Conditionals, Inference, and Evidentiality

Karolina Krzyżanowska[1], Sylvia Wenmackers[1], Igor Douven[1], and Sara Verbrugge[2]

[1] Faculty of Philosophy, University of Groningen, Oude Boteringestraat 52, 9712 GL Groningen, The Netherlands
[2] Department of Psychology, University of Leuven, Tiensestraat 102, 3000 Leuven, Belgium

Abstract. At least many conditionals seem to convey the existence of a link between their antecedent and consequent. We draw on a recently proposed typology of conditionals to revive an old philosophical idea according to which the link is inferential in nature. We show that the proposal has explanatory force by presenting empirical results on two Dutch linguistic markers.

1   Introduction

Although various theories of conditional sentences have been proposed, none of them seems to account for all the empirical data concerning how people use and interpret such sentences.[3] It is almost universally agreed that no theory of conditionals counts as empirically adequate if it validates sentences like these:

(1) a. If badgers are cute, then $2 + 2 = 4$.
     b. If weasels are vegetables, then unicorns hold Frege in particularly high esteem.

It is easy to understand why people are reluctant to accept conditionals like (1a) and (1b): the antecedents of those conditionals have nothing to do with their consequents. And it seems that using a conditional construction is meant to convey the existence of some sort of link between the content of the if-clause and the content of the main clause.

But what kind of link might this be? According to an old philosophical idea, the link is inferential in nature. If this idea is currently unpopular, at least in philosophy, that may be due to the fact that theorists have failed to recognize that the inferential connection need not be of one and the same type. We draw on a recently proposed typology of conditionals to reinstate the old idea, at least for a large class of conditionals that linguists have aptly termed "inferential conditionals." We buttress our proposal by presenting experimental results concerning two Dutch linguistic markers.

---

[3] For a survey of various accounts of conditionals and problems they are facing see, for instance, Edgington [1] or Bennett [2].

2 Inferential Conditionals

Probably the most general distinction to be made when it comes to conditionals is that between indicative and subjunctive conditionals. In this paper, we will only be concerned with indicative conditionals.[4] For many theorists this is only the beginning of a typology, though there is little unanimity as to what the typology should further look like. What has become common in linguistics, but is rarely encountered in the philosophical literature, is to classify indicative conditionals as *inferential* and *content* conditionals.[5] The class of content conditionals is not particularly well defined—its members are sometimes loosely said to describe relations between states of affairs or events as they happen in reality—but those will not concern us here. We are going to limit our attention to inferential conditionals, that is conditional expressing a reasoning process, having the conditional's antecedent as a premise and its consequent as the conclusion, for example: "If she has not had much sleep recently, she will perform poorly on her exams" or "If he owns a Bentley, he must be rich." Arguably, these constitute a common, if not the most common, type among the conditionals we encounter in natural language.

That a conditional sentence can be considered as a kind of "condensed argument" [8, p. 15] is not altogether new to philosophy; it can be traced back at least to Chrysippus, a stoic logician from the third century BC. He is believed to have held the view that a conditional is true if it corresponds to a valid argument [9]. Obviously, if we limit our understanding of a valid argument only to the classical deductive inference, we can easily find counterexamples to the above claim. Yet deduction is not the only type of reasoning people employ, and a plausible theory of inferential conditionals should not neglect this fact.

Although linguists have proposed various finer-grained typologies of inferential conditionals (see, e.g., Declerck and Reed [6]), most of these stem from grammatical distinctions. However, we are interested in a specific typology recently presented in Douven and Verbrugge [10] who acknowledge the variety of inferential relations between a conditional's antecedent and its consequent. A first distinction these authors make is that between *certain* and *uncertain* inferences, where certain inferences guarantee the truth of the conclusion given the truth of the premises while uncertain inferences only tend to make the truth of the conclusion likely given the truth of the premises; the former are standardly referred to as "deductive inferences." Douven and Verbrugge further follow standard philosophical usage in grouping uncertain inferences into *abductive* and *inductive* ones, where the former are inferences based on explanatory considerations and the latter are inferences based on statistical information. More exactly, in an abductive inference we infer a conclusion from a set of premises because

---

[4] The distinction between indicative and subjunctive conditionals is not as clear-cut as one might wish, but the conditionals we used in our materials were all uncontroversial cases of indicative conditionals. Henceforth, indicative conditionals will be referred to simply as "conditionals."

[5] See, among others, Dancygier [3,4], Dancygier and Sweetser [5], Declerck and Reed [6], and Haegeman [7].

the conclusion provides the best explanation for those premises; for example, we may infer that Sally failed her exam from the premises that Sally had an exam this morning and that she was just seen crying and apparently deeply unhappy: that she failed the exam is the best explanation for her apparent unhappiness. Inductive inferences rely on information about frequencies (which may be more or less precisely specified); for instance, we infer that Jim is rich from the premise that he owns a Bentley because we know that by far the most Bentley owners are rich. The validity of both abductive and inductive inferences is a matter of controversy and ongoing debate. It is largely uncontested, however, that people do engage in these types of inferences on a routine basis.

Douven and Verbrugge's typology of inferential conditionals follows the aforesaid typology of inference. That is to say, they distinguish between certain (or deductive) and uncertain inferential conditionals, and then divide the latter class further into abductive and inductive inferential conditionals. More specifically, they propose the following:

Definition 1. *"If p, then q" is a deductive inferential (DI, for short) / inductive inferential (II) / abductive inferential (AI) conditional if and only if q is a deductive / inductive / abductive consequence of p.*

(Various formalizations of the inductive and abductive consequence relations have been offered in the literature, though, like Douven and Verbrugge in their paper, we refrain from committing to any in particular.) Douven and Verbrugge note also that, often, the inference may rely on the antecedent p together with background assumptions that are salient in the context in which the conditional is asserted or evaluated. Such conditionals are called *contextual* DI, AI, or II conditionals, depending on the type of inference involved.[6]

Definition 2. *"If p, then q" is a contextual DI / II / AI conditional if and only if q is a deductive / inductive / abductive consequence of $\{p, p_1, \ldots, p_n\}$, with $p_1, \ldots, p_n$ being background premises salient in the context in which "If p, then q" is asserted or evaluated.*

Douven and Verbrugge do not claim that their typology is *correct* and the ones that so far have been propounded by other theorists are *incorrect*. Indeed, it might even be odd to think that there are *natural kinds* of conditionals that a typology should try to chart. What they do claim is that their typology is exceedingly simple and that it is non-ad hoc in that it relies on a time-tested distinction between types of inference. More importantly still, they show in their 2010 paper that the typology has considerable explanatory force by recruiting it in service of testing a thesis, first proposed by Adams [11] and championed by many since, according to which the acceptability of a conditional is measured by the probability of its consequent conditional on its antecedent.

---

[6] As Douven and Verbrugge [10, p. 304] note, in contextual AI conditionals, the consequent need not always be the best explanation of the antecedent. It may also be that the consequent is, *in light of the antecedent*, the best explanation of one of the background assumptions.

We expand here on the main experiment of Douven and Verbrugge's paper, because it served as the starting point for our own empirical work. In their experiment, Douven and Verbrugge divided the participants into two groups, asking one group to judge the acceptability of ten DI, ten AI, and ten II conditionals and the other group to judge the corresponding conditional probabilities. This is an example of a question asking for the acceptability of an AI conditional:[7]

> CONTEXT: Judy is waiting for a train. She is looking for her iPod. It is not in her coat. She suddenly sees that the zipper of her bag is open. She cannot remember having opened it. It is announced that there are pickpockets active in the train station.
>
> CONDITIONAL: If Judy's iPod is not in her bag, then someone has stolen it.
>
> Indicate how acceptable you find this conditional in the given context:
>
> Highly unacceptable   1   2   3   4   5   6   7   Highly acceptable

The corresponding question asking for the conditional probability is this:

> CONTEXT: Judy is waiting for a train. She is looking for her iPod. It is not ain her coat. She suddenly sees that the zipper of her bag is open. She cannot remember having opened it. It is announced that there are pickpockets active in the train station. Suppose that Judy's iPod is not in her bag.
>
> SENTENCE: Someone has stolen Judy's iPod.
>
> Indicate how probable you find this sentence in the given context:
>
> Highly improbable   1   2   3   4   5   6   7   Highly probable

The results obtained in this experiment show that Adams' thesis holds for DI conditionals at best, and that for AI conditionals the most that can be said is that acceptability and conditional probability are highly correlated; for II conditionals not even that much is true.

This already shows that the typology of inferential conditionals proposed by Douven and Verbrugge has considerable explanatory force. Here, we aim to extend the case for this typology by relating it to two putative evidential markers. Before we move on to report our experimental results on these markers (in section 4), we briefly discuss the concept of an evidential marker.

3   Evidential Markers

In some languages, it is common for speakers to communicate information about the evidential grounds for the contents of their assertions. Some languages also possess a rich arsenal of prefixes, suffixes, particles, and other linguistic items for

---

[7] See Appendix A of Douven and Verbrugge [10] for the full materials used in this experiment.

42

this purpose.[8] European languages do not encode evidentiality grammatically, but this is not to say that speakers of European languages do not have the resources to indicate evidential grounds, or that they never use those resources. Sentences like:

(2) Adam works hard.

do not indicate the speaker's source of information at all—the speaker can actually share an office with Adam and see him working hard on an everyday basis, but she could also have inferred (2) from Adam's work output or simply heard someone else say so.

But other sentences do suggest the speaker's evidential grounds. For instance, when we are wondering about the translation of a phrase in Latin and we know that Susan studied classical languages for a number of years, we might say

(3) Susan should be able to translate this phrase.

This assertion would seem odd if we knew that (say) the phrase is from a text which Susan recently published in English translation. Similarly, when an English speaker tries to call her friend but does not get an answer, she may infer that her friend is out and express the resulting belief by saying:

(4) She must be out.

Were she to see her friend walking on the street, her assertion of (4) would again seem odd or even inappropriate.

Some authors have argued that the modal auxiliary verb "must" makes the assertion weaker than the one without it.[9] But this is not generally true. For sometimes "must" seems to indicate the necessity of what has been asserted: if one knows that Mary has put a bottle of wine either in the fridge or in the cupboard, and one has checked that it is not in the cupboard, it seems natural for one to conclude:

(5) It must be in the fridge.

As noticed by von Fintel and Gillies [17,18], what (4) and (5) have in common in the first place is that they signal the presence of an inference. Specifically, the verb "must" indicates that the speakers' grounds for their assertions are inferential, and hence indirect, but as they also argue, this need not mean that

---

[8] On the basis of data from 32 languages, Willett [12] proposed a taxonomy of markers encoding main types of sources of information. The main distinction is between direct (perceptual) and indirect access, and the latter can be further divided into other speaker's reports and inference. Since then, various aspects of evidentiality in language have been investigated, like for instance a developmental study by Papafragou and colleagues [13] on Korean speaking children's learning of evidential morphology and their ability to reason about sources of information.

[9] See e.g. Karttunen [14, p. 12], Groenendijk and Stokhof [15, p. 69] or Kratzer [16, p. 645]

these grounds are weak or inconclusive. Again, some confusion could have been avoided if the variety of inference relations had been attended to.

Of course, sometimes we convey information about our evidential grounds in more direct ways, as when we say that we *saw* that John crossed the street, or that it *seems* to us that Harriet is worried, or by the use of such words as "probably," "presumably," "possibly," "apparently," "allegedly," "putatively," and so on. But in this paper we focus on "must" and "should" in their roles as evidential markers, or rather, we focus on their Dutch counterparts "moet wel" ("must"[10]) and "zal wel" ("will," "should"). Evidential markers can serve a number of purposes. For instance, they may be used to indicate the *source* of the speaker's evidence: whether it is perceptual evidence, or evidence from testimony, or evidence from some third type of source still. They may also be used to indicate the *quality* of the evidence (e.g., indicate how reliable the source was). We will be mainly interested in the question of whether "moet wel" and "zal wel" play any distinctive role in signaling the *kind* of inference that is involved in making whatever evidence the speaker has bear on the content of her assertion. Can anything systematic be said about whether these markers go better with some type or types of inference than with others?

To clarify this question, note that the inference underlying the assertion of (4) in our example is most plausibly thought of as being abductive, that is, as an inference to the best explanation: that the friend is out is the best explanation for the evidence that the speaker has, to wit, that her friend does not answer the phone. In the example of Susan, it rather seems to be some form of inductive reasoning that warrants the assertion of (3): the people we met in our lives who had studied classical languages for a number of years were typically able to translate Latin phrases; given that Susan studied classical languages for a number of years, we expect her to be able to translate the designated phrase. Naturally, the inferential connection between evidence and grounds for assertion may also be deductive, as in the case of (5). The question we are interested in is whether the use of "moet wel" and "zal wel" gives us any indication as to what kind of inference (if any at all) led the speaker to feel warranted in making the assertion she did on the basis of the evidence she had.

4 Experiment: Linguistic Markers

Our experiment makes use of the typology of inferential conditionals discussed above. We look at a number of instances of the various types whose degrees of acceptability have been ascertained in previous research and we look whether these degrees are affected by inserting "moet wel" or "zal wel" into the sentences. We encountered the putative English counterparts of these markers already in our examples involving (4) and (3). "Must" and "zal wel" have also been described as inferential markers in the literature. As for the latter, Verbrugge [20] established a close connection between "zal wel" and inferential conditionals.

---

[10] "Wel" is a positive polar marker which has no counterpart in English; see Nuyts and Vonk [19, p. 701]. In German, "bestimmt" comes close.

Specifically, in an elicitation task in which they were requested to complete conditionals whose antecedents were given, participants tended to come up with a significantly higher number of inferential conditionals (as opposed to content conditionals) when they were in addition requested to use "zal wel" in the consequents than when they were not. As for "must," Dietz [21, p. 246] notes that in "It must be raining," the auxiliary indicates that the speaker only has (what he calls) "inferential evidence," and no direct observational evidence, that it is raining; see also the papers by von Fintel and Gillies cited earlier as well as Anderson [22], Papafragou [23], and Nuyts and Vonk [19]. However, so far researchers have not considered differentiating between the various types of inference by means of the said markers. Might not one marker go better with the expression of one type of inference and the other with the expression of another type of inference? Even more fundamentally, is "must" really an inferential marker? We are not aware of any empirical evidence that warrants a positive answer. Philosophers may be convinced that "must" can serve as an evidential marker, but philosophers were also convinced that the acceptability of a conditional is equal to the corresponding conditional probability, and—as Douven and Verbrugge showed—empirical evidence gives the lie to that thought.

To investigate the aforementioned questions, we used the materials of the experiment described in the previous section. We inserted "moet wel" and "zal wel" into the consequents of the conditionals used in Douven and Verbrugge's main experiment and checked whether this made a difference to acceptability ratings and their correlations with probability ratings. We were particularly interested in the effect the presence of the auxiliaries has on these correlations for different types of conditionals.

4.1 Method

PARTICIPANTS
Fifty seven students of the University of Leuven took part in the experiment.

DESIGN
The type of conditional (DI / AI / II) was manipulated within subjects. The different lexical markers were manipulated between subjects.

MATERIALS AND PROCEDURE
All materials were in Dutch, the participants' mother tongue. Thirty items were presented in a booklet. Every participant had to evaluate ten abductive, ten inductive, and ten deductive items. Items were randomized per booklet and the booklets were randomly distributed in the lecture hall. The items consisted of the same context–sentence pairs that were presented to the participants in the main experiment of Douven and Verbrugge [2010] who were asked to judge the acceptability of conditionals, except that the conditionals now contained the markers "moet wel" and, respectively, "zal wel." For instance, to the AI conditional "Als Judy's iPod niet in haar tas zit, dan heeft iemand die gestolen" ("If Judy's iPod is not in her bag, then someone has stolen it"), whose acceptability participants in Douven and Verbrugge's experiment had been asked to grade, corresponded

in our experiment the AI conditionals "Als Judy's iPod niet in haar tas zit, dan moet iemand die wel gestolen hebben" ("If Judy's iPod is not in her bag, then someone must have stolen it") and "Als Judy's iPod niet in haar tas zit, dan zal iemand die wel gestolen" ("If Judy's iPod is not in her bag, then someone will have stolen it").

Participants ($N = 30$) in the condition "moet wel" were asked to judge the acceptability of the conditionals containing "moet wel." Participants ($N = 27$) in the condition "zal wel" were asked to judge the acceptability of the conditionals containing "zal wel." The instructions were the same as the ones used in the acceptability condition of Douven and Verbrugge's experiment.

RESULTS

Comparisons with the condition investigating the probability were set up. We computed the mean per sentence (ten abductive sentences, ten inductive sentences, and ten deductive sentences) over the participants. We thus obtained a mean for each of the thirty sentences. Then we computed the correlations between the condition with marker ("zal wel" / "moet wel") and the mean probabilities per item obtained on the basis of Douven and Verbrugge's experiment.

For the thirty sentences in the experiment, probabilities as obtained in Douven and Verbrugge's experiment and acceptability of the sentences with "zal wel" were highly correlated: $N = 30$, Spearman $R = .837712$, $t(N - 2) = 8.116914$, $p < .0001$. For the thirty sentences in the experiment, probabilities as obtained in Douven and Verbrugge's experiment and acceptability of the sentences with "moet wel" were highly correlated: $N = 30$, Spearman $R = .855871$, $t(N - 2) = 8.756641$, $p < .0001$.

We next considered different types of conditionals. For the abductive sentences, probability and "moet wel" were highly correlated: $N = 10$, Spearman $R = .993902$, $t(N - 2) = 25.49522$, $p < .00001$. We obtained similar results for "zal wel": $N = 10$, Spearman $R = .936175$, $t(N - 2) = 7.532386$, $p = .0001$. For the inductive sentences, correlations did not reach significance level; for "moet wel" and "zal wel": $N = 10$, Spearman $R = .612121$, $t(N - 2) = 2.189453$, $p = .059972$. For the deductive sentences, probability was highly correlated with "moet wel": $N = 10$, Spearman $R = .814593$, $t(N - 2) = 3.972223$, $p < .01$. For probability and "zal wel," the correlation did not reach significance level: $N = 10$, Spearman $R = .613985$, $t(N - 2) = 2.200141$, $p = .058981$.

Comparison with the results from Douven and Verbrugge's experiment showed that, overall, the markers had little effect on the perceived acceptability of the conditionals as well as, correspondingly, on the correlation between acceptability and probability (see Table 1). However, splitting the results for the various types of conditionals was more revealing. It appeared that, while for II conditionals, adding either of our two markers had virtually no effect on the correlation between acceptability and probability, adding them to the AI conditionals did increase the said correlation, even to the extent of yielding a near-to-perfect correlation for the AI conditionals containing "moet wel" (for no marker, $R = .8997$; for "zal wel," $R = .936175$; and for "moet wel," $R = .993902$). For DI conditionals,

Table 1. Correlations with and without markers (* = marginally significant; for all other results $p < .01$)

	All	DI	II	AI
No marker	.851102	.818182	.620064*	.899700
"zal wel"	.837712	.613985	.612121*	.936175
"moet wel"	.855871	.814593	.612121*	.993902

adding "moet wel" had almost no effect, but adding "zal wel" led to a considerable decrease of the correlation between acceptability and probability.

4.2 Discussion

These results confirm that "moet wel" and "zal wel" are inferential markers, given that (i) inserting either of them in the AI conditionals has the effect of increasing the correlation between acceptability and probability, and (ii) inserting "zal wel" in the DI conditionals has the effect of decreasing that correlation. It is no surprise that inserting "moet wel" in the DI conditionals does not have a similar effect: like "must" in English, "moet wel" can also serve as an alethic modality, and may thus be naturally interpreted in a DI conditional as underlining the necessity of the inference.

5   Concluding Remarks

The typology proposed in Douven and Verbrugge [10] helps to explain why adding inferential markers to conditionals makes the systematic kind of difference that we found in our experiment. That is further support for the thought that this typology is of theoretical significance. The experimental work reported here is part of a larger project. Experiments concerning the English "must," "should," and "will" are currently being undertaken, and the results are to be compared to the results of the experiment reported above. A further avenue for future research concerns applications of these markers. For instance, if some markers can be used as a kind of litmus test for distinguishing between various types of conditionals, then testing for the effect that adding these markers to conditionals has may help us in classifying conditionals whose type is controversial. Perhaps the most important part of the project is to see whether the current typology of conditionals can serve to ground a new semantics and/or pragmatics of conditionals. Once it is recognized that various types of inferential connection may be involved, it becomes quite plausible to claim that at least many conditionals require for their acceptability the existence of an inferential link between antecedent and consequent. Whether this is then to be taken as a brute pragmatic fact, or whether it has a deeper explanation in terms of truth conditions, is the question we ultimately hope to answer. The investigations reported here are meant as a first step toward that answer.

References

1. Edgington, D.: On conditionals. Mind **104**(414) (1995) 235–329
2. Bennett, J.: A Philosophical Guide to Conditionals. Oxford University Press, Oxford (2003)
3. Dancygier, B.: Conditionals and Predictions: Time, Knowledge and Causation in Conditional Constructions. Cambridge University Press, Cambridge (1998)
4. Dancygier, B.: Classyfying conditionals: Form and function. English Language and Linguistics **7** (2003) 309–323
5. Dancygier, B., Sweetser, E.: Mental Spaces in Grammar. Conditional Constructions. Cambridge University Press, Cambridge (2005)
6. Declerck, R., Reed, S.: Conditionals: A Comprehensive Empirical Analysis. Mouton de Gruyter, Berlin/New York (2001)
7. Haegeman, L.: Conditional clauses: External and internal syntax. Mind & Language **18**(4) (2003) 317–339
8. Woods, M.: Conditionals. Oxford University Press, Oxford (2003)
9. Sanford, D.H.: If P, Then Q: Conditionals and the Foundations of Reasoning. Routledge, London (1989)
10. Douven, I., Verbrugge, S.: The Adams family. Cognition **117** (2010) 302–318
11. Adams, E.W.: The logic of conditionals. Inquiry **8** (1965) 166–197
12. Willett, T.: A cross-linguistic survey of the grammaticization of evidentiality. Studies in Language **12**(1) (1988) 51–97
13. Papafragou, A., Li, P., Choi, Y., Han, C.: Evidentiality in language and cognition. Cognition **103** (2007) 253–299
14. Karttunen, L.: "Possible" and "Must". In: Syntax and Semantics. Academic Press, New York (1972)
15. Groenendijk, J.A., Stokhof, M.J.: Modality and conversational information. Theoretical Linguistics **2** (1975) 61–112
16. Kratzer, A.: Modality. In von Stechow, A., Wunderlich, D., eds.: Semantics: An International Handbook of Contemporary Research. de Gruyter, Berlin (1991) 38–74
17. von Fintel, K., Gillies, A.: An opinionated guide to epistemic modality. Oxford Studies in Epistemology **2** (2007) 32–63
18. von Fintel, K., Gillies, A.S.: *Must* ... stay ... strong! Natural Language Semantics **18**(4) (2010) 351–383
19. Nuyts, J., Vonk, W.: Epistemic modality and focus in Dutch. Linguistics **37**(4) (1999) 699–737
20. Verbrugge, S.: Frequency and lexical marking of different types of conditionals by children and adults in an elicitation task. In: Proceedings of the EuroCogSci 07 Conference. (2007)
21. Dietz, R.: Epistemic modals and correct disagreement. In García-Carpintero, M., Kölbel, M., eds.: Relative Truth. Oxford University Press, Oxford (2008) 23–262
22. Anderson, L.B.: Evidentials, paths of change and mental maps: Typologically regular asymmetries. In Chage, W.L., Nichols, J., eds.: Evidentiality: The Linguistic Coding of Epistemology. Ablex, Norwood NJ (1986)
23. Papafragou, A.: Inference and word meaning: The case of modal auxiliaries. Lingua **105** (1998) 1–47