The grammar of discourse: The case of then *

María Biezma
University of Konstanz

Abstract This paper presents an analysis of then that appeals to discourse structure. I argue that then does not make any truth-conditional contribution but rather coordinates an anaphoric relation between consecutive discourse-moves. The meaning contribution of then is captured in terms of felicity conditions: then requires that two propositions enter into a ‘causal explanatory claim’-relation in which one (the antecedent) provides the ‘reasons’ for the other (the consequent). This proposal targets then within conditionals and across discourse and considers discourse structure as a key factor in explaining its behavior.

Keywords: Pragmatics, semantics, conditionals, causal explanations, discourse markers, then.

1 Desiderata

It is not news that lexical items do not necessarily contribute truth-conditional meaning. Non-truth conditional meaning contributions of lexical items have been said to include reference to the participants’ epistemic state (e.g., evidentials), the status of entities in the discourse (e.g., indefinites) or to contribute to other meaning dimensions (see for example Potts 2005 on the expressive dimension). In this paper I provide support for the view that lexical items also operate at the level of discourse structure (also a type of non-truth-conditional meaning) and that to account for their meaning we need to consider the grammar of an articulated discourse system.

It has already been claimed that the primary function of some lexical items is “information structural”. The case was made in Beaver & Clark 2008 for exclusives, whose main function is argued to be to comment on the current question under discussion (in the framework of Roberts 1996; Büring 2003). Biezma 2009, 2011 argued that discourse itself and the internal functioning of discourse is responsible for enriched meanings such as the “cornering effect”, and Biezma & Rawlins 2012a, 2014 argue that there is a pronominal system at the discourse-level and that discourse

*I would like to thank the audience of SALT 24 at NYU, specially to Scott AnderBois, Peter Klecha, Sarah Murray and William Starr. Special thanks also to Ana Arregui for discussion throughout the development of this project.

©2014 Biezma
pronouns are anaphoric with discourse moves. Similarly, for German, Rojas-Sponda 2014 claims that the lexical item überhaupt signals that the utterance addresses a previous discourse question.

In this paper I present further support for the importance of discourse structure in the construction of meaning by investigating non-temporal then. I argue that then is a discourse marker establishing an anaphoric relation with the previous discourse move. The remainder of the paper is organized as follows: in §2 I present the relevant data regarding then both in conditionals and across discourse; §3 reviews previous proposals; and in §4 I argue that then is a discourse marker that imposes felicity conditions: then establishes an anaphoric relation between the information gained from the previous discourse move and information gained from the subsequent move. I argue that then requires that antecedent and consequent establish a ‘causal explanatory claim’.

2 Data

2.1 Then in conditionals

To keep the discussion manageable, only the presence of then in canonical conditionals will be targeted (constructions consisting of an if-clause, the antecedent, and a main clause, the consequent). The literature often treats conditional antecedents (syntactically) as free relatives that adjoin to the matrix clause (see Bhatt & Pancheva 2006 for the syntax of conditionals), i.e., as correlatives, and in this tradition then is the correlative proform. The question arises as to why then is possible in some conditionals but not in others. One of the conclusions of this paper is that the (un)availability of then is not in itself an argument supporting different syntactic analysis of conditionals. I argue that when then is not available it is due to the fact that the meaning brought about by then is implausible.

Then can appear in hypothetical conditionals: conditional structures in which it is understood that the antecedent merely provides sufficient conditions for the truth of the consequent.

(1) If you finished your homework, then you can go play outside.

However, there are some conditionals in which then is not possible:2

1 I won’t discuss similar markers in other languages (e.g., Spanish entonces) for space reasons.
2 Examples (2a)-(2c) are from von Fintel (1994). Von Fintel (1994) also reports the following example with the corresponding judgement:

(i) If Smith were the Devil Incarnate, (# then) the Sheriff wants him.

However, speakers of English do not find this conditional good (with or without then). A way to amend the example would be as follows:

(ii) If Smith were the Devil Incarnate, (then) the Sheriff would want him.
The grammar of discourse: The case of then

(2) a. Even if Smith is dead, (#then) the Sheriff wants him.
b. If Smith is dead or alive, (#then) the Sheriff wants him.
c. Whether Smith is dead or alive, (#then) the Sheriff wants him.
d. If you are hungry, (#then) there is pizza in the fridge.

Besides the observation that then is not available in some conditionals, it is also observed in the literature (see Davis 1983) that the presence of then triggers a difference in meaning:

(3) a. If you open the refrigerator, it will not explode.
b. If you open the refrigerator, then it will not explode.

Davis (1983) claims that the utterance of (3a) conveys that your opening of the fridge won’t cause it to explode, whereas this is not conveyed by the utterance of (3b), whose interpretation is one in which your opening of the fridge will prevent it from exploding. As we will see, Davis’s claim is too strong and given appropriate contextual information (3a) can have the same interpretation as (3b) (see §4.5). But in any case, there is a clear contrast in (3).

The impossibility of then in some conditionals together with the meaning difference observed in (3) calls for an analysis of then in which it makes a meaning contribution beyond establishing an anaphoric relation.

2.2 Then across discourse

Then does not appear only in conditional constructions. We can also find then across discourse with a meaning similar to the meaning we find in conditionals:

(4) Child: I finished my homework.
    Dad: Well, then you can go play outside. [then-clause]

As in the case of conditionals, for then-clauses to be felicitous we need an ‘antecedent’ from which the utterance of the then-clause follows. I will keep referring loosely to what precedes the then-clause as the antecedent and the clause embedded under then as the consequent. Terminology will be made precise in §4.

Notice that the then-clause uttered in (4) is not a conditional structure in which the antecedent has been elided (and can be recovered discursively from the previous

Notice however that in this case, then is good. Iatridou 1994 also points out that only if conditionals do not accept then:

(iii) Only if it is sunny (# then) will I visit you.

Notice however that Only if it is sunny, will I then visit you is fine, hence indicating that the (iii) is bad possibly for independent syntactic reasons.
utterance). That is, the *then*-clause in (4) is not equivalent to the conditional in (1). To the utterance of the conditional in (1) the child can respond *Don’t you believe me?*, but this response is not possible after the utterance of the *then*-clause in (4). The protest is possible after the utterance of the conditional in (1) because the utterance of the conditional triggers the inference that it is not known that the antecedent is true. By uttering the conditional instead of just a declarative sentence the speaker leaves open the possibility that the addressee did not finish his/her homework (despite his claim). The protest is not possible in (4) because the utterance of the (plain) *then*-clause does not trigger the same inference.

Notice that the nature of the antecedent to *then*-clauses may differ, as may differ the nature of the embedded clauses. Regarding assertions, questions and imperatives, all combinations are possible but for one (I am leaving out the combination in which the clause type is the same in the examples below for space reasons):³

<table>
<thead>
<tr>
<th>(5)</th>
<th>Assertion - Imperative</th>
<th>(6)</th>
<th>Imperative - Assertion</th>
<th>(7)</th>
<th>Question - Assertion</th>
<th>(8)</th>
<th>Assertion - Question</th>
<th>(9)</th>
<th>Imperative - Question</th>
<th>(10)</th>
<th># Question - Imperative</th>
</tr>
</thead>
<tbody>
<tr>
<td>A:</td>
<td>I’m cold.</td>
<td>A:</td>
<td>He must confess! Cut off his fingers!</td>
<td>A:</td>
<td>What does “lambda” mean?</td>
<td>A:</td>
<td>He was cranky this morning.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B:</td>
<td>Then put on a sweater!</td>
<td>B:</td>
<td>Then you really are a sadist.</td>
<td>B:</td>
<td>Then you didn’t understand the lecture.</td>
<td>B:</td>
<td>Then, you told him?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The interpretations of the examples above are that you have to put on a sweater because you are cold (and want to be warmer); that you are a sadist because only sadists order such torture; that you didn’t understand the lecture because you are asking that very question; whether I should deduce that you told him because he was cranky; whether I should accept that the press was right because you are ordering such a butchery. An analysis of *then* has to derive all those readings and explain why (10) is infelicitous.

³ The case of exclamatives is discussed in footnote 10.
The grammar of discourse: The case of then

Notice that the parallel to Davis’s examples in (3) with plain then clauses have similar interpretations:

(11)  A: I’m gonna open the refrigerator.
      B: Sure!, it will not explode.
(12)  A: I’m gonna open the refrigerator.
      B: Sure!, then it will not explode.

The interpretation of (11) is one in which your opening of the refrigerator won’t make it explode. Notice however that in this case we lack the reading in which your opening of the refrigerator will prevent it from exploding (which given a very rich context could be obtained in Davis’s then-less-conditional example (3a)). The missing reading in (11) is the reading available in (12): your opening of the refrigerator will prevent it from exploding, just like in the conditional example (3b).

In light of the discussion above, it seems that the meaning of then is the same in conditionals and across discourse, and that differences in interpretation (such as the differences between (1) and (4) and between the possible interpretations for Davis’s example (3b) (and the parallel (12) just discussed) and (3a) need to be explained in terms of then being in different grammatical environments.

3  Previous proposals

In this paper I claim that then carries a meaning and is not just a proform. This approach is defended also in Iatridou 1994 and von Fintel 1994. Paying attention only to (canonical) conditionals, Iatridou and von Fintel make proposals that intuitively could be said to have a “causal-like” flavor. Iatridou (1994) proposes that the utterance of the conditional If p, then q is equivalent to the utterance of the conditional if p, q with the presupposition that not all ¬p worlds are q worlds. (This is not exactly the causation intuition since it leaves open the possibility of there being ¬p worlds that are q worlds.) In Iatridou’s account we can derive the impossibility of then in all conditionals in which the antecedent exhaustifies logical space since the inference forced by the presence of then is not compatible with the conditional itself: when the antecedent exhaustifies logical space, the conditional claim is that the consequent is true no-matter what and there are no worlds in which the consequent is not true. This is similar to the case of biscuit conditionals (BCs). Iatridou 1994 considers them conditional assertions and, since the consequent is true

4 For Lewis causation is a relation between events: e₁ causes e₂ iff had e₁ not occurred, e₂ would have not occurred (Lewis 1986b).
5 Iatridou 1994 calls this inference “presupposition”, but she warns the reader that she is using the term loosely.
6 Any strengthening of the inference would have to be made pragmatically.
no matter what, it is also impossible for there to be worlds in which the consequent is not true, hence clashing with the inference triggered by then.  

Von Fintel 1994 builds on Iatridou’s work and proposes that the presence of then triggers the conventional implicature that only the antecedent-worlds are consequent-worlds with the requirement that there be alternatives to the antecedent. In this proposal the presence of then conveys that if the antecedent were not true, the consequent would not be true. Like Iatridou’s proposal, von Fintel’s proposal explains the impossibility of then in conditionals in which the antecedent exhaustifies logical space and in BCs, in which the consequent is true regardless of the antecedent. These proposals are not concerned with then-clauses across discourse. However, regardless of how they could be amended to cover those cases, they face challenges within the conditional empirical domain alone. Consider the following conditionals:

(13) I’m certainly taking a job, but so far all the opportunities require me to start working at 7:00 a.m. Yes, as incredible as it sounds, if I take the Taco Bell job, then I also start at 7:00 a.m.

(14) If Jim had asked Jack for help, then there would (have to) have been no quarrel yesterday.

According to Iatridou 1994, the presence of then triggers the presupposition that there are alternatives to the antecedent in which the consequent is not true. However, the speaker uttering (13) does not assume such a thing and yet can include then in the utterance. The example in (14) is a backtracker conditional: a conditional in which the consequent temporally precedes the antecedent (see Lewis 1979 for discussion). The speaker uttering (14) is not trying to convey that only worlds in which Jim asked Jack for help are worlds in which there was no quarrel yesterday. Contrary to the prediction made by the proposal in von Fintel (1994), the interpretation of (14) is that the relevant worlds in which Jack asked John for help are also worlds in which there was no previous quarrel, but not that only in worlds in which Jim asks Jack for help was there no quarrel yesterday.

Even though the previous proposals fail to make the right predictions in some cases, there is something quite appealing about the flavor of causation somehow implicit in those analyses. It seems to respond to the intuition behind the ‘because’ paraphrase in examples (5)-(9). In the following sections I argue that then indicates that antecedent and consequent together convey a ‘causal explanatory claim’.

4 Then: An analysis based on causal explanations

4.1 Then is a discourse marker

Looking at the paraphrases of the examples discussed above, the meaning contribution of then does not appear to be backgrounded, i.e., entailed by discourse
The grammar of discourse: The case of *then*

participants’ shared knowledge at the time of utterance. Let us conclude therefore that *then* is not a presupposition trigger. There is also a problem with placing *then* in the conventional implicature dimension, since its contribution is not speaker-oriented (Potts 2005) (i.e., *then*-clauses do not present commitments expressed by the speaker). At the same time, the contribution of *then* is not truth-conditional. One way to see this is to apply Zimmermann’s (2008) diagnosis. Part of the empirical evidence concerns the behavior of *then* with sentence-types other than declaratives. If the meaning of *then* were truth-conditional, its presence would affect the truth-conditional meaning of the sentences it appears in, but it doesn’t:

(15) Deciding what dress to wear for her first date.
   
   A: I don’t like the red dresses.
   
   B: Then, do you want to wear the blue dress or the black dress?

Assuming that the truth-conditional meaning of a question is the set of possible answers (Hamblin 1958, 1973; Karttunen 1977), and that the possible answers for an alternative question are the alternatives spelled out in the individual disjuncts (Belnap & Steel 1976; Karttunen & Peters 1976; Higginbotham 1991; Bartels 1999; Biezma & Rawlins 2012b), the addition of *then* does not modify the meaning of the question embedded under it (the answers are the same). The same can be seen with imperatives:

(16) A: I’m cold.
   
   B: Then, put on a sweater!

Whatever the semantics of imperatives, whether modal or not (see Schwager 2006; Portner 2004, 2007; Eckardt 2011; Condoravdi & Lauer 2012 for different views), there is no difference in the interpretation of the imperative with or without *then*.

What is *then* then? Looking at the examples above, *then* links the embedded clause to the previous discourse. In (15) *then* indicates that the question is motivated/follows from information gained in the previous discourse move; in (16), *then* indicates that the order follows/is motivated by information gained upon A’s previous discourse move. The function of *then* can be characterized as some kind of discourse function. However, it is not a discourse particle, if we take discourse particles to establish “a link between the proposition expressed by an utterance and the knowledge and belief systems of the discourse participants” (Zimmermann 2011: 2013). Looking at (16) it is clear that *then* does not add anything regarding the degree of certainty towards B’s suggestion to A to put on a sweater.

The function of *then* seems to agree with that of a discourse marker: it signals “the relationship of the basic message to the foregoing discourse” (Fraser 1990). This is what we see in the examples above. But notice that the propositions targeted
by *then* are not necessarily the content proposition of the previous utterance and the content proposition of the clause embedded under *then*. Let’s take (16) for instance. B’s utterance can be paraphrased by ‘I’m suggesting that you put on a sweater because you want to be warmer’ (and I clearly consider that putting on a sweater is the best way for you to get warmer). That A wants to be warmer is an inference triggered by A’s claim that s/he is cold. With B’s utterance we learn (in addition to the conventional meaning of the imperative) that B suggests that A puts on a sweater. The claim that A is cold and the inference that s/he wants to be warmer are crucial to evaluate the *then*-claim in relation with the consequent (see §4.4).

Let us consider the example in (17):

(17)  
A: Are our kids’ soccer uniforms red or blue?  
B: Oh!, then they finally joined the team. I see...

The example in (17) shows that what triggers B’s utterance is not the propositional content that the kids’ soccer uniforms are red or that the kids’ soccer uniform is blue (the conventional meaning of the question, Hamblin 1958, 1973; Karttunen 1977). B’s utterance is triggered by the information that the kids have a soccer uniform, which is presupposed in A’s utterance. This example also shows that it is not the relation between the conventional meaning of the clauses that is marked by *then* as being ‘causal-like’ in nature: it is not that the kids having a soccer uniform “causes” them to join the team (more specifically, it is not the case that had the uniform not been red (or blue) the kids would have not joined the team). B’s utterance in (17) could be rather paraphrased as ‘I am justified in claiming that the kids joined the soccer team because I just learned that they have soccer uniforms for this year’s league (and according to my knowledge of how things work, they have uniforms only when they join the team)’.

In sum, *then* is a discourse marker that signals that the utterance of the embedded clause is in some sense motivated by the preceding discourse move. This will be made more precise below.

### 4.2 Causal explanatory claims

As we have noticed in the backtracking-like examples, the presence of *then* does not convey that the antecedent \( p \) is a “cause” for the consequent \( q \). However, there is something quite intuitive about linking the presence of *then* with a causal-like

---

7 Temporal relations here are similar to those found in backtracking examples. In (17) the temporal anchoring of the proposition embedded under *then*, that the kids joined the team, precedes the acquisition of the uniform. Notice that Lewis (1979) already warns us about trying to exemplify causation in these circumstances since it is usually the case that events causally depend on earlier events and not the other way around.
The grammar of discourse: The case of *then*

relation (as evidenced by our paraphrases with *because*). The thesis I defend in this paper is that the presence of *then* indicates that the antecedent is a causal explanation of the consequent.

The difference between causation and causal explanation was defended by Davidson 1967 on the grounds of a difference in logical form. However, Beebee 2004 provides additional arguments (independent of logical form) supporting the view that causal explanation is not necessarily a report of causation. Beebee (2004) adopts Lewis’s (1986b) concept of explanation according to which “to explain an event is to provide some information about its causal history” (Lewis 1986b) and she considers that the causal history of an event is a “relational structure”. There are ways of providing information about the history of an event beyond providing its causes: “explanatory information comes in many shapes and sizes”. For example, “an explainer well might be unable to specify fully any particular event in the history, but might be in a position to make existential statements. He might say, for instance, that the history includes and event of such-and-such-kinds [...] in other words, he might make an existential statement to he effect that the history includes a pattern of events of a certain sort” (Lewis 1986a: 240).

The proposal I present below argues that the presence of *then* indicates that there is a dependency between antecedent and consequent: its presence indicates that the combination of antecedent and consequent establishes a causal explanatory claim. We are going to build on the idea of a causal explanatory relation even though strictly speaking we are dealing with propositions and not events. We are targeting the intuition of what it means to provide justification for the truth of the proposition.

Given this proposal, *then* makes a conventional contribution distinct from that of conditionals. In the case of conditionals, and adopting a Lewis/Stalnaker analysis, (roughly) a conditional (without *then*) if $p$, $q$ simply conveys that in the most similar worlds to the actual world (best worlds) in which $p$ is true, $q$ is also true. If we think about this in a dynamic perspective, the conditional claim only provides a description of the worlds in the context set. The presence of *then* adds to this conditional claim that it is because of $p$ that $q$ (for specific examples see §4.4).

In an account in which conditionals containing *then* are causal explanatory claims, the antecedent provides ‘causal information’ about the consequent. This explanation is grounded in laws. In such an account it is predicted that conditionals in which the antecedent is orthogonal to the consequent, *then* is not possible: the

---

8 Beebee 2004 adopts a “network model” (see also Steward 1997), based on Lewis’s work, in which the causal relation is taken to be a directed segment and the causal relate as nodes. This model takes history to be a vast causal network that can be represented as an inverted tree. See Schaffer 2008 for an overview.

9 The relevant notion of orthogonality is the one in Lewis 1988: “Two subject matters $M_1$ and $M_2$ are orthogonal iff, roughly, any way for $M_1$ to be is compatible with any way for $M_2$ to be. If we think of
antecedent will not have any causal link to the consequent.

(2) a. Even if Smith is dead, (#then) the Sheriff wants him.
   b. If Smith is dead or alive, (#then) the Sheriff wants him.
   c. Whether Smith is dead or alive, (#then) the Sheriff wants him.

When the antecedent exhaustifies logical space, it does not provide a “causal explanation” for the consequent, rendering the presence of then infelicitous.

4.3 Implementation

The analysis of then will be implemented by appealing to discourse, which will provide the flexibility to deal with then across discourse and in conditionals. The information gained in discourse moves will be modelled using Gunlogson’s (2001) commitment slates, which allow us to track the information available to individual discourse participants.

Let us consider that a discourse move $M_i$ is the utterance of a sentence structure syntactically headed by a force operator ([A[S]]; [Q[S]] or [Imp[S]]). Following Gunlogson 2001, 2008 let us define a commitment slate as the set of propositions taken to be true by a discourse participant. Every discourse participant has his/her own commitment slate and these are bound to being consistent (the commitment slates do not contain conflicting propositions).

\[(18) \quad \text{Commitment Slate} \cos_{B,M_n} = \{ p : B \text{ commits to } p \text{ after } M_n \}\]

Notice that commitments made by participants do not need to correspond to explicit utterances. Following Gunlogson (2008), an agent also implicitly commits to entailments, presuppositions, and (non-cancelled) implicatures. In general, “anything that an agent publicly treats as true for the purposes of the discourse will qualify as a commitment.” Associated to the commitment slate we find the commitment state, the set of worlds in the intersection of the propositions in the participant’s $\cos$. With the aid of the notion of commitment slate, let us now define information gain:

\[(19) \quad \text{Information gain:} \quad I_{B,M_i} = \{ p : p \in \cos_{B,M_i} \land p \notin \cos_{B,M_{i-1}} \}, \text{ where } M_{i-1} \text{ is the move immediately preceding } M_i \text{ and } \cos_{B,M_i} \text{ is B’s commitment slate after } M_i.\]

subject matters as equivalence relations, orthogonality means that for any worlds $w$ and $v$ there is a world $u$ such that $M_1(u, w)$ and $M_2(u, v)$. If we think of subject matters as partitions, orthogonality means that $M_1$ and $M_2$ cut across each other: each cell of $M_1$ intersects each cell of $M_2$. Subject matters are connected iff they are not orthogonal.” (Lewis 1988: 167)
After a discourse move has taken place, we usually learn lots of things beyond the propositional content of the utterance (see Stalnaker 2002). For example, after A’s utterance in (17) B knows that A wonders what color the kids’ uniform is (A uttered a question on that matter), but also that the kids have a soccer uniform (use of the possessive) and that the soccer uniforms can only be red or blue (by virtue of the presupposition triggered by alternative questions). All this information was not necessarily known by B before A’s utterance (indeed in the context of utterance in the example, none of this seems to be known by B before A’s utterance), and in that case, all that information is now part of the information gain after A’s utterance.

With these notions in place we can now make a proposal for the meaning of then. Given our conclusion that then is a discourse marker, even though it does not affect the truth-conditions of the utterance, it imposes felicity conditions between antecedent and consequent. We will build on the intuition that then signals a ‘causal explanatory claim’ between antecedent and consequent. As noted earlier, causal explanatory claims depend on patterns of events of a certain sort. Let us assume that relevant regularities are law-like relations. Following Lewis (1994), laws are generalizations about collections of truths. Which collection of truths is selected to draw the generalization from is the result of a competition (between different collection of truths) with respect to strength (a truth is stronger if it tells us more about the world), simplicity (or systematicity) and fit (as summarized in Weatherson 2010, the collection of laws that entails that what does happen has at earlier times a higher chance of happening has better fit than a collection that entails that it has a lower chance of happening). In order for there to be a law-like dependence between two propositions, the truth of one has to at least increase the probability of the truth of the other. Laws are contingent (Lewis 1986c) and underlie causal explanations (Lewis 1994), which is where we want to arrive in our investigation of then.

With these ingredients in hand, we can now propose the denotation in (20). I follow Collins 1989; Iatridou 1991; Iatridou & Kroch 1993; von Fintel 1994 in assuming that then is placed in SpecCP.

(20) Let A be the speaker, g an assignment function, P an ordering source formed by regularities/laws and MB a modal base (Kratzer-style conversational backgrounds), $w_@$ the utterance world, and $\text{MAX}_{P(w_@)}(X)$ the $P$-best worlds in a set of worlds $X$,

$$[[\text{cp} \text{ Then } [M_{i+1}]]]g(w_@) = [M_{i+1}](w_@$), \text{ defined only if}$$

$$\forall w \in \text{MAX}_{P(w_@)}(\cap \text{MB}(w_@)) \cap g(i), g(i+1)(w) = 1$$

Where for any discourse move $M_j$, $g(j) \in I_{A,M_j}$.

According to the definition above, then does not change the truth-conditions of the embedded clause, but imposes felicity conditions on the utterance of such clause. According to (20), then is anaphoric with the previous discourse move. It coordinates
a (salient) piece of information (identified by the assignment function) gained from
the utterance of the preceding move —the antecedent (either the propositional content
of the preceding move or any inference triggered by it)— and a piece information
gained by the utterance of the then-clause (also identified by the assignment function)
—the consequent. Then requires the two to be in a law-like modal relation. The
assignment function needs to find a proposition in the set of information gain after the
previous discourse that can be in a conditional relation with information gained after
the utterance of the clause embedded under then, taking into account an ordering
source based on laws/regularities (which could be epistemic). Upon the utterance of
the then-clause, participants reconstruct (part of) what the speaker learned after the
previous move. The constraint on the antecedent is that it stands in a law-dependence
with the consequent, which is retrieved from the information gained after the then
clause is uttered. In sum, then drives us to identify two propositions that stand in an
appropriate law-like relation.

A welcome prediction from the definition in (20) is that for then to be felicitously
included in an utterance, there must be a previous discourse move from which the
appropriate content can be recovered. This cashes out the anaphoric nature of then.

4.4 Case studies

Then in conditionals: I adopt a Lewis/Stalnaker/Kratzer analysis for indicative
conditionals. Under this account if-clauses are restrictors of a modal operator in the
main clause. Following Kratzer (1981, 1986), when there is no overt operator, there
must be a covert necessity (epistemic) one. Under this account, a conditional If p, q
states roughly that given the speaker’s knowledge, in the relevant antecedent-worlds,
the consequent is true. Following Heim (1992), the dynamic implementation of this
proposal is as follows:

\[(21) \quad Sim_w(p) = \{w \in W : w' \in p \text{ and } w' \text{ resembles } w \text{ no less than any other} \]
\[\text{world in } p} \]

\[(22) \quad \text{For any context set } c, \text{ clause } \phi, \text{ and clause } \psi
c + \{[[if \phi] \psi] = \{w \in c : Sim_w(c + \phi) + \psi = Sim_w(c + \phi) \}

According to (22), upon uttering a conditional we update the main context set with
the proposition denoted by the antecedent clause. In the next step the claim is that in
the resulting (temporary) context the consequent is true. At the end of the conditional
update, all the worlds ‘lost’ in the temporary context (those in which the consequent
is not true) are also lost in the main context. According to this definition, the context
update in conditionals is a two-step process: the context is updated first with the
antecedent and then with the consequent.
As stated in (22), conditionals simply describe what the worlds in the context set are like. The semantics of conditionals in itself does not require that there be a correlation between antecedent and consequent. A conditional can be true even if there isn’t a law-like dependence between the antecedent and consequent proposition. For example, in the case of the unconditional Whether it rains or not, Jones wears his hat, in which the antecedent exhaustifies logical space, we only learn that in all the worlds in the context set Jones wears his hat (see Rawlins 2013). There is no law-like link between the weather and the wearing of the hat by Jones (the antecedent is true in all possible worlds—its truth does not raise the probability of the consequent being true). Then is similar to conditionals in the sense that it coordinates a relation between two propositions, but it requires a more particular relation between the “antecedent” and ‘consequent’. Given (20), there must be a salient law-like link between the two. In cases in which this does not hold, as in Whether it rains or not, Jones wears his hat, then is not possible.

Let us examine the following epistemic example:

(23) If there is light in John’s room, then he is home.

We can map each context update in conditionals to the two discourse moves:

(24) $M_1$ There is light in John’s room.
    $M_2$ John is home.

Upon the utterance of the if-clause we learn/assume (temporarily) that there is light in John’s room, and after the utterance of the consequent we learn that he is home.

(25) a. $I_{B, M_1} = \{\text{There is light in John’s room}\}$
    b. $I_{B, M_2} = \{\text{John is home}\}$

The conditional (without then) has as truth-conditions that given what the speaker knows, in the relevant antecedent worlds, the consequent is true. This could be an accidentally true description of the worlds in the domain or it could be an inference based on epistemic generalizations. The felicity conditions imposed by then ensure a law-like link between antecedent and consequent. Given the regularities of John’s behavior (he always turns off the lights when he is not home), the lights being on allows us to conclude that he is home. If the conditional itself is interpreted in relation to this kind of generalizations, the presence of then would be superfluous.

In the presence of then, we understand that it is because the antecedent is true that we can claim the consequent is true.

Then across discourse: In the case of then across discourse, the felicity conditions impose a modal relation between the antecedent and the consequent:

(26) A: I’m cold.
Biezma

\[ M_2 \]

**B**: Then put on a sweater!

By uttering the *then*-clause, B states that given \( M_1 \), s/he committed to the fact that A is cold and that A wants to be warmer.

(27) \[ I_{B, M_1} = \{ \text{A is cold; A wants to be warmer} \} \]

By uttering the *then*-clause \( (M_2) \), in addition to the conventional meaning of the imperative, B indicates that the best worlds according to what s/he knows are worlds in which A puts on a sweater (paraphrased below using *should*):

(28) \[ I_{B, M_2} = \{ \text{B should put on a sweater} \} \]

In the example above, *then* appeals to epistemic regularities. Given what I know, since you are cold, it is true that you should put on a sweater. The evaluation of the *should*-statement is made taking into account your wish to be warmer. Given what I know, it is true that in the best possibilities (in which you are warmer), you put on a sweater (see Arregui 2010; Kratzer 2012 regarding deontic modals embedded under epistemic operators).

### 4.5 Predictions

**The antecedents and consequents of then**: The first prediction is that *then*-clauses cannot be uttered out of the blue, since they wouldn’t have an antecedent and hence *then* would not find what it needs for its interpretation. Given speakers’ intuitions, this prediction is borne out. The other prediction refers to the kind of clauses that are good antecedents and consequents. Since *then* takes its antecedent and consequent from the commitment set, when a question is asked the only thing that makes its way to the commitment set (besides possible presuppositions triggered within it or other inferences and entailments) is that the speaker asked a question. The same happens when the speaker utters an imperative, what makes its way to the commitment set is that the speaker ordered/suggested something (that the speaker considers the worlds in which X is true to be the best). Hence, the antecedent or consequent of *then* in these cases is about the fact that X was asked or that X was suggested.

The question is why (10) is not possible.\(^\text{10}\)

(10) # Question - Imperative

**A**: What does lambda mean?

\(^{10}\) Exclamatives (not to be confused with sentence exclamations), don’t seem to be possible in *then*-clauses. One possible explanation could be that exclamatives have a discourse contribution different from that of assertion. According to Rett 2011, the exclamative *How tall John is!* is a property of degrees. This, I take is the key to explaining the impossibility of *then*-exclamative clauses: the only thing that makes its way to the information gain is that the speaker does not expect a degree d of X to be achieved, which does not make for a good consequent.
The grammar of discourse: The case of \textit{then}

\textbf{B:} # Then study today’s lesson! You will find out, hopefully!
\textbf{B:} √ Study today’s lesson! You will find out, hopefully!

The utterance of the imperative without \textit{then} is felicitous and is interpreted as a refusal to provide an answer. To explain the infelicity in the \textit{then}-case in (10) we need to take into account that to ask a question is to demand an answer (the speaker wants the addressee to provide an answer), i.e., in this sense questions are like commands (Lewis 1969). Taking into account the felicity conditions imposed by \textit{then}, if the utterance were felicitous we would be accepting that the best possible worlds (according to some laws) in which the speaker demands to obtain an answer are worlds in which the addressee refuses to provide one! This is not an accepted interpretation.

\textbf{Davis’s (1983) examples:} Recall the contrast between conditionals with and without \textit{then} from Davis 1983:

\begin{enumerate}
\item If you open the refrigerator, it will not explode.
\item If you open the refrigerator, then it will not explode.
\end{enumerate}

We saw that the preferred interpretation for (3a) is one in which your opening of the fridge won’t cause it to explode. However, given the right contextual conditions (3a) could also have the interpretation that your opening of the fridge will prevent it from exploding. Imagine that the fridge is making a strange noise, has a very hot surface and is in danger of exploding. You phone John, who knows a lot about fridges and how they work, and you explain the situation to him. The utterance of (3a) in that context is interpreted as (3b) would be: your opening of the fridge will prevent it from exploding. However, as we saw, (3b) only has this last reading.

The current account makes the right predictions regarding these examples. That we can have two readings for (3a) is due to the two possible scopes of negation: \textit{if }\neg q \textit{ and }\neg (\textit{if } p, q). Negation can scope over the consequent (narrow scope), in which case the reading is one in which your opening of the refrigerator will prevent it from exploding, or over the entire conditional (wide scope), in which case the reading is one in which your opening of the refrigerator won’t cause it to explode. Even though English speakers prefer the wide scope reading for (3a) (probably for common sense reasons), the narrow-scope reading is available, as we have seen. However, the wide scope reading is not available in (3b). Why? One could think that \textit{then} blocks movement preventing negation from taking wide scope. However, this explanation is hard to maintain:

\begin{enumerate}
\item If his mother is angry, then every boy is upset.
\end{enumerate}

Arguably, (29) has a reading in which the pronoun in the antecedent is bound by the quantifier in the consequent. Under standard approaches, this means that the
quantifier phrase takes wide scope over *then*, which wouldn’t be a barrier in this case. So, saying that *then* blocks movement would only apply to negation.

A closer inspection of the data presents another explanation. It seems that when the context allows us to reconstruct a dependence between the antecedent and the consequent with negation we prefer the narrow-scope reading (see (31) below). However, if the reading resulting from narrow scope is implausible because we do not recover any dependence between antecedent and consequent,\(^1\) we can always fall back on a wide scope reading. After all, a wide-scope reading only states that there is no correlation between \(p\) and \(q\) (\(\neg [if\ p,\ q]\)). As expected, when it is not clear whether we can establish a dependence between antecedent and consequent we have perfectly ambiguous conditionals, (30).

(30) If McDonald’s buys Wendy’s, Wendy’s employees won’t lose their jobs.

The utterance of (30), without contextual information, is ambiguous between McDonald’s purchase of Wendy’s preventing any job loss (\(if\ p,\ \neg q\)) and the statement that it is not the case that there is any correlation between McDonald’s purchase of Wendy’s and the loss of jobs (\(\neg [if\ p,\ q]\)). Notice that the difference between (30) and (3a) is independence: when independence is obvious, we obtain a wide-scope reading for negation, when we do not know, it’s ambiguous. When it is (contextually) obvious that there is a link between antecedent and consequent we obtain a narrow-scope reading even in the absence of *then*:

(31) Wendy’s is on the verge of bankruptcy and they will have to let many employees go, but McDonald’s need to expand their market and they will need lots of workers. If McDonald’s buys Wendy’s, Wendy’s employees won’t lose their jobs.

Claiming that when the conditional is independent we obtain a wide-scope reading for negation amounts to saying that the preferred reading is one in which negation has narrow scope. Why would we prefer a narrow-scope reading? Speakers prefer to interpret conditionals as establishing a dependence between antecedent and consequent,\(^1\) —that would be the obvious reason for the speaker to put them together! Indeed, that is why the presence of *then* seems superfluous many times. Having wide-scope negation simply denies that there is any kind of dependence, hence being disfavoured. Notice also that if \(p,\ \neg q\) is more informative than \(\neg (if\ p,\ q)\) since the former entails the later. Given that stronger claims (more informative statements) are preferred, the narrow scope reading is preferred also for this reason. This explanation

\(^{11}\) Antecedent and consequent are *independent*. Notice that this is not what Bennett 2003 has in mind for what he dubs *independent* conditionals: “[In independent conditionals] one can get the consequent from the antecedent without input from any matters of particular fact” (Bennett 2003: 16).

\(^{12}\) Indeed, if possible, speakers tend to pragmatically strengthen the conditional claim interpreting them not just as sufficient but necessary conditions (conditional perfection, see Geis & Zwicky 1971 a.o.).
for the relative scope of negation allows us to explain the contrast in (3). Because 
prima facie there is no correlation between opening a fridge and a fridge-explosion 
given what we standardly know about fridges, the preferred reading for (3a) is one 
in which negation has wide scope (the claim is that there is no correlation between 
the truth of the antecedent and the consequent), but contextual information can make 
prominent the narrow-scope reading.

The question now becomes why in (3b) the reading with wide-scope negation is 
not available. Recall that when negation takes wide-scope over the conditional (\(\neg[if 
p, q]\)) the meaning is basically that there is no correlation between antecedent and 
consequent. The result of interpreting a conditional containing then with wide-scope 
interpretation is then a contradiction. Truth-conditionally, \(\neg(if \ p, then \ q)\) states that 
the truth of q is independent from the truth of p. However, the felicity conditions 
imposed by then require that p and q be dependent (there is a law-like dependence).

**BCs and bactrackers:** The inclusion of then in a conditional prevents a BC reading. 
The inclusion of then imposes a particular type of modal relation between the 
antecedent and the consequent. But a conditional is a BC when such a relation is not 
possible, and we interpret that the consequent is true regardless of the consequent 
(in BCs we infer that antecedent and consequent are orthogonal). The presence of 
then makes impossible such interpretation and hence when then is present the BC 
reading is not available.

In the case of backtrackers (If Jim had asked Jack for help, there would have 
to have been no quarrel yesterday), (roughly) the presence of then is possible 
because the antecedent provides information about what makes the consequent more 
probable. More precisely, the account proposed in this paper predicts the availability 
of then since the felicity conditions simply impose that given what we know, the 
best worlds according to Jim’s regular behavior similar to the actual world in which 
he asks Jack for help are worlds in which they did not quarrel.

**When the consequent presupposes the antecedent:** Iatridou 1994 argues that then 
is not allowed when the consequent presupposes the antecedent:

\[(32) \text{ If Mary bakes a cake, (#then) she gives some slice of it to John.} \]

The judgement reported above with respect to then may be too strong for English 
speakers. Although some speakers find (32) degraded, others do not find a difference 
between the presence/absence of then. But even those who find (32) degraded, report 
improvement if the future marking is inserted as in (33), which is perfectly fine.

\[(33) \text{ If Mary bakes a cake, then she will give some slice of it to John.} \]

This suggests that there is a different explanation for (32) independent from 
preparation. It may be that the reason that some speakers find (32) degraded is 
that then may be read as establishing a rather redundant causal explanation: that
John gets a slice of the cake that Mary baked depends on Mary baking the cake. While the same could hold of (33), there is a difference: in the case of (33) we do not know whether Mary will bake a cake, in the case of (32), however, we do know that Mary baked a cake (at least once). Beyond establishing a rather trivial causal explanation, the utterance of (33) provides information about the epistemic state of the speaker (we don’t know whether Mary will bake a cake). The utterance of (32) only establishes a trivial causal explanation, and under this reading the presence of then is thus degraded.

**Modal subordination:** Notice the contrast between the following examples: 13

(34) a. I wish the bus were on time. ?(Then) I’d get home before midnight.
    b. I hope the bus comes on time. ??(Then) I’ll get home before midnight.
    c. You can’t go outside! (#Then) You’ll freeze.

That speakers prefer then in (34a) and (34b) is expected. The presence of then helps the addressee anticipate that there is a modal connection between information gained after the utterance of the first clause and information gained by the utterance of the second clause. In the absence of then the speaker has to work out why the two clauses were uttered consecutively and pragmatically supply a posteriori the connection that then would provide conventionally in an immediate manner. The contrast between (34a) and (34b), the latter being more degraded, is explained by the fact that the use of would in (34a) announces that there is an implicit antecedent (see Kasper 1992) and we can anticipate the need to recover it in the adjacent linguistic context, establishing a modal relation even without the presence of then.

The impossibility of (34c) is derived from the fact that the felicity conditions require us to understand that the best worlds (according to some laws) in which you are not allowed to go outside are worlds in which you end up freezing (roughly, that you will end up freezing because you are not allowed to go outside). However, the intended message is not quite that, but rather the opposite.

5 Conclusion

In this paper I have presented arguments supporting the idea that there are lexical items whose meaning contribution requires taking into account discourse structure. In particular, I have argued that then establishes an anaphoric relation between two consecutive discourse moves. Then imposes a modal felicity condition: that two propositions within the information gain after each move stand in a law-like relation. The presence of then requires that the sequence of moves establish a causal explanatory claim.

13 Thanks to Peter Klecha for presenting me the puzzle in (34).
The grammar of discourse: The case of then

References


Collins, C. 1989. Extraction from conditionals. Ms. MIT.


von Fintel, Kai. 1994. Restrictions on quantifier domains: University of Mas-


The grammar of discourse: The case of then

455–476.


María Biezma  
University of Konstanz  
Postfach 185  
Universitätstraße 10  
78457 Konstanz  
maria.biezma@uni-konstanz.de