

## ***That-trace revisited: an indirect dependency analysis***

The difficulty of long-distance (LD) extraction of subjects in English (and several other languages), traditionally known as the *that*-trace effect, is well known and much debated: in case of non-subject extraction, complementizers are optionally retained, whereas they need to be left out in case of subject extraction (see (1) vs. (2)):

1. Who do you think (that) Tom saw \_\_\_\_ ?
2. Who do you think (\*that) \_\_\_\_ saw Tom?

Two main types of syntactic analyses to explain this can be distinguished: those that argue that null complementizers license subject extraction in ways that overt complementizers do not (Rizzi 1990, Rizzi & Shlonsky 2007 and many others), and those arguing that complementizerless clauses have less structure, thereby facilitating LD subject extraction one way or the other (Ishii, 2004; Brillman & Hirsch, to appear; Douglas, 2017; Bošković, 2016 and many others). I offer a novel analysis, claiming that LD subject extraction in English does not involve LD movement proper, in other words: I claim that subjects cannot be LD extracted at all. Specifically, I argue that the embedded clause is a type of pseudo-relative (cf. McCawley, 1981; 1998) and the *wh*-phrase in the matrix clause its head, as illustrated in (3):

3. Who do you think [[<sub>t<sub>who</sub></sub> [  $\emptyset$  saw Tom]]

Syntactically, the highest *wh*-phrase functions as the object of the matrix verb, but semantically it is the subject of the embedded clause. The embedded CP is typically a zero relative, but can also be introduced by a relative *wh*-phrase or *that*. I relate this construction to the partial *wh*-movement construction in German, adopting the so-called indirect dependency approach as proposed by Felser (2001). English pseudo-relatives are known to be limited in terms of their distribution: they occur predominantly with existential predicates and *it*-clefts, whereas LD subject questions occur with all sorts of bridge predicates. Following a suggestion by Koster (2009) for Dutch partial *wh*-movement and *wh*-copy constructions, I therefore assume that the matrix clause is a concealed cleft with the ‘*it is*’ part phonologically deleted, and that the matrix predicate + subject are actually a parenthetical insert. The latter has also been argued by Bayer (2005), who provides evidence that parenthetical inserts can involve *do*-support. Example (2) thus has the underlying structure in (4):

4. Who ~~is it~~ do you think [<sub>t<sub>who</sub></sub> [  $\emptyset$  saw Tom]]

One of the upshots of the current approach is that it is able to explain several puzzling variants of LD subject extraction. First of all, it has been pointed out repeatedly that the highest *wh*-phrase in (apparent) LD subject extractions can be spelled out as *whom* (Fowler & Fowler 1922; Jespersen 1949; Quirk et al. 1985; Kayne 1980, 2005; Quinn 2005, Armstrong & Mackenzie 2012). This phenomenon is completely unexpected under standard analyses of LD subject extraction, but follows naturally from the proposed analysis: accusative pronouns are most natural in *it*-clefts, even though there are still some prescriptivists that argue otherwise (cf. Johnson, 2015). Based on data from the Corpus of Contemporary American English (COCA, Davies 2008 -...) I show that *whom* occurs just as often for alleged LD extracted subjects as it does for ‘true’ LD extracted objects. Secondly, there are also cases in which the embedded clause in a subject LD dependency is introduced by a *wh*-phrase or *that*, as the examples below from the COCA show.

5. Could you just talk to us about **who** you think **who** \_\_\_\_ is to blame, sir?
6. Bottom line, **what** is the issue you think **that** \_\_\_\_ will win it for you?

Such constructions are again totally unexpected under standard accounts of LD subject extraction: (4) looks like a case of wh-copying, and (5) as a violation of the *that*-trace filter. It is well-known that there are speakers that indeed allow cases like (5) (Sobin, 1987) and this is particularly troublesome for previous analyses of the *that*-trace effect. Examples like (4) are much less known and discussed, but are also attested (see Armstrong & Mackenzie 2012). I attribute the fact that constructions like (4) and (5) are much rarer than (2) to the fact that the zero relativization strategy appears most natural for pseudo-relatives. Unlike normal subject relatives, subject pseudo-relatives allow zero relativization (cf. McCawley, 1998 – a phenomenon also known as subject contact clauses). McCawley states that this is in fact the most common option for pseudo-relatives. My analysis thus gives the following explanation for the *that*-trace effect: the strong preference for deleting *that* in subject LD dependencies is simply due to the embedded clause being a pseudo-relative, in which the option of a zero relative operator is preferred. Concluding, I argue that even English, the poster-child for LD movement, shows examples of indirect dependency formation by which LD-movement proper is avoided.

## References

- Armstrong, N. & Mackenzie, I. (2012). *Standardization, Ideology and Linguistics*. Retrieved from <http://www.palgraveconnect.com/pc/doi/10.1057/.0001>.
- Bayer, J. (2005). Was beschränkt die Extraktion?: Subjekt Objekt vs. Topic Fokus. In F. J. d'Avis (Ed.), *Deutsche Syntax: Empirie und Theorie* (pp. 233-257). Göteborg: Acta Universitatis Gothoburgensis.
- Bošković, Ž. (2016). On the timing of labeling: Deducing Comp-trace effects, the Subject Condition, the Adjunct Condition, and tucking in from labeling. *The Linguistic Review*, 33(1), 17-66.
- Brillman, R., & Hirsch, A. (to appear). *An anti-locality account of English subject/non-subject asymmetries*. In Proceedings of the 50th meeting of the Chicago Linguistic Society, Chicago.
- Davies, M. (2008 - ). *The Corpus of Contemporary American English: 450 million words, 1990-present*. Retrieved from <http://corpus.byu.edu/coca/>.
- Douglas, J. (2017). Unifying the *that*-trace and anti-*that*-trace effects. *Glossa: A Journal of General Linguistics*, 2(1): 60. 1–28.
- Felser, C. (2001). Wh-expletives and secondary predication: German partial wh-movement reconsidered. *Journal of Germanic Linguistics*, 13, 5-38.
- Fowler, H & Fowler, F. (1922). *The King's English*. Oxford: Clarendon Press.
- Ishii, T. (2004). The Phase Impenetrability Condition, the Vacuous Movement Hypothesis, and *That-t* Effect. *Lingua*, 114, 183-215.
- Jespersen, O. (1949). *A modern English grammar on historical principles*. Copenhagen: E. Mungsgaard.
- Johnson, E. S. (2015). *It-clefts and their processing*. (Master's thesis). Retrieved from [https://scholarworks.moreheadstate.edu/msu\\_theses\\_dissertations/29/](https://scholarworks.moreheadstate.edu/msu_theses_dissertations/29/)
- Kayne, R. S. (1980). Extensions of binding and case-marking. *Linguistic Inquiry*, 11(1), 75-96.
- Kayne, R. S. (2005). *Movement and silence*. Oxford: Oxford University Press.
- Koster, J. (2009). *IM not perfect*. (Unpublished Manuscript). University of Groningen, Groningen. Retrieved from <http://www.let.rug.nl/koster/1999.htm>.
- McCawley, J. D. (1981). The syntax and semantics of English relative clauses. *Lingua*, 53(2-3), 99-149.
- McCawley, J. D. (1998). *The syntactic phenomena of English*. University of Chicago Press, Chicago.

- Quinn, H. (2005). The distribution of pronoun case forms in English (Vol. 82). Amsterdam /Philadelphia: John Benjamins Publishing Company.
- Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). *A contemporary grammar of the English language*. London: Longman.
- Rizzi, L. (1990). *Relativized Minimality*. Cambridge, MA: MIT Press.
- Rizzi, L. & Shlonsky, U. (2007). Strategies of subject extraction. In U. Sauerland & H. Gartner (Eds.), *Interfaces + Recursion = Language? Chomsky's Minimalism and the View from Syntax-Semantics* (pp. 115–160). Berlin: Walter de Gruyter.
- Sobin, N. (1987). The variable status of Comp-trace phenomena. *Natural Language & Linguistic Theory*, 5(1), 33-60.