Construction Grammar vs. Lexical Grammar: 
A case study of the modal load in *if*-conditionals

Costas Gabrielatos  
Lancaster University

University of Murcia, 1-2 December 2011
Motivation

Corpus based examinations of the *modal load* (i.e. extent of modal marking) in *if*-conditionals in the written BNC (Gabrielatos 2007, 2010) have revealed that they have a significantly higher modal load than:

- average
- concessive conditionals with *even if* and *whether*,
- indirect interrogatives with *if* and *whether*,
- non-conditional constructions with *when* and *whenever*
- conditionals with other subordinators (*assuming, in case, on condition, provided, supposing, unless*).

Is this due to ...

- the semantic preference of the lexical item *if*? (LG)
- the semantic make-up of *if*-conditional constructions? (CxG)
Why the particular theories?

• Both take into account ...
  … meaning (semantic **and** pragmatic)
  … lexical **and** grammatical elements

• Main difference ...
  … LG gives clear prominence to lexis over grammar
  … CxG accounts for both in a balanced way
    -- in fact, it posits no distinction.
Data: random samples

Source: written BNC; approx 1000 s-units each.

- **S-units**
  - Estimation of the average frequency of modal marking in written British English (baseline);

- **Non-conditional constructions, taken collectively;**

- **Conditional constructions with** assuming, if, in case, provided, supposing, unless

- **Conditional-concessive constructions with** even if and whether;

- **Indirect interrogative (non-conditional) constructions with** if and whether;

- **Constructions with** when and whenever (used as conjunctions)
  - They have been presented as synonymous with unmodalised if conditionals (e.g. Athanasiadou & Dirven, 1996: 617, 1997: 62; Palmer, 1990: 174-175).
Data: random samples

Source: written BNC; approx 1000 s-units each.

- S-units
  - Estimation of the average frequency of modal marking in written British English (baseline);
- Non-conditional constructions, taken collectively;
- Conditional constructions with assuming, if, in case, provided, supposing, unless;
- Conditional-concessive constructions with even if and whether;
- Indirect interrogative (non-conditional) constructions with if and whether;
- Constructions with when and whenever (used as conjunctions)
  - They have been presented as synonymous with unmodalised if conditionals (e.g. Athanasiadou & Dirven, 1996: 617, 1997: 62; Palmer, 1990: 174-175).

I am grateful to Stefan Evert (University of Osnabrück) and Neil Millar (University of Birmingham) for help with the regular expressions.
Modal Load
The interaction of two complementary metrics

Modal Density
Modalisation Spread
# Modal Density

<table>
<thead>
<tr>
<th>Definition</th>
<th>Average number of modal markings per clause.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expression</td>
<td>Number of modal markings per 100 clauses. (%)</td>
</tr>
<tr>
<td>Utility</td>
<td>Helps comparisons between samples by normalising for the complexity of the constructions in each.</td>
</tr>
</tbody>
</table>

(Gabrielatos, 2008, 2010)

**Lexical Density:**

- The average number of content words per clause (Halliday, 2004: 654-655).
- The percentage of the tokens in a text that are content words (Ure, 1971).
Modal density may not be enough

• A high MD may be the result of a number of heavily modalised constructions in the sample.
  
  – If you live in the Wallingford area and have a railway interest perhaps you might like to join this enthusiastic group and give them a few hours of your time. [CJ7 109]

• In such a case, a sample might show a high MD (relative to another sample) despite a large proportion of constructions in it being modally unmarked.
## Modalisation Spread

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Definition</strong></td>
<td>Proportion of constructions that carry at least one modal marking.</td>
</tr>
<tr>
<td><strong>Expression</strong></td>
<td>Proportion (%) of modalised constructions.</td>
</tr>
<tr>
<td><strong>Utility</strong></td>
<td>Corrects for heavily modalised constructions in the sample.</td>
</tr>
</tbody>
</table>

**Spread:**

- The proportion of corpus speakers who use a particular language item (Gabrielatos & Torgersen, 2009; Gabrielatos et al., 2010).
Why don’t we just calculate modal markings per X number of words?
## Words vs. opportunities

<table>
<thead>
<tr>
<th></th>
<th>Modals</th>
<th>Words</th>
<th>Clauses</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) If we <strong>could</strong> keep to a blue theme for leotards it <strong>would</strong> make a lovely contrast with the scarves. [KAF 72]</td>
<td>2</td>
<td>19</td>
<td>2</td>
</tr>
<tr>
<td>(2) If you are worried or have questions about the illness, <strong>try</strong> to find someone you <strong>can</strong> trust to talk to about it. [CJ9 2271]</td>
<td>2</td>
<td>22</td>
<td>4</td>
</tr>
</tbody>
</table>

- **Words:** (1) and (2) are fairly equally modalised (10.5% and 9.1% respectively)

- **Clauses (MD):** (1) has twice the MD of (2) (100 and 50 respectively)
Relevant quantitative findings
(written BNC - estimations)

• On average (written BrE), we can expect…
  … about three modal markings per ten clauses (MD=27.7).
  … about 40% of s-units to be modalised (MS=40.9).

• About 85% of *if* tokens are subordinators of conditional constructions. The rest are subordinators of indirect interrogatives.

• *If*-conditionals account for about 80% of all conditional construction tokens.

Written BrE is fairly heavily modalised to start with

The word *if* is not a ‘free agent’

They are excellent candidates for a case study
Modal Load comparisons
Indication that the ML is explained by the constructional nature – not the SP of *if*. 
Most conditionals cluster together (irrespective of subordinator) ... though not all.

Constructions with *if* do not cluster together.

... though not all.

Indirect interrogatives cluster together (irrespective of subordinator).
Conditional constructions with *if do* cluster together

Still ...
The ML of whole constructions may not reflect the SP of *if* within the usual short collocation span of 4-5 words.
When we look at the immediate co-text of *if*, the ML of *if-cnd* and *if-q* is comparable.

An indication that ML is explained by the SP of *if* -- not constructional nature.
When we look at the immediate co-text of *if*, the ML of *if-cnd* and *if-q* is comparable. An indication that ML is explained by the SP of *if* -- not constructional nature.

But this doesn’t explain why *even-if* should have ML significantly lower than *if-q* and *if-cnd* ... 

... or why *whether-q* should have ML very close to *if-q* and *if-cnd*.
When we look at the immediate co-text of *if*, the ML of *if-cnd* and *if-q* is comparable.

An indication that ML is explained by the SP of *if* -- not constructional nature.

More importantly, this co-text has to be defined grammatically.

But this doesn’t explain why *even-if* should have ML significantly lower than *if-q* and *if-cnd* ...

... or why *whether-q* should have ML very close to *if-q* and *if-cnd*.
Overall, the ML clustering of subordinate parts seems to support neither a lexis-only nor a construction-family-only explanation.
The matrix parts of constructions with *if* show significant differences in their ML.
The matrix parts of constructions with *if* show significant differences in their ML.

However, substantial differences are also seen between constructions of the same nature.
The scatterplot shows how balanced the ML is between the subordinate and matrix parts of each construction.

**If-constructions:**
- If seems to have different SP in different constructions

**Lexis + Grammar**

**Polysemous if**
Hypothesis: *if* and *whether* are polysemous:

- \(i_{\text{cnd}} - i_{\text{q}}\)
- \(w_{\text{cc}} - w_{\text{q}}\)

Assumption: The ML of the subordinate parts of the relevant constructions is a reflection of the respective subordinators’ semantic preference.

For the hypothesis to stand, the subordinate parts of *if*-cnd must have substantially different ML from those of *if*-q -- with the same holding between *whether*-cc and *whether*-q.
Polysemy doesn’t seem to be supported
The ML of *if*-conditionals cannot be regarded as reflecting on the semantics of *if* alone, but the interaction of its semantics with the semantics of the constructions of which it is a component part.

Which, if any, of the two theories can better accommodate this?
Lexical Grammar

**Lexical Item / Extended Unit of Meaning**
(Sinclair, 1996: 75, 90; Stubbs, 2009: 123-126)

Components
- **The core (a word or phrase)**
- **Its collocates**
- Its semantic preference
- **Its semantic prosody**
- Its colligations

Lexis independent of grammar

In its current form, LG cannot explain the ML patterns
“[F]requent co-selections of a content word and an associated grammatical frame” (Stubbs, 2002: 238).

“[T]he grammatical company a word keeps” (Hoey, 1997: 8; also Sinclair, 2004: 174).
“The statement of meaning at the grammatical level is in terms of word and sentence classes or of similar categories and of the interrelation of those categories in colligations. Grammatical relations should not be regarded as relations between words as such – between *watched* and *him* in ‘I watched him’ – but between a personal pronoun, first person singular nominative, the past tense” (Firth, 1968: 181)

“[F]requent co-selections of a content word and an associated grammatical frame” (Stubbs, 2002: 238).

“[T]he grammatical company a word keeps” (Hoey, 1997: 8; also Sinclair, 2004: 174).
A hybrid of semantic preference and colligation:

“The mutual attraction holding between a sentence class ... and a semantic category” (Gabrielatos, 2007: 2).

If-conditionals can be seen as modal colligations.
**Semantic Colligation**

*A hybrid of semantic preference and colligation:*

“The mutual attraction holding between a sentence class ... and a semantic category” (Gabrielatos, 2007: 2).

*If*-conditionals can be seen as *modal colligations*

However ...

Not all conditionals have high ML.

The construct doesn’t fully account for the bi-partite structure of conditionals.

The construct is a reduced version of a construction.
Constructions

“Conventionalised pairings of form and function” (Goldberg, 2006: 1)

“Symbolic units” with particular features pertaining to their form and meaning (Croft & Cruse, 2004: 257).

Formal properties:
morphological, phonological, lexical, syntactic

Meaning properties:
semantics, (potential) pragmatic uses

(Croft & Cruse, 2004: 258; Fillmore et al., 1988: 501; Fried & Östman, 2004: 18-21)
Accounting for the ML of different conditionals

Interacting dimensions differentiating between members of the family of conditional constructions:

- The modal marker of the protasis.
- The semantic function of the conditional (largely determined by the modal marking of the apodosis).
- The nature of P-A link (direct or indirect).
- The P-A syntactic link (subordination or co-ordination).

(Gabrielatos, 2010: 323-324)
Thank you

For details and references, please see: