definitions, criteria and examples to help in the disambiguation process. For the purpose of the present study, we have grouped some similar senses together, in order to avoid rare values. The list of discourse relations and other pragmatic functions performed by *and* in our sample is as follows: addition (simple conjunction of two contents, no extra inference); adversative (contrastive or concessive relation); consequence (the second segment is the result of the first one); punctuation (the connective is used to stall, as a filler); specification (the segment gives a detail or an example); structuring (the connective performs some topic-structuring or turn-taking function); temporal (the two segments are chronologically ordered) (Crible, 2017).

19   In addition to *and*, we report on the distribution of *but* and *so* as baselines. These two high-frequency connectives typically express relations of contrast and concession (merged into "adversative" here) and of consequence, respectively, and are therefore relevant to the present study. If our hypothesis also applies to *but* and *so*, the effect of genre on the meaning variation of connectives will be further evidenced. The sample contains 966 tokens of *and*, 431 tokens of *but* and 414 tokens of *so*. We present their relative frequency per thousand words in Table 1.

	Informal	Semi-formal	Formal
And	265 (9.73)	618 (17.53)	101 (6.43)
But	214 (7.86)	172 (4.88)	45 (2.87)
So	166 (6.1)	240 (6.81)	8 (0.51)

Table 1 – Frequency of *and*, *but* and *so* in the sample (normalised per thousand words)

3.2.   Results: underspecification across genres

20   As a reminder, we expect the proportion of the core meaning of *and* to be larger in formal genres, whereas informal genres should favour its underspecification and present more types of discourse relations. Figure 1 shows the proportion of the basic meaning of *and* along with *but* and *so* (addition, adversative and consequence, respectively) in the three types of settings represented in the sample. We can see that, as expected, *and* is more often underspecified in informal genres (orange and yellow bars). The slight (and unexpected) difference between informal and semi-formal genres can be explained by the frequent use of *and* in specification relations in the semi-formal genre of interviews, where speakers are often asked to describe a given topic in further detail. Overall, the large gap between informal (e.g., conversations) and formal (e.g., political speech) brings substantial support for our hypothesis.

———

3. Sense disambiguation for connectives is a challenging and partly subjective task. The limits of the validity of this method should be borne in mind in the remainder of this study, and are discussed at length in Crible and Degand (2019).
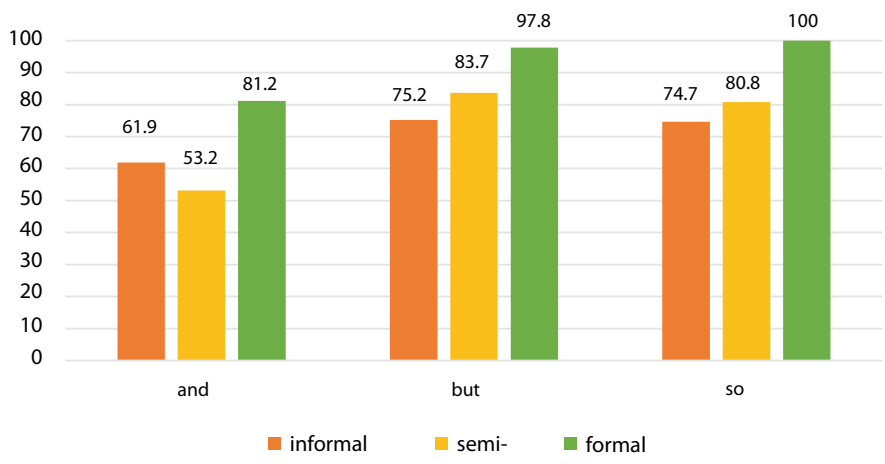
Figure 1 – Proportion of the basic meaning across genres for *and*, *but* and *so*
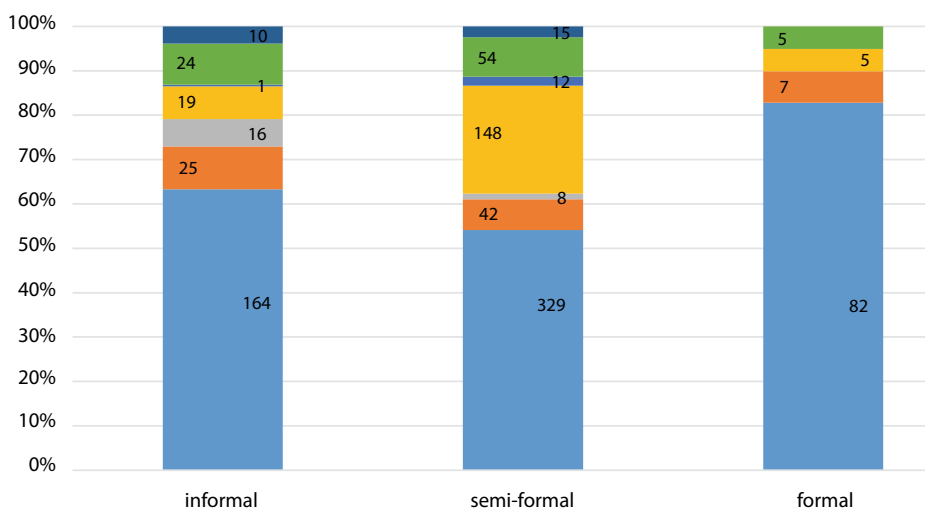
Figure 2 – Sense distribution of *and*

21   The hypothesis is also true for *but* and *so*, which confirms a consistent genre effect on connective use. The basic meaning takes up a much larger proportion in the formal genres, close or up to 100% of the uses for *but* and *so*. Our genre hypothesis therefore seems to be confirmed, as far as this sample is concerned.

22   We can zoom in on *and* to see which particular relations or functions are expressed besides its core additive meaning (Figure 2). Up to six other functions are expressed by *and* in the sample, primarily consequence and specification relations as well as structuring functions. Such underspecified uses are almost non-existent in the formal data, where no adversative (i.e., contrastive) uses can be observed.

However, *and* expresses consequence at a similar rate in all three settings. This could relate to the fact that consequence is not a very complex relation and can easily be inferred from underspecified *and*, therefore such uses are not affected by genre conventions, contrary to the adversative uses.

23   In sum, both our hypotheses are confirmed by this corpus study: underspecification is more frequent in informal genres, and *and* cannot be used in contrastive relations in the formal genre (it is already rare in informal genres). There is thus an apparent relationship between functional spectrum and genre in production. We will now test whether these observations hold in experimental conditions, both from the viewpoint of production and perception (acceptability).

4.   Production: connective insertion

24   In this first experiment, we are measuring participants' preferences in the connectives that they consider the best possible fit between pairs of sentences representing relations of contrast or consequence. These two relations were selected on the basis of their distribution in the corpus study: consequence is an underspecified use of *and* which is much less frequent than the basic additive meaning but can still be found in formal genres; contrastive uses, however, were not found in formal genres. Consequence and contrast thus differ in their formality when they are expressed by *and*. By manipulating the discourse relation and the text genre, we aim to reproduce the production effects found in corpus data. We expect that participants will avoid using *and* as a connective to mark contrastive meanings in the formal genre, but may be more likely to use *and* in a contrastive context in informal genres.

4.1.   Materials

25   Materials for this experiment (and the next) were extracted from the chat component of the Loyola CMC corpus (Goldstein-Stewart et al., 2008), which contains recordings of American university students. This corpus was selected because it includes data from computer-mediated communication (chat conversations), which combines features of writing and features of informal language. Since our experiments use materials in their written format, this affordance of the Loyola CMC corpus was privileged over corpora containing speech exclusively (such as the DisFrEn dataset used in the corpus study above). In the corpus, the sentences were originally related by various connectives, including *and*, *but*, *so*, *however* or *therefore*. We only extracted cases of explicit relations (i.e., with an explicit connective). Some sentences were slightly modified to better match standard English, in spelling, lexicon or syntax.

26   From this original data, we selected relations of contrast and of consequence. To do so, we checked whether all uses of *but* and *however* corresponded to either contrast or concession, and whether *so* and *therefore* always expressed consequence. For

relations originally marked by *and*, because of its tendency towards underspecification, we pre-tested the interpretation of sentences in a crowd-sourced disambiguation task by naïve participants. The materials to be disambiguated consisted in 80 relations originally marked by *and* from the Loyola CMC corpus, divided into four lists of 20. In each list, each relation was pre-annotated by an expert as either additive (10 cases), adversative (5 cases) or consequential (5 cases). The sentences were presented without the original *and*. Participants had to choose between the following list of options: *in addition*, *plus*, *therefore*, *as a result*, *by contrast*, *whereas*, *nevertheless*, *yet*, in addition to manual answers if none of the options was selected. Each list was seen by 20 remunerated participants on Prolific Academic. We then measured the agreement between participants for the inserted connectives, and only selected relations with an agreement score of 80% or higher.

27   For the present study, we ended up with 24 different items, balanced across relation type and original connective as follows: 10 relations of consequence (4 *and*, 4 *so*, 2 *therefore*), 10 relations of contrast (4 *and*, 4 *but*, 2 *however*), 4 distractors (2 temporal and 2 conditional relations). We then manipulated the formality of the sentences by creating a second, more formal version of each pair, compared to the original "chat" version (conversational, spoken-like). This second version was meant to represent comments to online press articles (not conversational, more written-like). The genre of comments was selected because of its assumed familiarity to participants, as opposed to more formal genres such as newspaper articles, which most of our participants will never have produced. However, the specific genre is not as relevant to our study as the linguistic features that are associated with it, and our main aim was to create more formal versions of the materials from the chat corpus. To do so, we made systematic changes to the punctuation (capital letters and punctuation signs added), verb contractions (e.g., "won't" into "will not") and lexical register (e.g., "locked up in jail" into "imprisoned").

28   In addition to these linguistic changes, the visual display of the text on screen was also different for the two genres: the chat versions were presented as lines in a chatbox; the comment versions were presented as a paragraph, with a title containing the pseudonym of the writer, the location and the date. Previous and following sentences were blurred (see Figures 3 and 4).
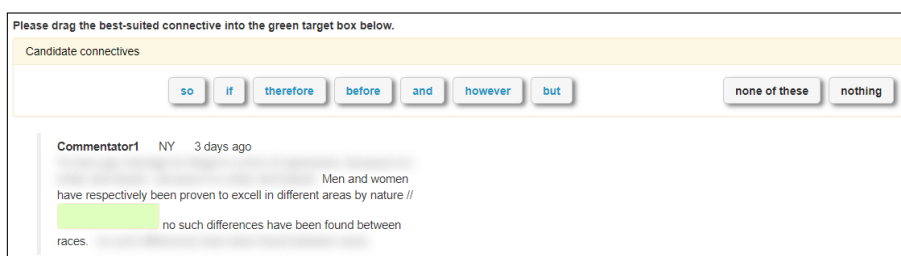

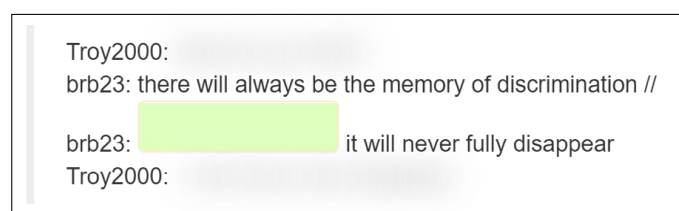Figure 3 – Display for the comments materials

Figure 4 – Display for the chat materials

4.2.   Procedure

29   In the experiment, the original connectives were removed from the sentences, rendering the relations virtually implicit. The participants had to drag and drop a connective from a list of options (cf. Figure 3) including *and*, *but*, *so*, *however*, *therefore*, *if*, *before*, as well as a manual entry option ("none of these") and a null option ("nothing"). The experiment was programmed on LingoTurk (Pusse et al., 2016) and run on Prolific Academic. Two groups of 30 native English speakers were recruited and remunerated (8€ per hour).

30   Answers were first coded as either correct or incorrect depending on whether they matched the target relation. The type of connective inserted was then classified into underspecified (*and*), appropriate (*but*, *so*) or strong (*however*, *therefore*). In the following, we only report the data from correct insertions, excluding distractors, wrong insertions (e.g., *however* in a consequence relation) and null insertions (no connective), which amounts to 2,030 insertions.

4.3.   Results

31   As a reminder, we are testing for an effect of relation type (fewer *and*-insertions in contrastive relations; HP1) and of genre (fewer *and*-insertions in comments; HP2). Table 2 presents the proportions of connective types (underspecified, appropriate or strong) across conditions. The mosaic plots in Figures 5 and 6 show the Pearson residuals for the two independent variables: the colours on the graph show the deviation from the expected frequency (blue for fewer observed than expected, red for more observed than expected), based on a chi-square test. These graphs allow us to identify which specific combinations of variables were significantly more or less frequent than expected. We observe a significant effect of relation type, with much fewer insertions of underspecified *and* in contrast than in consequence, which confirms our first hypothesis. As for the effect of genre, no difference is observed for underspecified *and* between chat and comment versions. However, there are significantly more strong connectives (*however*, *therefore*) in comments than in chat conversations for both relations (38.78% vs. 27.88% for consequence; 42.39% vs. 30.22% for contrast). The second hypothesis is therefore only partially confirmed: strong connectives are more typical of formal genres, but underspecified connectives are equally produced in both settings.

	chat		comment		Total
Connective	conseq.	contr.	conseq.	contr.	
Underspecified	31.86%	11.38%	29.18%	12.50%	20.54%
Appropriate	40.27%	58.40%	32.04%	45.11%	44.38%
Strong	27.88%	30.22%	38.78%	42.39%	35.07%
Total	**452**	**536**	**490**	**552**	**2,030**

Table 2 – Connective types inserted across conditions
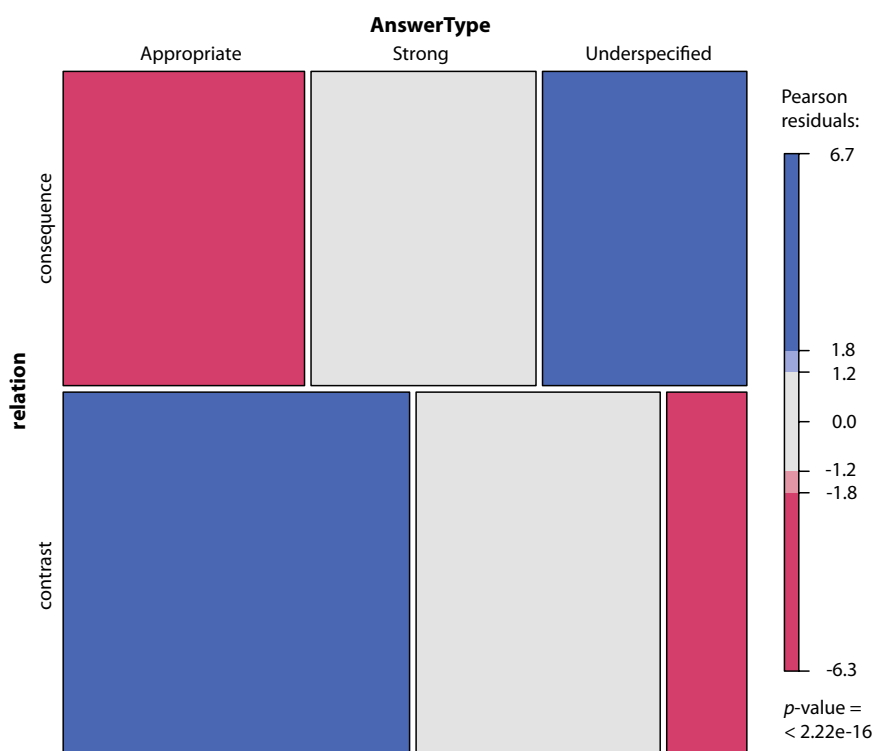
Figure 5 – Distribution of Pearson residuals for connective type across relations

³² If we compare the inserted connectives to the original connective in the corpus materials, we notice that *and* was chosen in 20.54% of all trials, which is a substantial proportion given the conditions of the experiment (*and* facing equally frequent connectives conceptually closer to the target relation). *And*-insertions are particularly frequent when the original connective was either *and* itself (29.47%) or *so* (25.85%), followed by *therefore* (16.45%), which leaves much smaller proportions for the two contrastive connectives *but* (7.16%) and *however* (5.99%). The high proportion of original *so* replaced by *and* is telling of the strong compatibility between *and* and consequence relations, which can be explained by the cognitive simplicity of this relation and the ease with which humans infer causality, even from underspecified contexts.

AnswerType
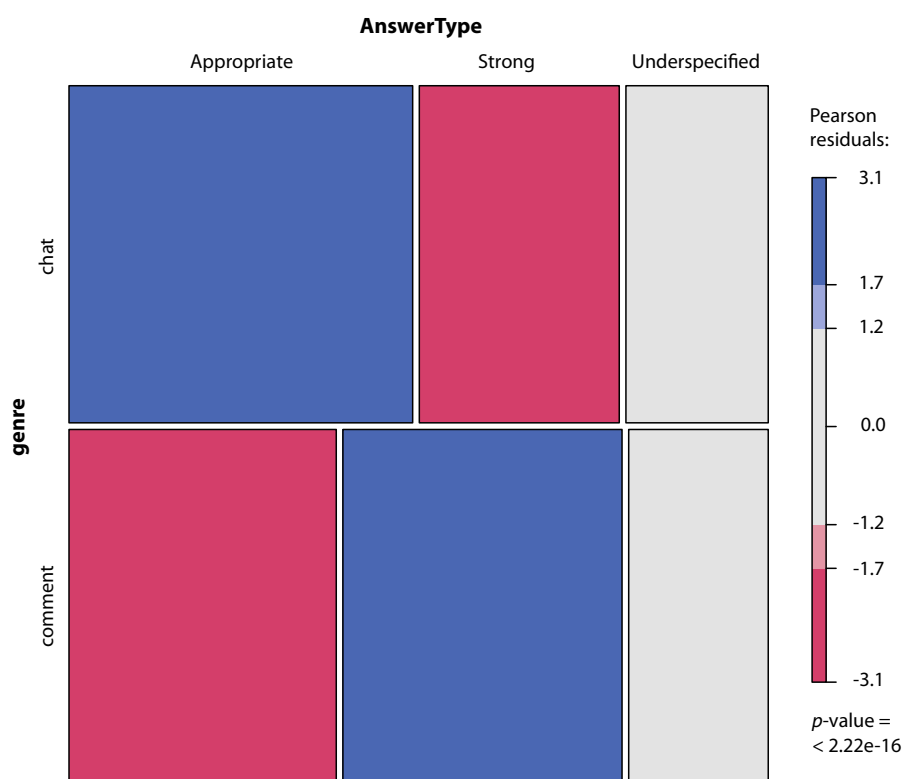
Figure 6 – Distribution of Pearson residuals for connective type across genres

33   Finally, we ran a mixed-effect logistic regression in order to test the significance of relation type and genre on whether participants inserted *and* or another connective (Appendix 1). The regression returns a significant negative effect of contrastive relations on insertions of *and* ($p < 0.001$), which confirms that underspecified contrast is not a preferred choice of participants in this production task. Genre was not significant. This first production experiment thus confirms the corpus findings and seems to validate our HP1 (relation type) but not HP2 (genre). We will now test whether the same effects hold in an acceptability task, thus comparing production with perception.

5.   Acceptability: forced choice task

34   In this second experiment, we are reducing the participants' preferences to a binary choice between a weaker and a stronger version of the same pair of sentences. With the same experimental materials and conditions, we are looking to further confirm that underspecified connectives are dispreferred over a stronger competitor in contrastive relations and in more formal texts.

Figure 7 – Display of the forced-choice task

5.1.   Materials

35   The materials for this study are the same as for the previous one. To ensure the validity of the analysis, we doubled the size of the lists and added 12 more test sentences (plus 12 distractors) per list of materials. They were created following the same method as the others (extracted from the Loyola CMC corpus, pre-tested by naive participants who disambiguated the meaning of *and*, changed into a more formal version). We now have four lists of 48 items: 16 consequence (6 *and*, 6 *so*, 4 *therefore*), 16 contrast (6 *and*, 6 *but*, 4 *however*) and 16 distractors (8 causal, 8 temporal). Half of the items are in the chat version, the other half in the comment version. The conditions are balanced across lists.

5.2.   Procedure

36   Participants were presented with a split screen containing two versions of the same text, only the connective changed: the original connective from the corpus (*and*, *but*, *so*, *however*, *therefore* plus distractors) was confronted to either a weaker or a stronger alternative. There was only one degree of strength difference between the two connectives in each trial (e.g., *and* vs. *but*, *but* vs. *however*, but not *and* vs. *however*). The connectives were highlighted in red on the screen so as to draw the participants' attention to them (see Figure 7). The participants had to click on the button underneath the version they thought was more acceptable.

37   The experiment was programmed on LingoTurk and run on Prolific Academic. Thirty native English-speaking participants were recruited and remunerated (8€ per hour). Each participant only saw each item once in one condition, and only did one list. Weaker and stronger alternatives were balanced across conditions and lists (*therefore* and *however* only have weaker alternatives).

5.3.   Results

38   We report results for the test sentences, excluding filler sentences (causal and temporal relations), which amounts to 7,104 answers. We start by exploring whether participants systematically preferred the stronger option in contrast and in comments, according to the hypotheses. Overall, stronger options were preferred in around two

thirds of the trials. The data presents very small differences between conditions, but all in the expected direction: stronger versions are preferred in 64.86% of contrastive relations vs. 62.95% for consequence, and in 65.74% of the comment materials vs. 62.08% of the chat materials. A mixed-effect logistic regression (with random slope on the genre variable) indicates that only the genre effect is significant ($p < 0.01$), thus only confirming HP2 on the whole dataset (Appendix 2). This finding corroborates the result of the production experiment, where we found a preference for strong connectives in comments.

³⁹ However, this analysis concerns all weak vs. strong alternatives, including cases where *and* was not an option. Given the focus on *and* in this study, we now report a regression model on a subset of the data including only trials where *and* was an option against *but* or *so* (see Appendix 3). This mixed model returns a significant negative fixed effect of the contrastive relation on the selection of *and* ($p < 0.05$), and a marginally significant negative effect of the comment genre ($p < 0.1$). In other words, participants prefer *but* over *and* in contrast, and tend to prefer *but* and *so* over *and* in comments, thus confirming both our hypotheses, although the effect size is quite small.

6. General discussion

⁴⁰ The three studies partially converge in identifying a stable effect of relation type and genre on the production and perception of underspecified relations of consequence and contrast marked by *and*. The incompatibility of *and* with contrastive contexts is robust across all tasks. However, the association between underspecified *and* and informal genres, observed in the corpus, is only marginally significant in the perception study, and was not found in the production experiment, where the formal genre only affects the proportion of strong connectives such as *however*.

⁴¹ While we had expected to find a consistent genre effect, we only observed this effect in the corpus study, and found a marginal trend in the expected direction in the acceptability judgement task. There are several explanations for this small difference between the two genres. One obvious limitation of this study is that the genres included in the experiments do not differ strongly: they are both forms of online computer-mediated communication, and therefore lack the high formality and professionalism of printed texts such as academic or newspaper articles. Although comments are neither conversational nor spontaneous, the difference between chats and comments is more a matter of degree, one of them being more spoken-like (chat), the other more written-like (comments). A more contrasted choice of genres could lead to larger differences in connective choice.

⁴² A related issue is the unnatural setting of any experiment: behaviour observed in an experimental setting differs from natural behaviour outside the lab, primarily due to the so-called "Hawthorne effect", whereby participants overestimate what is expected of them and therefore try to perform "better" than they would normally do. In the context of psycholinguistic experiments, this could translate into highly

normative answers. In our case, participants may have preferred stronger options over *and* across the board, regardless of the genre, as an artifact of this normative effect. This factor, although hard to test, cannot be ruled out in the explanation of our small effect size.

43   Moreover, the linguistic manipulation of the original items to turn them into more formal versions was minimal for some sentences, in order to make items comparable and make sure that the inferred coherence relation for the two versions remained identical. In particular, when the original sentences were short, there was no obvious step to be taken to strengthen the difference between the two versions, as in Example [4] below.

[4a]   my parents are old devout catholics // they would totally support a more progressive Church

[4b]   My parents are old and devout Catholics // they would absolutely support a more progressive Church.

44   In fact, we observe a large variance between items (see regression coefficients for the *and* subset in Appendix 3). Making further modifications to the sentences could address this issue, at the risk of creating large differences in the interpretation of the discourse relation.

45   Another explanation for the small effect size could be found in the presence of segment-internal features that compensate for the underspecification of *and*, thus making the utterance acceptable. There is a growing trend of research into discourse signals besides connectives, which tends to show that connectives often co-occur with other devices that help interpret the discourse relation (Crible, in press; Das & Taboada, 2018; Hoek et al., 2018; Morera et al., 2010). Our hypothesis was based on the fact that underspecification creates a peak in information density and constrains the listener to make an extra interpretation effort. However, if the lack of informativeness of *and* is immediately compensated by other elements in its co-text, then it is possible that the interpretation cost remains low. Under such conditions, *and* can occur in underspecified relations, even in formal genres. Some of our materials include such compensation discourse signals:

[5]   an Iraqi whose family was killed in the gas attacks will love us for hanging Saddam // one who has to commute to Baghdad for work every day and wait in a 3+hour checkpoint line will be frustrated

[6]   some people like things the way they are // many don't

[7]   I know that in my discussion group everyone else was against the Patriot Act // it seems here that everyone is for it

46   In [5], the contrast is further marked by the syntactic parallelism between the two sentences. In [6], it is reinforced by the negation and the verb elision in the

second segment. In [7], a cluster of signals including the antonyms "against" and "for", the difference in verb tense (past vs. present) and the deictic marker "here" converge in signalling contrast. Corpus studies of these phenomena are developing, but experimental approaches are still lacking (Grisot & Blochowiak, 2017).

47   Another point open for discussion is the restriction to two factors of variation in this study. We have shown that underspecification does not apply equally to all relations and all genres. However, there might be other reasons why speakers are vague in their discourse, which were not accounted for here. For instance, we may lack the cognitive resources to come up with a more specific connective because of a physiological factor or a sudden trouble in lexical access; we can be deliberately vague to deceive our audience, or because we do not want to commit to a specific interpretation and our idea itself is vague, or simply for stylistic reasons and individual preferences.

48   Finally, this study points at a partial overlap between production and perception data, and between corpus-based and experimental methods. While triangulation of evidence is desirable in research, our results suggest that, as far as underspecification is concerned, it is not trivial to achieve. The strong differences that we found in corpus data were only partially confirmed experimentally. It is generally argued that language production and comprehension are highly intertwined mechanisms (e.g., Pickering & Garrod, 2007). Less is known for perception tasks such as acceptability judgements (see Didirková et al., 2018). For future work, using comprehension tasks on the same data would further extend our understanding of how we produce and interpret language.

7.   Conclusion

49   In this paper, we investigated the effect of discourse relation type and genre on the production and perception of underspecified relations of contrast and consequence signalled by *and*. We combined a corpus study, a production experiment and a perception experiment in order to test two hypotheses: (1) that *and* is more compatible with relations of consequence than of contrast, due to factors of cognitive complexity and conceptual differences; (2) that *and* is more compatible with informal than formal genres, because of requirements of recipient design. The corpus study confirmed that *and* is more often underspecified in informal settings, and that it does not express contrast in formal settings.

50   Our production experiment (connective insertion task) only confirmed the effect of relation type, with very few insertions of *and* in contrastive relations, but no effect of genre. The perception experiment (forced-choice task) converges in further identifying a stable effect of relation type (contrastive uses of *and* are dispreferred), while the effect of genre is only marginal.

51   This study addresses a frequent phenomenon, viz. underspecification, which has been somewhat overlooked in the literature so far, typically focused on more

specific connectives. It also presents data from computed-mediated genres, which are only emerging in discourse analysis and in particular in experimental studies, which traditionally use newspaper articles (see Degand & Van Bergen, 2018, for an exception). It therefore constitutes an innovative addition to the existing body of research on discourse connectives. We call for further investigation of the production and comprehension of underspecification, jointly with other related phenomena (e.g., other discourse signals, cf. Das & Taboada, 2018), within or outside the field of discourse.

References

ASHER, N. & LASCARIDES, A. 2003. *Logics of Conversation*. Cambridge – New York – Melbourne: Cambridge University Press.

ASR, F.T. & DEMBERG, V. 2012. Measuring the Strength of Linguistic Cues for Discourse Relations. In E. HAJIČOVÁ, L. POLÁKOVÁ & J. MÍROVSKÝ (eds.), *Proceedings of the Workshop on Advances in Discourse Analysis and its Computational Aspects – ADACA*. Stroudsburg: Association for Computational Linguistics: 33-42. Available online: https://www.aclweb.org/anthology/W12-4703/.

ASR, F.T. & DEMBERG, V. 2015. Uniform Information Density at the Level of Discourse Relations: Negation Markers and Discourse Connective Omission. In M. PURVER, M. SADRZADEH & M. STONE (eds.), *Proceedings of the 11th International Conference on Computational Semantics – IWCS*. Stroudsburg: Association for Computational Linguistics: 118-128. Available online: https://www.aclweb.org/anthology/W15-0117/.

ASR, F.T. & DEMBERG, V. 2016. "But" vs. "Although" under the Microscope. In A. PAPAFRAGOU, D. GRODNER, D. MIRMAN & J.C. TRUESWELL (eds.), *Proceedings of the 38th Annual Conference of the Cognitive Science Society – COGSCI 2016*. Austin: Cognitive Science Society: 366-371. Available online: https://cogsci.mindmodeling.org/2016/papers/0075/paper0075.pdf.

BARTH-WEINGARTEN, D. 2014. Dialogism and the Emergence of Final Particles: The Case of "And". In S. GÜNTHNER, W. IMO & J. BÜCKER (eds.), *Grammar and Dialogism: Sequential, Syntactic, and Prosodic Patterns between Emergence and Sedimentation*. Berlin: De Gruyter: 335-365.

BIBER, D. 2006. *University Language: A Corpus-Based Study of Spoken and Written Registers*. Amsterdam – Philadelphia: J. Benjamins.

CAIN, K. & NASH, H.M. 2011. The Influence of Connectives on Young Readers' Processing and Comprehension of Text. *Journal of Educational Psychology* 103 (2): 429-441.

CRIBLE, L. 2017. Discourse Markers and (Dis)fluencies in English and French: Variation and Combination in the DisFrEn Corpus. *International Journal of Corpus Linguistics* 22 (2): 242-269.

CRIBLE, L. 2018. *Discourse Markers and (Dis)fluency. Forms and Functions across Languages and Registers*. Amsterdam – Philadelphia: J. Benjamins.

CRIBLE, L. in press. Weak and Strong Discourse Markers in Speech, Chat and Writing: How Signals Compensate for Ambiguity in Explicit Relations. *Discourse Processes*.

CRIBLE, L., ABUCZKI, A., BURKŠAITIENĖ, N., FURKÓ, P., NEDOLUZHKO, A., RACKEVIČIENĖ, S., VALŪNAITĖ OLEŠKEVIČIENĖ, G. & ZIKÁNOVÁ, Š. 2019. Functions and Translations of Discourse Markers in TED Talks: A Parallel Corpus Study of Underspecification in Five Languages. *Journal of Pragmatics* 142: 139-155.

CRIBLE, L. & DEGAND, L. 2019. Reliability vs. Granularity in Discourse Annotation: What Is the Trade-off? *Corpus Linguistics and Linguistic Theory* 15 (1): 71-99.

DAS, D. & TABOADA, M. 2018. Signalling of Coherence Relations in Discourse, beyond Discourse Markers. *Discourse Processes* 55 (8): 743-770.

DEGAND, L. & VAN BERGEN, G. 2018. Discourse Markers as Turn-Transition Devices: Evidence from Speech and Instant Messaging. *Discourse Processes* 55 (1): 47-71.

DIDIRKOVÁ, I., CRIBLE, L. & SIMON, A.C. 2018. Impact of Prosody on the Perception and Interpretation of Discourse Relations: Studies on "Et" and "Alors" in Spoken French. *Discourse Processes* 56 (8): 619-642.

EVERS-VERMEUL, J. & SANDERS, T.J.M. 2009. The Emergence of Dutch Connectives: How Cumulative Cognitive Complexity Explains the Order of Acquisition. *Journal of Child Language* 36 (4): 829-854.

FRANK, A. & GOODMAN, N. 2012. Predicting Pragmatic Reasoning in Language Games. *Science* 336 (6084): 998.

GOLDSTEIN-STEWART, J., GOODWIN, K.A., SABIN, R.E. & WINDER, R.K. 2008. Creating and Using a Correlated Corpus to Glean Communicative Commonalities. In N. CALZOLARI, K. CHOUKRI, B. MAEGAARD, J. MARIANI, J. ODIJK, S. PIPERIDIS, D. TAPIAS (eds.), *Proceedings of the 6th International Conference on Language Resources and Evaluation – LREC 2008*. Luxembourg: European Language Resources Association: 3029-3035. Available online: https://www.aclweb.org/anthology/L08-1198/.

GÓMEZ TXURRUKA, I. 2003. The Natural Language Conjunction "And". *Linguistics and Philosophy* 26 (3): 255-285.

GRISOT, C. & BLOCHOWIAK, J. 2017. Temporal Connectives and Verbal Tenses as Processing Instructions: Evidence from French. *Pragmatics & Cognition* 24 (3): 404-440.

HOEK, J., ZUFFEREY, S., EVERS-VERMEUL, J. & SANDERS, T.J.M. 2017. Cognitive Complexity and the Linguistic Marking of Coherence Relations: A Parallel Corpus Study. *Journal of Pragmatics* 121: 113-131.

HOEK, J., ZUFFEREY, S., EVERS-VERMEUL, J. & SANDERS, T.J.M. 2018. The Linguistic Marking of Coherence Relations: Interactions between Connectives and Segment-Internal Elements. *Pragmatics & Cognition* 25 (2): 276-309.

HORN, L.R. 1984. Toward a New Taxonomy for Pragmatic Inference: Q-Based and R-Based Implicature. In D. SCHIFFRIN (ed.), *Meaning, Form, and Use in Context: Linguistic Applications*. Washington: Georgetown University Press: 11-42.

KITIS, E. 2000. Connectives and Frame Theory: The Case of Hypotextual Antinomial "And". *Pragmatics & Cognition* 8 (2): 357-409.

LEVY, R. & JAEGER, T.F. 2007. Speakers Optimize Information Density through Syntactic Reduction. In B. SCHÖLKOPF, J.C. PLATT & T. HOFMANN (eds.), *Advances in Neural Information Processing Systems 19: Proceedings of the 2006 Conference*. Cambridge – London: The MIT Press: 849-856.

Luscher, J.-M. & Moeschler, J. 1990. Approches dérivationnelles et procédurales des opérateurs et connecteurs temporels: les exemples de "et" et de "enfin". *Cahiers de linguistique française* 11: 77-104.

Mann, W.C. & Thompson, S.A. 1988. Rhetorical Structure Theory: Toward a Functional Theory of Text Organization. *Text* 8 (3): 243-281.

Morera, Y., De Vega, M. & Camacho, J. 2010. Differences in Continuity of Force Dynamics and Emotional Valence in Sentences with Causal and Adversative Connectives. *Cognitive Linguistics* 21 (3): 501-536.

Murray, J.D. 1997. Connectives and Narrative Text: The Role of Continuity. *Memory & Cognition* 25 (2): 227-236.

Nelson, G., Wallis, S. & Aarts, B. 2002. *Exploring Natural Language: Working with the British Component of the International Corpus of English*. Amsterdam – Philadelphia: J. Benjamins.

Pander Maat, H. 1999. The Differential Linguistic Realization of Comparative and Additive Coherence Relations. *Cognitive Linguistics* 10 (2): 147-184.

Petukhova, V. & Bunt, H. 2009. Towards a Multidimensional Semantics of Discourse Markers in Spoken Dialogue. In H. Bunt, V. Petukhova & S. Wubben (eds.), *Proceedings of the 8th International Conference on Computational Semantics – IWCS*. Stroudsburg: Association for Computational Linguistics: 157-168. Available online: https://www.aclweb.org/anthology/W09-3715/.

Pickering, M.J. & Garrod, S. 2007. Do People Use Language Production to Make Predictions during Comprehension? *Trends in Cognitive Sciences* 11 (3): 105-110.

Prasad, R., Dinesh, N., Lee, A., Miltsakaki, E., Robaldo, L., Joschi, A. & Webber, B. 2008. The Penn Discourse Treebank 2.0. In N. Calzolari, K. Choukri, B. Maegaard, J. Mariani, J. Odijk, S. Piperidis, D. Tapias (eds.), *Proceedings of the 6th International Conference on Language Resources and Evaluation – LREC 2008*. Luxembourg: European Language Resources Association: 2961-2968. Available online: https://www.aclweb.org/anthology/L08-1093/.

Pusse, F., Asad, S. & Demberg, V. 2016. LingoTurk: Managing Crowdsourced Tasks for Psycholinguistics. In J. DeNero, M. Finlayson & S. Reddy (eds.), *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Demonstrations – NAACL-HLT*. Stroudsburg: Association for Computational Linguistics: 57-61. Available online: https://www.aclweb.org/anthology/N16-3012/.

Sanders, T.J.M. 2005. Coherence, Causality and Cognitive Complexity in Discourse. In M. Aurnague, M. Bras, A. Le Draoulec, L. Vieu (eds.), *Proceedings of the First International Symposium on the Exploration and Modelling of Meaning – SEM-05*. Toulouse: Université de Toulouse: 105-114.

Sanders, T.J.M. & Noordman, L.G.M. 2000. The Role of Coherence Relations and their Linguistic Markers in Text Processing. *Discourse Processes* 29 (1): 37-60.

Sanders, T.J.M., Spooren, W.P.M. & Noordman, L.G.M. 1993. Coherence Relations in a Cognitive Theory of Discourse Representation. *Cognitive Linguistics* 4 (2): 93-133.

Schiffrin, D. 1987. *Discourse Markers*. Cambridge: Cambridge University Press.

Schiffrin, D. 2006. Discourse Marker Research and Theory: Revisiting "And". In K. Fischer (ed.), *Approaches to Discourse Particles*. Amsterdam – London – Paris: Elsevier: 315-338.

SCHMIDT, T. & WÖRNER, K. 2009. EXMARaLDA – Creating, Analyzing and Sharing Spoken Language Corpora for Pragmatic Research. *Pragmatics* 19 (4): 565-582.

SPOOREN, W.P.M. 1997. The Processing of Underspecified Coherence Relations. *Discourse Processes* 24 (1): 149-168.

TABOADA, M. 2006. Discourse Markers as Signals (or Not) of Rhetorical Relations. *Journal of Pragmatics* 38 (4): 567-592.

TABOADA, M. 2009. Implicit and Explicit Coherence Relations. In J. RENKEMA (ed.), *Discourse, of Course: An Overview of Research in Discourse Studies.* Amsterdam – Philadelphia: J. Benjamins: 127-140.

Appendix: regression coefficients

Production experiment

AnswerCorrect ~ Relation + Genre + (1|participant) + (1|item)

AIC	BIC	logLik	deviance	df.resid
1,869.0	1,897.0	-929.5	1,859.0	2,025

Random effects:

Groups	Name	Variance	Std.Dev.
Participant	(Intercept)	0.4764	0.6902
Item	(Intercept)	0.3691	0.6075

Number of obs: 2,030; groups: participants, 113; items, 20

Fixed effects:

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-0.86946	0.18895	-4.602	4.19e-06 ***
RelationContrast	-1.39728	0.15285	-9.141	< 2e-16 ***
GenreComment	-0.07565	0.14027	-0.539	0.59

Perception experiment (all data)

Selection ~ Relation + Genre + (1+genre|participant) + (1+genre|item)

AIC	BIC	logLik	deviance	df.resid
9,135.1	9,196.9	-4,558.5	9,117.1	7,095

Random effects:

Groups	Name	Variance	Std.Dev.	Corr
Participant	(Intercept)	0.4764	0.6902	
	GenreComment	0.05055	0.2248	-0.62
Item	(Intercept)	0.3691	0.6075	
	GenreComment	0.02169	0.1473	0.44

Number of obs: 7,104; groups: participants, 267; items, 32

Fixed effects:

	Estimate	Std. Error	*z* value	Pr(>|z|)
(Intercept)	-0.52509	0.08341	-6.295	3.08e-10 ***
RelationContrast	-0.00198	0.10511	-0.019	0.98497
GenreComment	-0.16817	0.06047	-2.781	0.00542 **

Perception experiment ("and" subset)

Selection - Relation + Genre + (1|participant) + (1|item)

AIC	BIC	logLik	deviance	df.resid
3,056.2	3,085.7	-1,523.1	3,046.2	2,716

Random effects:

Groups	Name	Variance	Std.Dev.
Participant	(Intercept)	0.3766	0.6137
Item	(Intercept)	0.5698	0.7549

Number of obs: 2,721; groups: participants, 267; items, 24

Fixed effects:

	Estimate	Std. Error	*z* value	Pr(>|z|)
(Intercept)	-0.67024	0.22066	-3.037	0.00239 **
RelationContrast	-0.60691	0.27701	-2.191	0.02846 *
GenreComment	-0.16509	0.09097	-1.815	0.06956 ·