If-conditionals and modality
A corpus-based investigation

Costas Gabrielatos

Lancaster University

Abstract (1): Motivation and focus

A number of studies on modality and/or conditionals have presented the claim that conditionals are intimately connected to modality (Comrie, 1986: 89; Dancygier, 1998: 72; Huddleston & Pullum, 2002: 741; Nuyts, 2001: 352; Palmer, 1986: 189; Sweetser, 1990: 141); however, the nature of that connection has not been investigated empirically. This paper reports on parts of a larger study which empirically tested the above claim – namely the corpus-based approach and metrics developed in the study, as well as some significant findings. More specifically, the paper examines whether, and to what extent, this relationship …

a. holds for all conditionals, irrespective of their subordinator (e.g. *if, in case, unless*);
b. extends to concessive-conditionals (e.g. *even if*);
c. is limited to conditionals (and concessive-conditionals), or extends to other constructions sharing subordinators with conditionals (e.g. indirect interrogatives with *if*).
d. holds for either of the two parts of bi-partite constructions (e.g. protasis and apodosis in conditionals; Fillmore, 1986).

In the case of *if*-conditionals, the paper also examines the extent to which this relationship applies to their two basic types, direct and indirect (Quirk et al., 1985: 1088-1097).

Abstract (2): Data and methodology

The methodology combines a corpus-based, quantitative approach with close analysis of the data for the purposes of the annotation of modal marking in all corpus samples, and the classification of *if*-conditionals. The study uses eleven random samples from the written BNC, containing the following:

a. All types of constructions, providing an indication of the average frequency of modal marking in written British English –which was used as the baseline;
b. Non-conditional constructions, taken collectively;
c. Conditional constructions (e.g. *assuming, if, unless*);
d. Conditional-concessive constructions with *even if* and *whether*;
e. Indirect interrogative (non-conditional) constructions with *if* and *whether*;
f. Constructions with *when* and *whenever* (used as conjunctions), as they have been presented as synonymous with unmodalised *if*-conditionals in some studies (e.g. Athanasiadou & Dirven, 1996: 617, 1997: 62; Palmer, 1990: 174-175).

Abstract (3): Findings

The analysis revealed that the modal load (i.e. the extent of modal marking) in conditionals as a construction family, and *if*-conditionals in particular, is significantly higher than the baseline and non-conditional constructions (taken collectively), as well as most, but not all, non-conditional constructions. More importantly, *if*-conditionals showed a distinctly higher modal load than other conditional constructions.

Overall, constructions of the same family tend to have similar modal load; however, this is not consistently the case with individual constructions within a family. Also, constructions across and within bi-partite families show different ratios of modal load in their two parts. More importantly, the protases of *if*-conditionals have a modal load at least equal to that of the baseline, and, in some cases, significantly higher – despite protases being already modally marked by *if*.

Motivation

A number of studies have claimed (directly or indirectly) that conditionals and modality are intimately related.

Motivation

- "[A] conditional never involves factuality, or more accurately [… it] never expresses the factuality of either of its constituent propositions" (Comrie, 1986: 89).

- "The presence of *if* in the construction marks the assumption in its scope as unassertable. As a result, the assumption in the apodosis […] is not treated as asserted either" (Dancygier, 1998: 72).

- "*If P (then) Q* is a weaker statement that *Q* on its own" (Huddleston & Pullum, 2002: 741).

- "Conditionals are not part of fact-stating discourse: conditionals, instead, express uncertainties" (Turner, 2003: 135).

- "Modality seems […] to be doubly marked in conditionals" (Palmer, 1986: 189).

- "Conditionals have an intimate link with the domain of epistemic qualification" (Nuyts, 2001: 352).

- "The conditional construction is conducive to the expression of modality" (Huddleston & Pullum, 2002: 744).

Motivation

- "[A] conditional never involves factuality, or more accurately [... it] never expresses the factuality of either of its constituent propositions" (Comrie, 1986: 89).

- "The _____ scope as u_____ is not treat_____

- "_If P_____ ston & Pullu_____

- "Con_____ nstead, express uncertainties" (Turner, 2003: 133).

- "Modality seems [...] to be doubly marked in conditionals" (Palmer, 1986: 189).

- "Conditionals have an intimate link with the domain of epistemic qualification" (Nuyts, 2001: 352).

- "The conditional construction is conducive to the expression of modality" (Huddleston & Pullum, 2002: 744).

> The claim seems plausible.
>
> However ...
>
> It has not been examined in depth.
>
> No study offers empirical evidence.

Overarching aim and questions

- To examine the claim using a corpus-based methodology.

- Is this putative connection manifested quantitatively?

  – What do we count?

  – What do we count within?

  – How do we count (metrics)?

  $\rightarrow$ Focus mainly methodological

Terms

Conditionals as "bipartite constructions"
(Fillmore, 1986: 164)

- Protasis (P) [Subordinate part]
  - The part of the construction containing the condition (e.g. the *if* part)

- Apodosis (A) [Matrix part]
  - The part of the construction containing the 'consequence' (the *then* part)

Modal load (ML)
- The extent of modal marking
- Two dimensions → two metrics.

Fundamental typological distinction:
Link between protasis and apodosis
(Quirk et al., 1985: 1088-1097, Gabrielatos, 2010: 236-264)

Direct conditionals (DIR)

The realisation, activation, actuality or factuality of the content of A depends on the realisation, activation, actuality or factuality of the content of P.

- If physicists had tried to discover a way to release nuclear energy before 1939, they would have worked on anything else rather than the field which finally led to the discovery of fission, namely radiochemistry. [B78 1973]

Indirect conditionals (IND)

What is contingent on P is not the content of A, but the relevance of its very uttering, or the wording of its content , or the accurate indication of (aspects of) the referent.

- He's not a bad sort, for a brother if you know what I mean. [AN7 3257]

Research Questions

- Do conditionals have a heavier ML than …
  - average (i.e. written BE seen as a whole)?
  - non-conditional constructions taken collectively?
  - non-conditional bi-partite constructions (e.g. *when*)?
  - concessive-conditionals (*even if, whether*)?
  - indirect interrogatives (*if, whether*)?

- Do all conditionals have comparable ML?
  - *assuming, if, in case, on condition, supposing, unless*

- Do DIR and IND *if*-conditionals have comparable ML?

Data: random samples

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

Written BNC; approx 1000 s-units each.

- S-units
  - Estimation of the average frequency of modal marking in written British English (baseline);
- Non-conditio
- Conditiona ~~provided,~~ *supposing, unle*
- 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).

Thank you,
Stefan Evert and Neil Millar
!!!!!!!!!!!!!!

What do we count *within*?

- In CL, we normally count within the corpus or sample.

  – e.g. occurrences per million words

- How do we account for construction tokens of *varying length* and/or *structural complexity* in the corpus or sample?

     (See Ball, 2004: 297-299; Halliday, 2004: 654)

Modal Density

Definition	Average number of modal markings per clause.
Expression	Number of modal markings per 100 clauses. (%)
Utility	Helps comparisons between samples by normalising for the complexity of the constructions in each.

(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).

Why don't we just calculate
modal markings
per X number of words?

Words vs. opportunities

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

- 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)

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

Definition	Proportion of constructions that carry at least one modal marking.
Expression	Proportion (%) of modalised constructions.
Utility	Corrects for heavily modalised constructions in the sample.

(Gabrielatos, 2010)

Spread:

- The proportion of corpus speakers who use a particular language item (Gabrielatos & Torgersen, 2009; Gabrielatos et al., 2010).

Relevant quantitative findings
(written BNC - estimations)

- About 85% of *if tokens* are in conditional constructions (protases).

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

- On average, we can expect…

  … 40% of s-units to be modalised.

  … three modal markings per ten clauses.

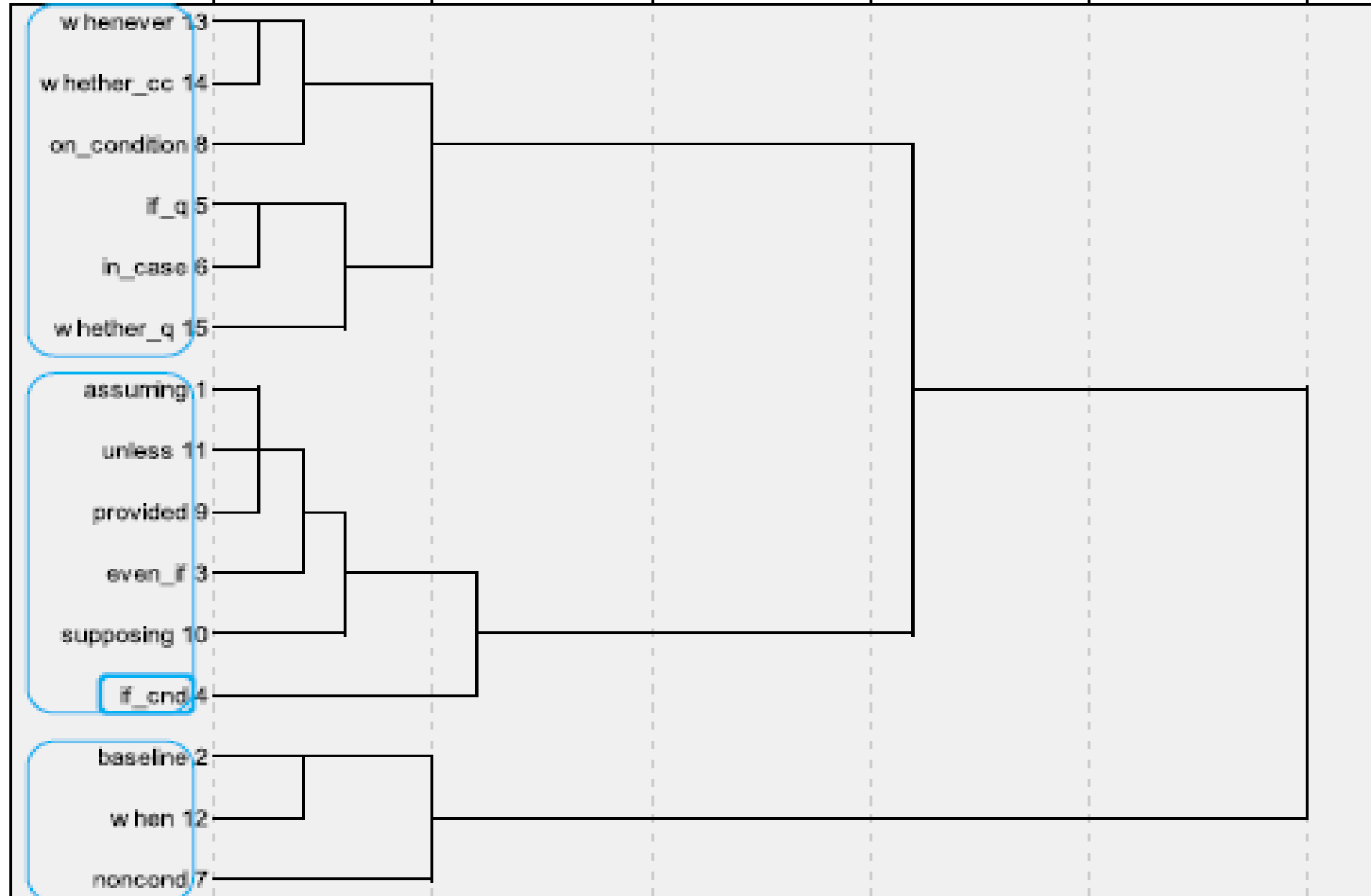# *Modal Load*

The interaction of MD and MS
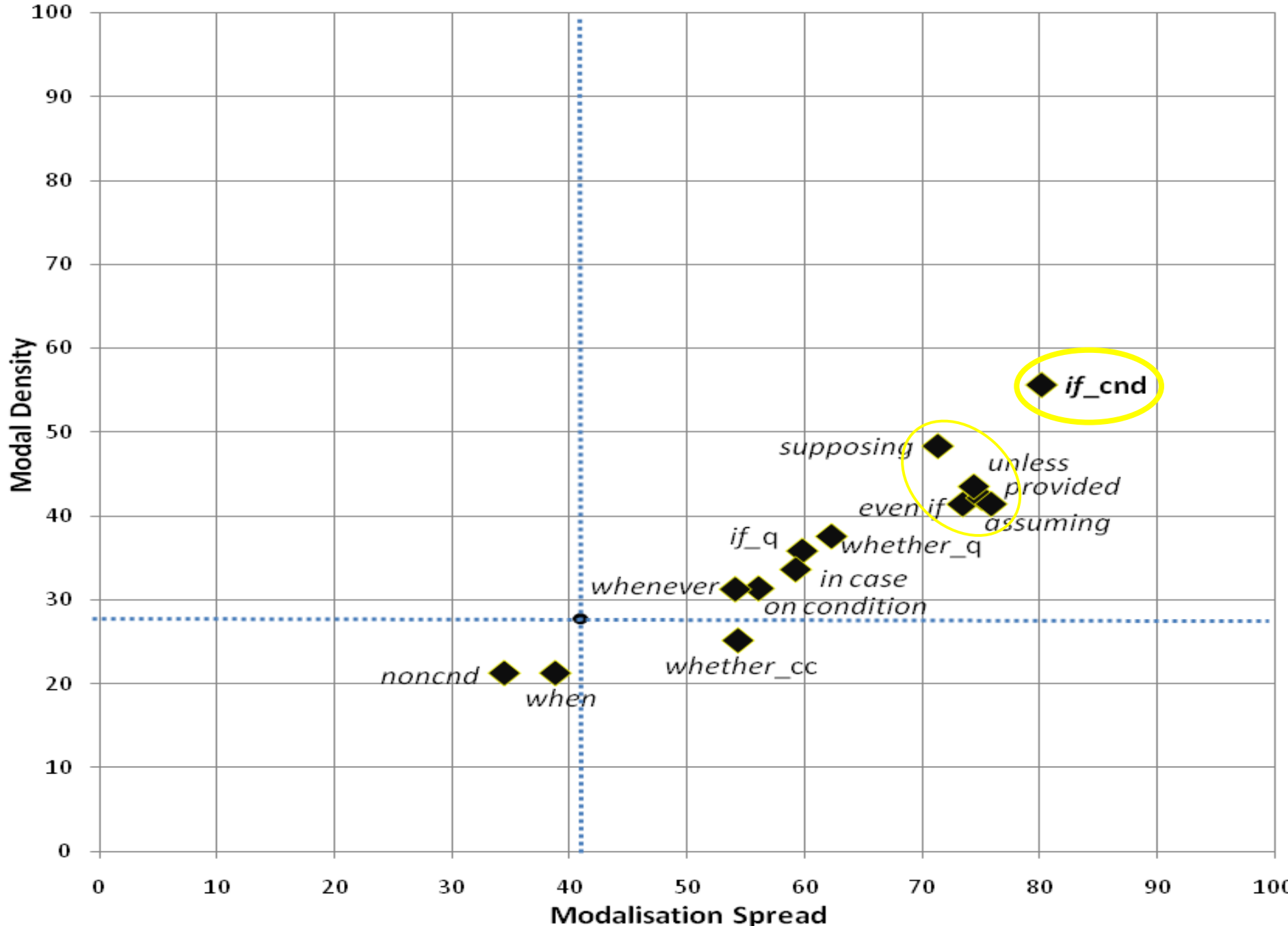
Subordinate part
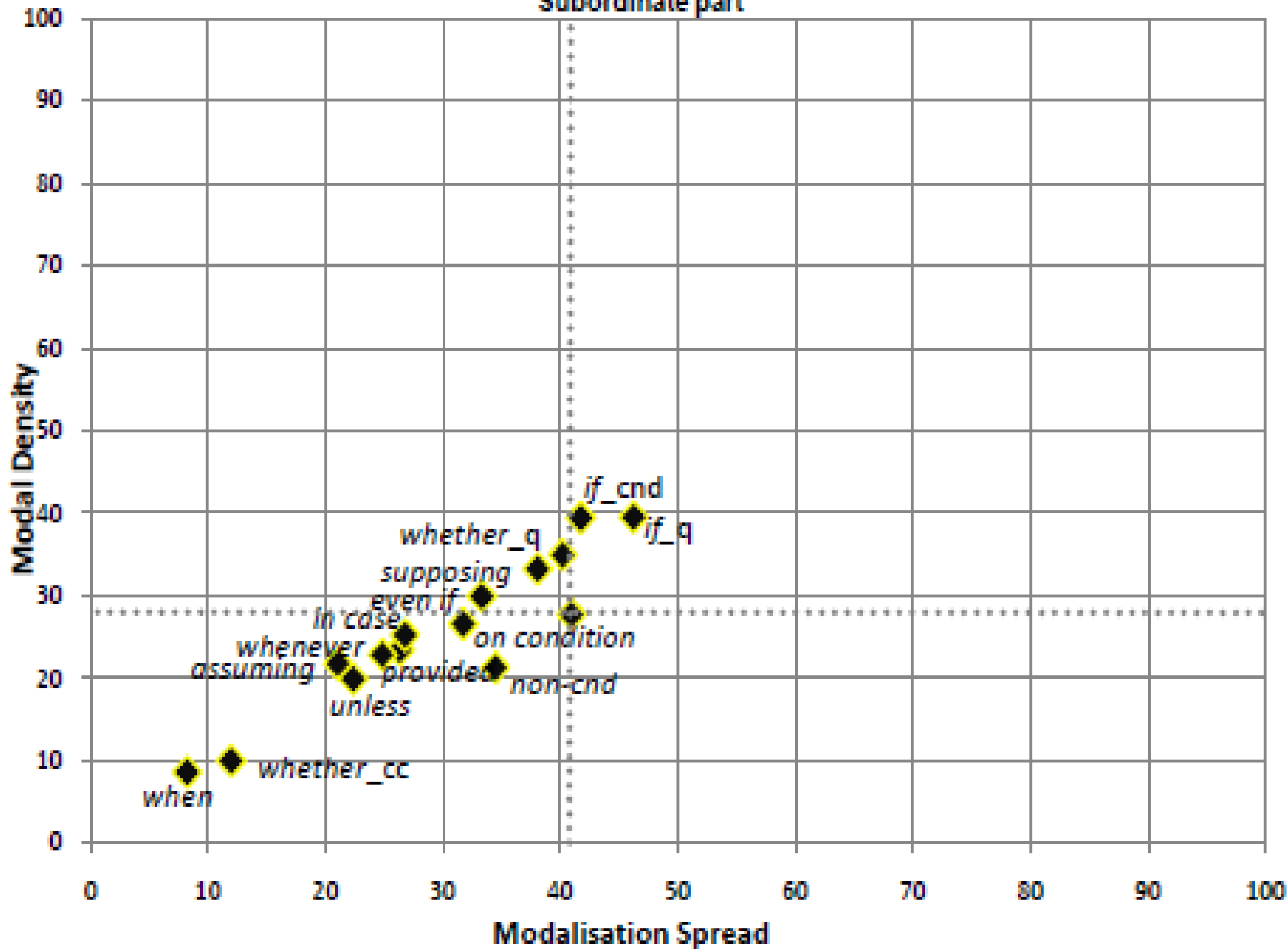
Constructions: MD and MS
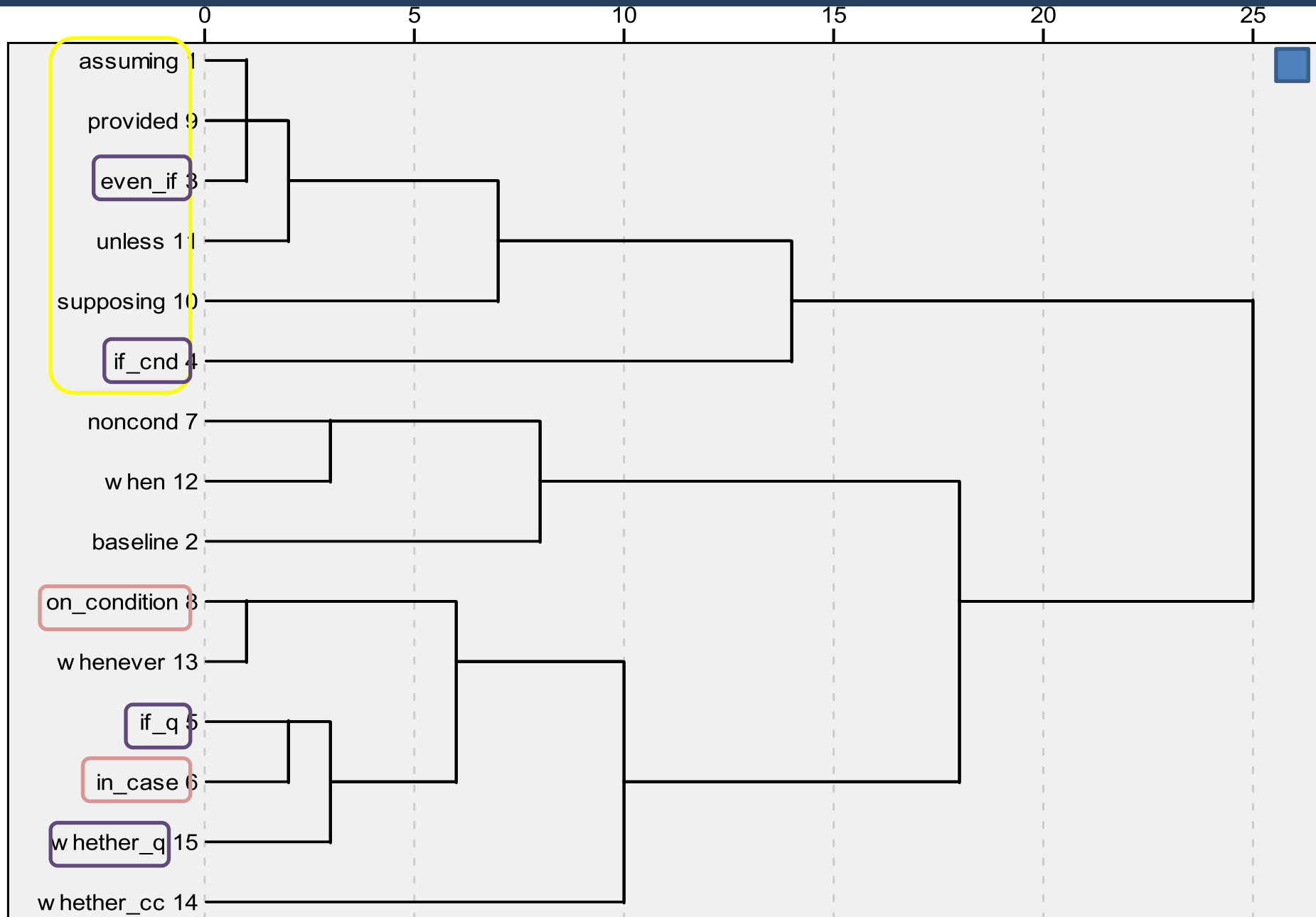
Constructions: MS
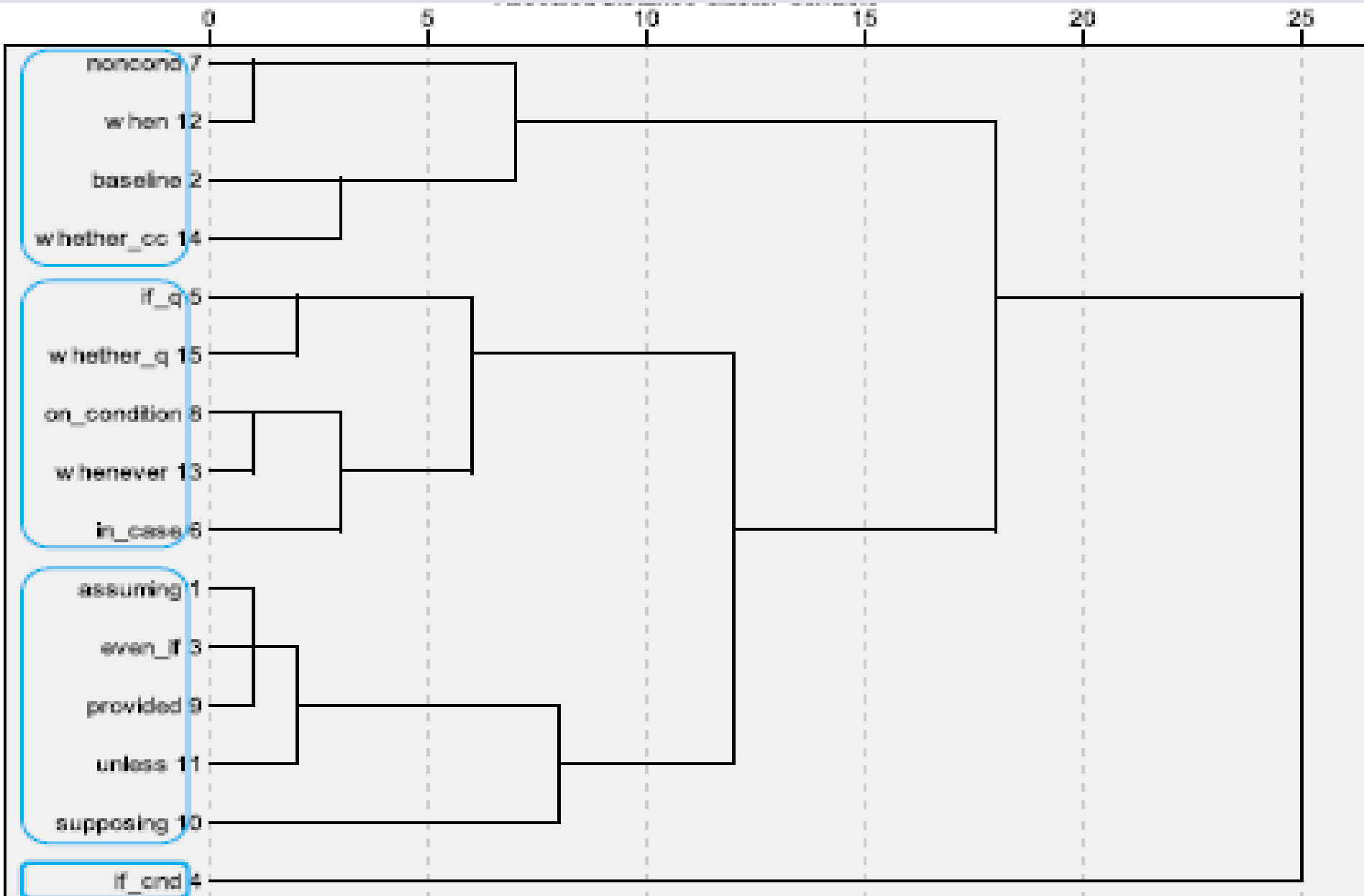
Emerging research question

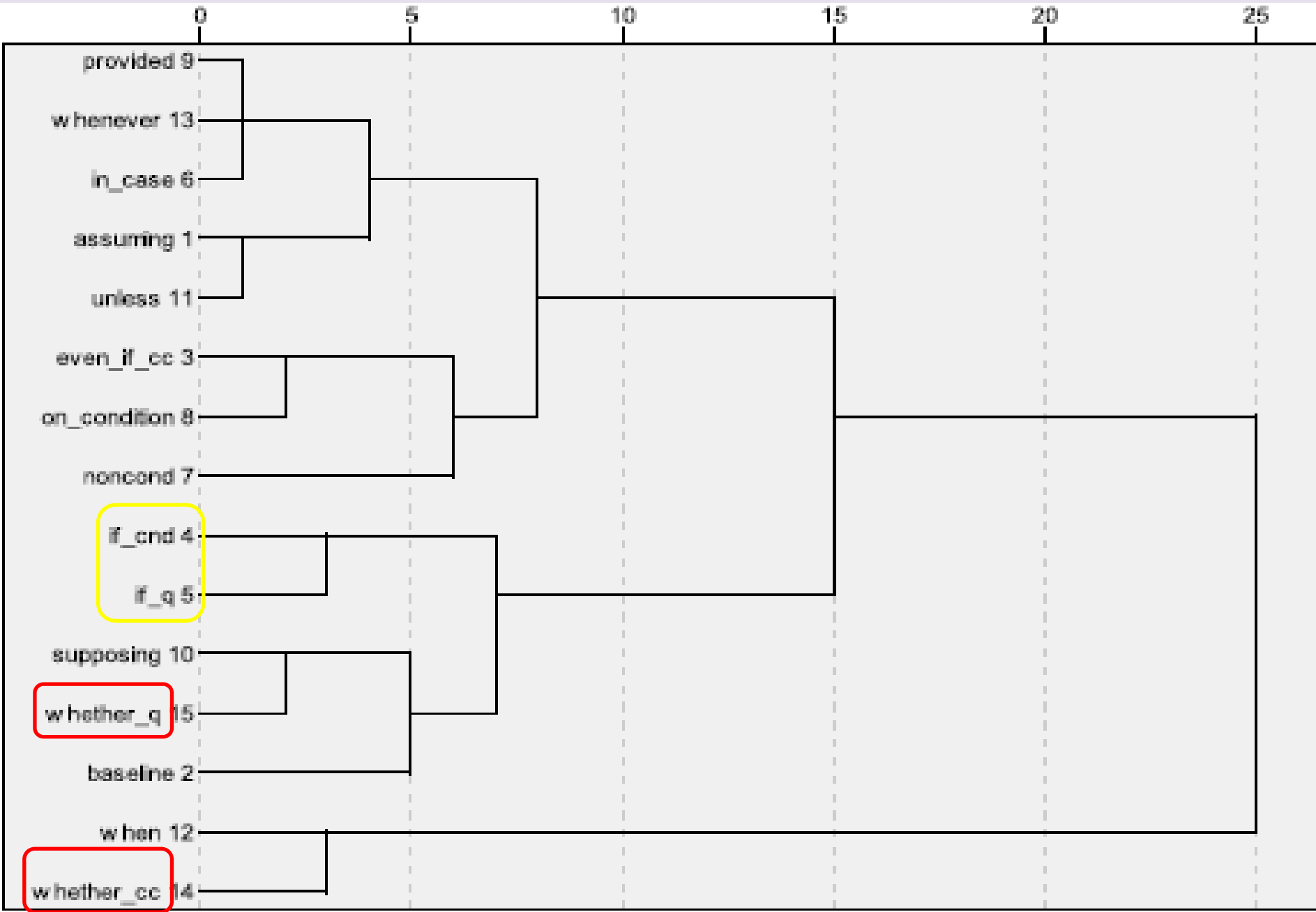Is the modal load due to the type of construction or the semantic preference of the subordinator?

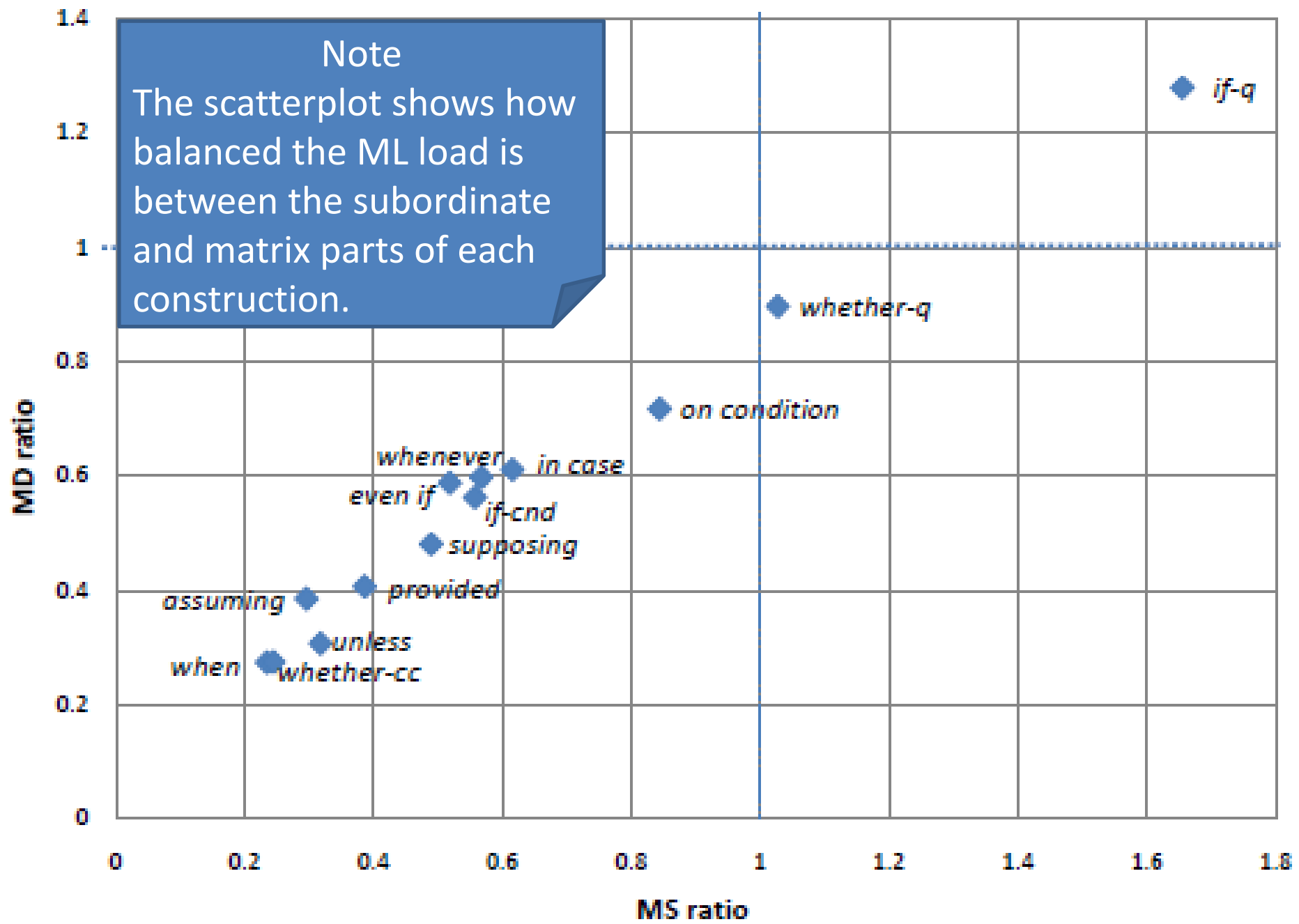Construction Grammar vs. Lexical Grammar

No time to discuss this in any detail
But here's a teaser ...

Modal load: Subordinate/Matrix ratio
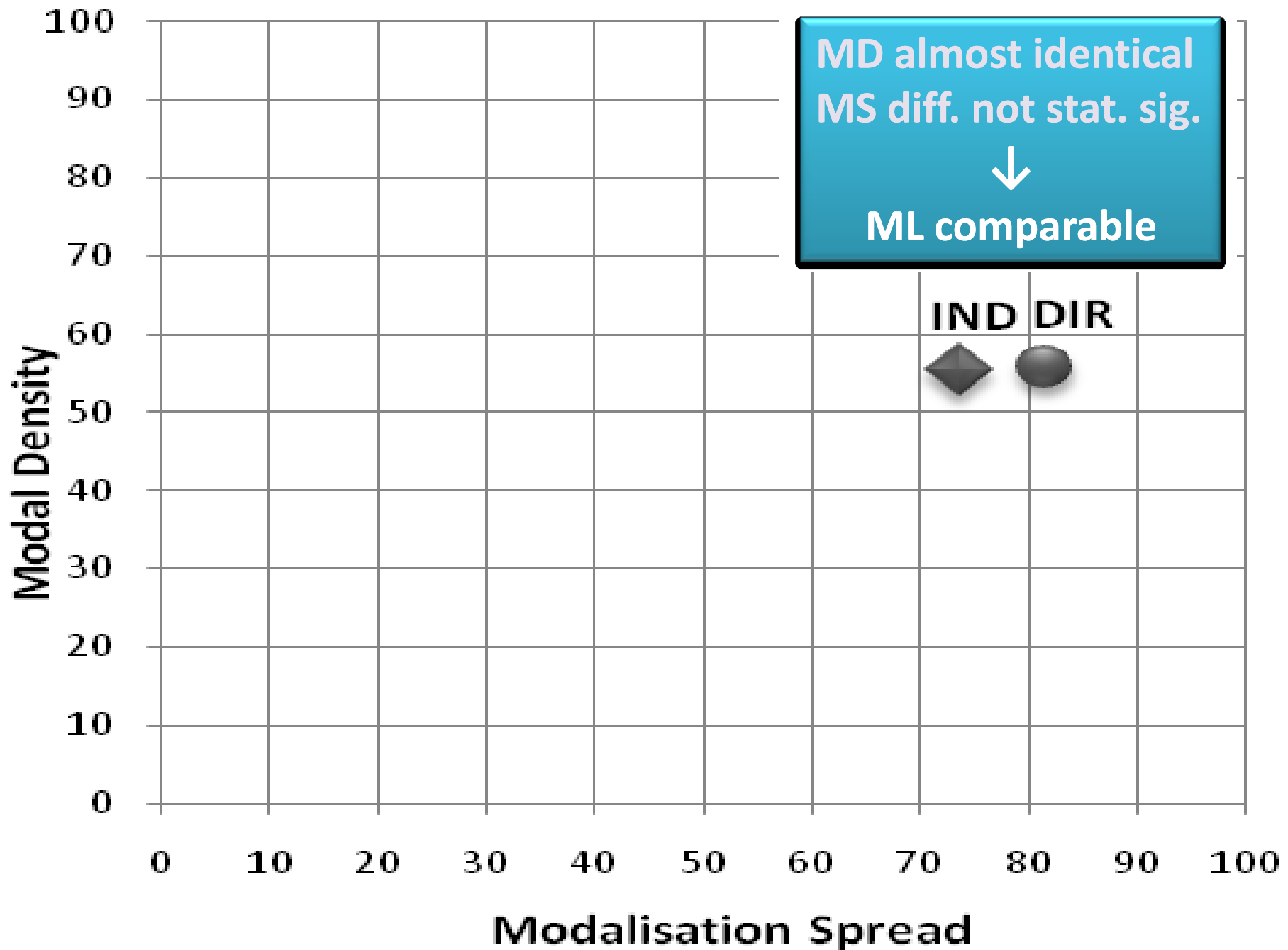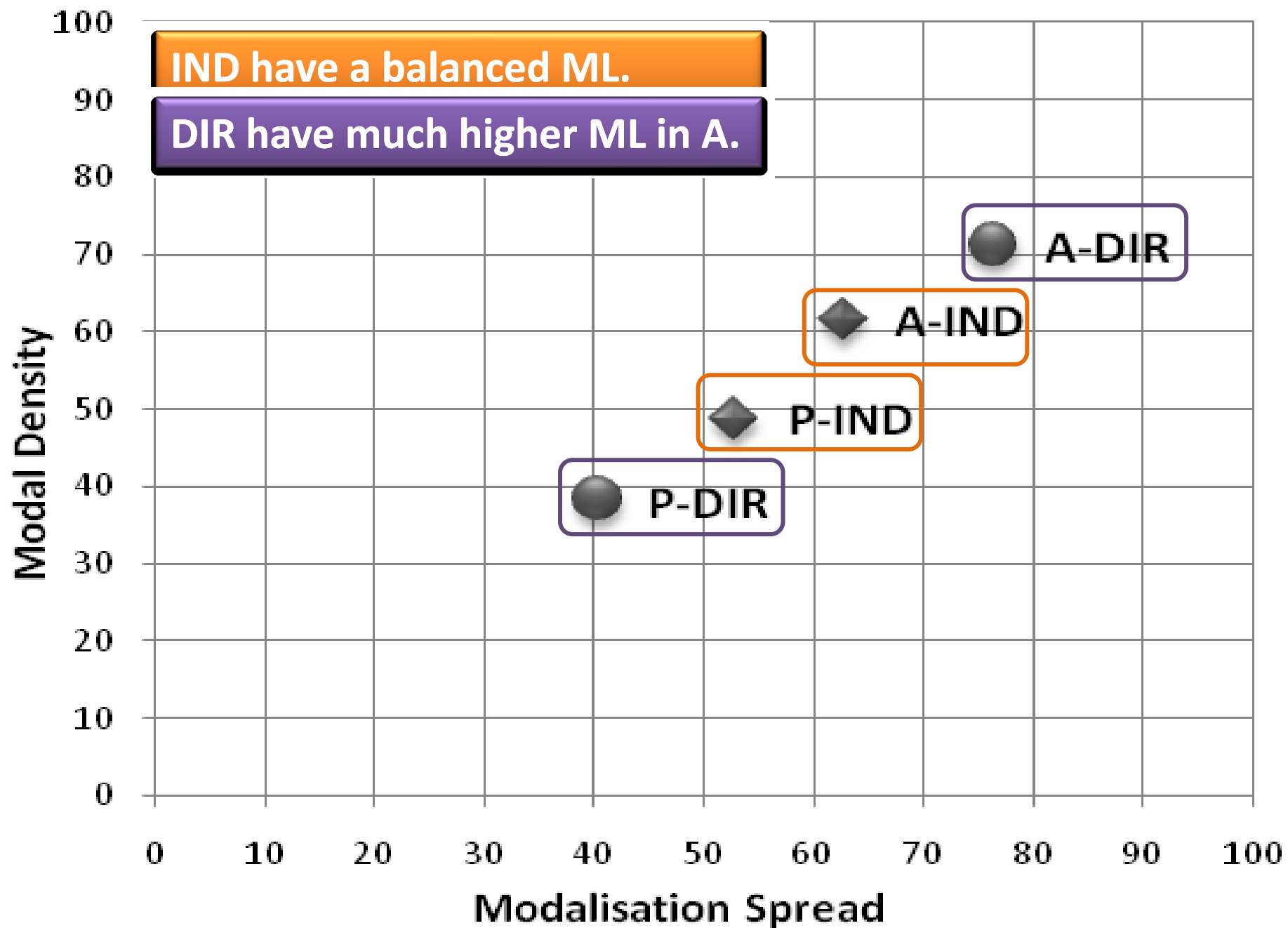
Modal load
of direct and indirect
if-conditionals

Comparison of modal load in P and A

ML of Protases in DIR vs. IND

- In IND …

    … the condition is usually caged in terms of the listener's permission or volition;

    … the semantic function is less important – it is more the pragmatic inferences that determine their function.

    →Their protases have cause to be modalised more frequently than in DIR.

- In DIR …

    … the condition does not necessarily need additional modalisation (usually *if* is enough);

    … the semantic function is carried out by overt modal marking in apodoses.

    → Their protases have less cause to be modalised.

Main points

- Conditional constructions, taken collectively, have a clearly higher modal load than …

  … average.

  … other bi-partite constructions.

- If-conditionals have a clearly higher modal load than …

  … other conditional constructions.

  … non-conditional constructions with *if.*

- Within if-conditionals …

  … the two main sub-types have comparable modal load,

  … but, the balance of ML in P and A reflects their uses.

References (1)

Athanasiadou, A. & Dirven, R. (1996). Typology of *if*-clauses. In E.H. Casad (Ed.), *Cognitive Linguistics in the Redwoods: The expansion of a new paradigm in linguistics*. Cognitive Linguistics Research 6 (pp. 609-654). Berlin: Mouton de Gruyter.

Athanasiadou, A. & Dirven, R. (1997). Conditionality, hypotheticality, counterfactuality. In A. Athanasiadou & R. Dirven (Eds.), *On Conditionals Again* (pp. 61–96). Amsterdam: John Benjamins.

Ball, C.N. (1994). Automated text analysis: Cautionary tales. *Literary and Linguistic Computing*, 9(4), 265-302.

Comrie, B. (1986) Conditionals: A typology. In E.C. Traugott, A. Meulen, J.S. Reilly & C.A. Ferguson (Eds.), *On Conditionals* (pp. 77-99). Cambridge: Cambridge University Press.

Dancygier, B. (1998). *Conditionals and Prediction: Time, knowledge and causation in conditional constructions*. Cambridge: Cambridge University Press.

Fillmore, C.J. (1986). Varieties of conditional sentences. *Eastern States Conference on Linguistics*, Vol. 3, 163-182.

Fillmore, C.J. (1990). Epistemic stance and grammatical form in English conditional sentences. *Papers from the Twenty-sixth Regional Meeting of the Chicago Linguistic Society, Volume 1, The Main Session*, 137-162.

Gabrielatos, C. (2010). *A corpus-based examination of English if-conditionals through the lens of modality: Nature and types*. PhD Thesis. Lancaster University.

Gabrielatos, C. & Torgesrsen, E. (2009). A corpus-based sociolinguistic analysis of indefinite article use in London English. ICAME 30, Lancaster, UK, 27-31 May 2009.

References (2)

Gabrielatos, C., Torgersen, E., Hoffmann, S. & Fox, S. (2010). A corpus-based sociolinguistic study of indefinite article forms in London English. *Journal of English Linguistics*, 38(4), 297-334.

Halliday, M.A.K. (2004). *An Introduction to Functional Grammar* (3rd edn.). London: Arnold.

Huddleston, R. & Pullum, G.K. (2002). *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.

Nuyts, J. (2001). *Epistemic Modality, Language, and Conceptualization*. Amsterdam: John Benjamins.

Palmer, F.R. (1986). *Mood and Modality*. Cambridge: Cambridge University Press.

Palmer, F.R. (1990). *Modality and the English Modals* (2nd edn). Cambridge: Cambridge University Press.

Quirk, R., Greenbaum, S., Leech, G. & Svartvik, J. (1985). *A Comprehensive Grammar of the English Language*. London: Longman.

Sweetser, E.E. (1990). *From Etymology to Pragmatics*. Cambridge: Cambridge University Press.

Turner, K. (2003). On neo- (and post-) Gricean conditionals. *International Journal of Pragmatics*, 14, 135-154.

Ure, J. (1971). Lexical density and register differentiation. In G. Perren & J.L.M. Trim (Eds.), *Applications of Linguistics*. *Selected papers of the second International Congress of Applied Linguistics, Cambridge 1969* (pp. 443-452). London: Cambridge University Press.