

Gap-less instrumental relative clauses in English

John Beavers
CSLI, Stanford University
210 Panama St
Stanford CA 94305-4115 USA
jbeavers@csli.stanford.edu

Emily M. Bender
Dept of Linguistics
University of Washington
Seattle WA 98195-4340 USA
ebender@u.washington.edu

LSA – Boston, MA
January 9, 2004

1 Introduction

- This paper concerns a class of infinitival verb phrases illustrated in (1):¹
 - (1) a. Give me a Fedex number [to return the package].
b. [@]Christ, and I can't find the money [to buy my kids new shoes]!
c. [@]The authorities claimed they were conspiring to overthrow the government, which is dominated by a different community, the Moors, but offered no evidence [to substantiate this claim].
d. [@]The Church of Ireland rejoined in 1860, because it lacked funds [to continue its own system].
- Interpreted as if they were instrumental infinitival relative clauses, i.e. containing the gapped constituent [*with* ___], with the head noun linked to its missing complement:
 - (2) Give me a Fedex number_{*i*} [to return the package with ____{*i*}].
- The examples in (1) lack gaps; we refer to these as **gap-less instrumental relative clauses**.
- Overlooked by previous treatments of English adnominals (Quirk et al. 1985, McCawley 1988, Sag 1997, Huddleston and Pullum 2002).
- We discuss both the properties of this construction and our methodology, which highlights the utility of integrating grammar engineering and corpus-based techniques into theoretical syntactic analysis.

¹Following Bender and Kathol (2001) we indicate attested examples with a [@]. In this paper, all such examples are from the British National Corpus.

2 Methodology

- We first found the construction while building a corpus-based dynamic treebank (Oepen et al. 2002):
 - Parsing a corpus (in this case, of simulated customer service email), using the LinGO English Resource Grammar (Flickinger 2000).
 - Manually selecting among the candidate parses proposed by the grammar, based on binary distinctions between parses.
 - Saving both the best parse and the choices taken to reach it.
- Sentences containing these gap-less relative clauses were parsed by the grammar, but only with anomalous subject-gap readings:

(3) #Give me a Fedex number_i [_i to return the package].
- To get a better sense of the range of possibilities for the construction, we pulled more examples out of the BNC (Burnard 1995):
 - Chunk-parse the corpus with the RASP parser (Briscoe and Carroll 2002)
 - Using `egrep`, search for NPs followed by *to* + a verb + another NP (to raise the proportion of genuine gap-less examples).

3 Properties

3.1 Distributional

- The head noun must be an instrument/enabler, as though marked by *with/du*e to/etc.:

(4) a. @Jose Lezama Lima’s *Paradiso*, a kind of portrait of the artist, recounting the childhood and adolescence of a young poet, elaborates a poetics in which words are the tools [to understand the world], [to make sense of history] and even [to achieve transcendence].

b. @“I couldn’t find the words [to tell Dorothy], so I said it was a wrong number.”

c. @The kit arrives complete with worms, a pack of starter compost [to get the process going], calcified seaweed [to keep the compost sweet and non-acidic]; there’s even a ready-torn supply of strips of newspaper to go in the bottom.
- Potentially, these could be analyzed as subject gaps, since in English instruments can be subjects.

(5) a. @Exercises [to develop thickness] include seated rowing.

b. You develop thickness with exercises (non-instrumental subject).

c. Exercises develop thickness (instrumental subject).

- However, not all instruments appearing in this construction can be subjects:
 - (6) a. [@]Any form of measurement [to obtain these calculations] is helpful; ie. calipers taken from the original.
 - b. *A form of measurement obtained these calculations.
- Likewise, infinitival *for...to* constructions as in (7a) and reflexive objects as in (7b) disambiguate subject control.
 - (7) a. [@]“You would be the sole owner of the business; my client would be merely an investor in the scheme, providing sufficient funds [for you to purchase your stock-in-trade], and paying the entire rental of the property.
 - b. [@]Given a large enough supply of machine catalogues and enough flexibility of mind [to give yourself different problems], an inventive woodworker could spend the entire winter playing this game.
- Therefore this is a distinct construction from instrumental subject gaps.
- *Reason*-infinitivals (8a) (Lasnik 1987), temporal modifiers (8b,c), and location modifiers (8d,e) appear to allow similar constructions:
 - (8) a. [@]...that awful moment he had so much reason [to dread].
 - b. [@]Surely here, if anywhere, was the time and place [to sacrifice some profit].
 - c. [@]The best time of day [to ride sledges] is at night, and shifting Odd-Knut in daylight needed explosives.
 - d. [@]She liked the raintimes, but this was no place [to savour them].
 - e. [@]“What a place [to earn a living],” she says.
- However, the limited lexemes that may head these constructions suggest analyses as special proforms (*reason*) or gapped nominals (e.g. suitably modified, *time* can serve as an NP adverbial):
 - (9) a. the reason/*explanation/*purpose [to leave]
 - b. He sacrificed some profit only this one time.
 - c. They rode sledges the earliest time of day they were allowed.
 - d. She savoured the raintimes this very spot.
 - e. You can’t earn a living this side of paradise.
- The instrumentals in (4) are more productive (*funds, number, compost, seaweed, etc.*) and may not occur nominally (*He understood the world *(with) these tools*).
- Other covert oblique gaps are ungrammatical:
 - (10) a. comitatives: I saw the man [to travel *(with)].
 - b. governed prepositions: I read the book [to talk *(about)].
 - c. goals of directed motion: I passed by the school [to walk *(to)].
 - d. agentive *by*: I met the man [to be killed *(by)]. (On intended reading)
 - e. inactive cause *from*: The doctor discovered the disease [to die *(from)].

- Finally, these constructions never allow long-distance dependencies:

- (11) a. #This is the number [to think Kim called Sandy].
 b. #This is the number [to try to call Sandy].

- The head noun must be an instrumental associated with the infinitival verb.

3.2 Semantic

- The semantics in many cases is exactly that of an instrument marked by *with* or *via*.
- Some examples involve something more like an enabling semantics:

- (12) c. @The writer didn't have the guts [to put a name to it]!
 b. @Vogtlander added, "any new tax should not just be a means [to raise additional revenues as an easy way out for governments with budget deficits]".
 c. @This gives a lighter, cleaner taste which doesn't overpower and illustrates the fact that it is not just the obvious oily herring, mackerel and salmon which have the body [to withstand a mustard onslaught].

- The difference correlates with the nature of the head, referential heads denoting artifacts (e.g. *money*, *number*) are instruments, abstract heads (e.g. *means*, *flexibility*, *wherewithal*, also *guts* and *body* in the senses here) are enablers.
- Therefore the semantics of the construction appears to be a general means/purpose disambiguated by the head noun (to instrument or enabler).
- Semantically similar to purposives (e.g. *Sandy stayed to watch TV*), where the VP[*to*] modifier ascribes a purpose for the event described by the VP modifiee.
- Presumably related to purposives, although the nominal nature of the head leads to the specific instrument/enabler semantics.
- **Summary:** The construction is an adnominal infinitival VP that links the head noun to some means/purpose role of the head of the VP without a gap.

4 Disambiguation - Isolating the Construction

- Many of the BNC examples had alternative analyses.
- Additional filtering isolated (and validated the existence of) true examples of this construction.

4.1 Subject gap infinitival relatives

- As noted above, some of the relevant data had analyses as instrumental subject gaps.
- We automatically eliminated this ambiguity by looking for *for...to* constructions with overt controllers or reflexive objects that disambiguate control as in (7) above.

4.2 VP[*to*] Complements

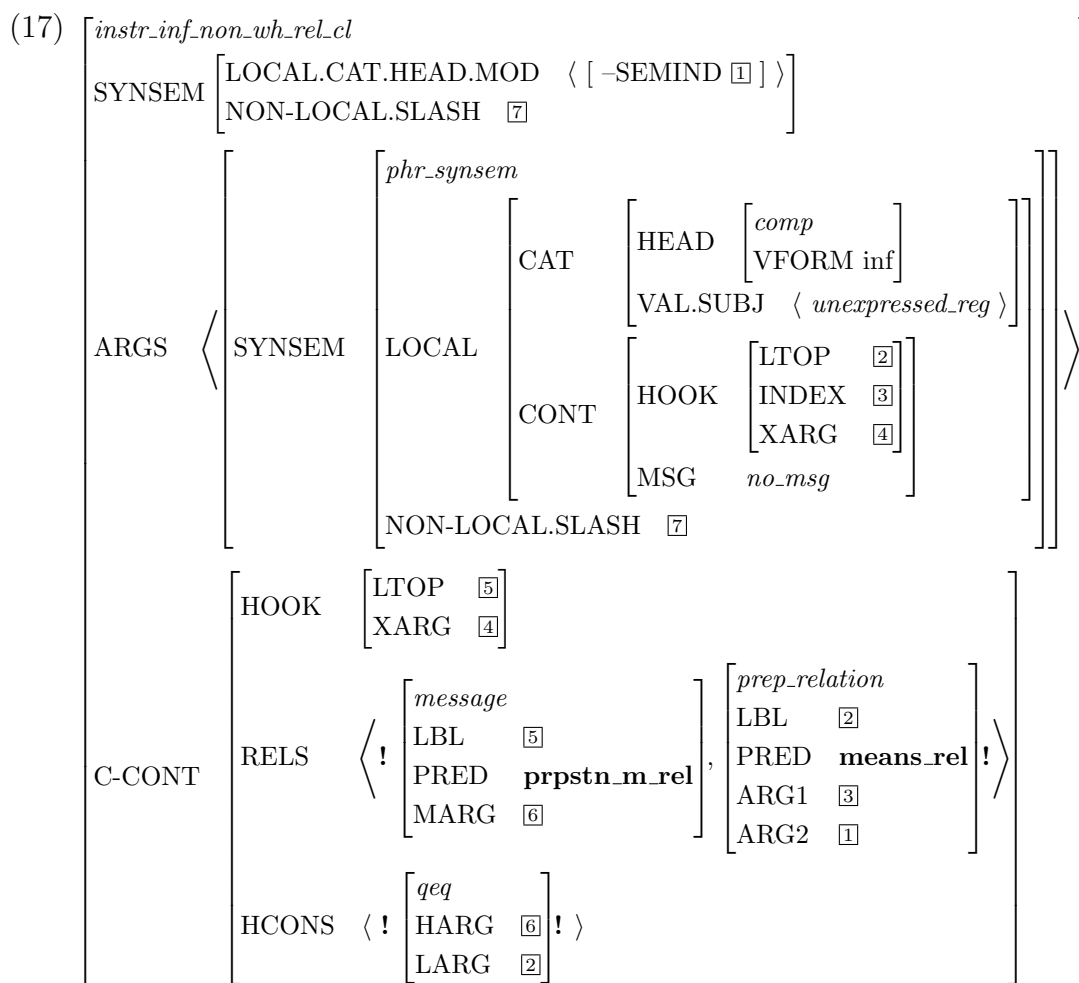
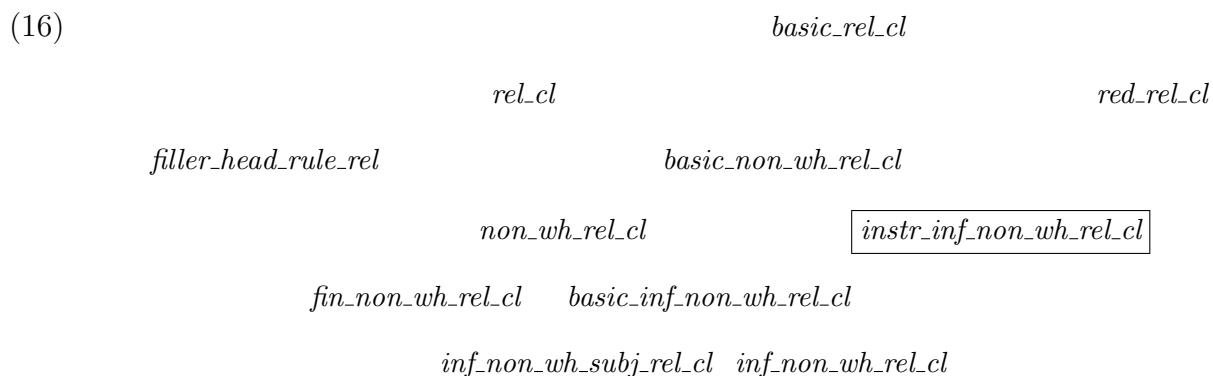
- Some VP[*to*]s are complements of a raising/equi verb or the head noun:
 - (13) a. [@]Levin earnestly considers Situationist film [to be an effective oppositional practice].
 - b. [@]They have been inferring that applying the opposite rudder has been the cause of reversing the spin, whereas the real cause of this happening must be the failure [to uninstall the wings].
- The VP[*to*] in (13a) is a complement of *consider* and in (13b) a complement of *failure* (which is deverbal and event denoting, cf. *We failed to uninstall the wings*).
- In addition to syntactic differences, these constructions do not offer means/purpose semantics (e.g. the purpose of the failure was not to uninstall the wings in (13b))
- Examples such as (13) were eliminated by filtering out specific raising/equi verbs and VP[*to*]-complement-taking nominals (e.g. *ability/failure/attempt/intention*).

4.3 Verbal purposives

- Many examples are ambiguously verbal purposives (with subtly different semantics):
 - (14) a. [@]It is said that the prison director steals the fuel [to warm his own house in Meknes].
 - b. [@]The producer, Sarah D Wilson comments — “the public and press reaction was so strong that we developed another seven programmes [to conclude the story].”
- Context occasionally disambiguates (e.g. purposive readings are infelicitous in (1c,d)).
- Automatic filtering can be done by searching for VP[*to*]s following subjects, where purposives are generally unacceptable (as adverbials):
 - (15) a. [@]Details of the operation [to remove the boat people] are now emerging.
 - b. [@]A trial scheme [to test the viability of the proposal] was called for at last weekend’s Bar conference.
- **Summary:** Searching for *for...to* infinitives following subject NPs headed by non-VP[*to*] complement Ns produced some unambiguous results, although for the remainder of this document we used a wider class of examples and ignore ambiguities.

5 Analysis (HPSG)

- Implemented within the LinGO English Resource Grammar.
- Constructional and lexical approaches are both possible, but the former is more compatible with the current state of the ERG.



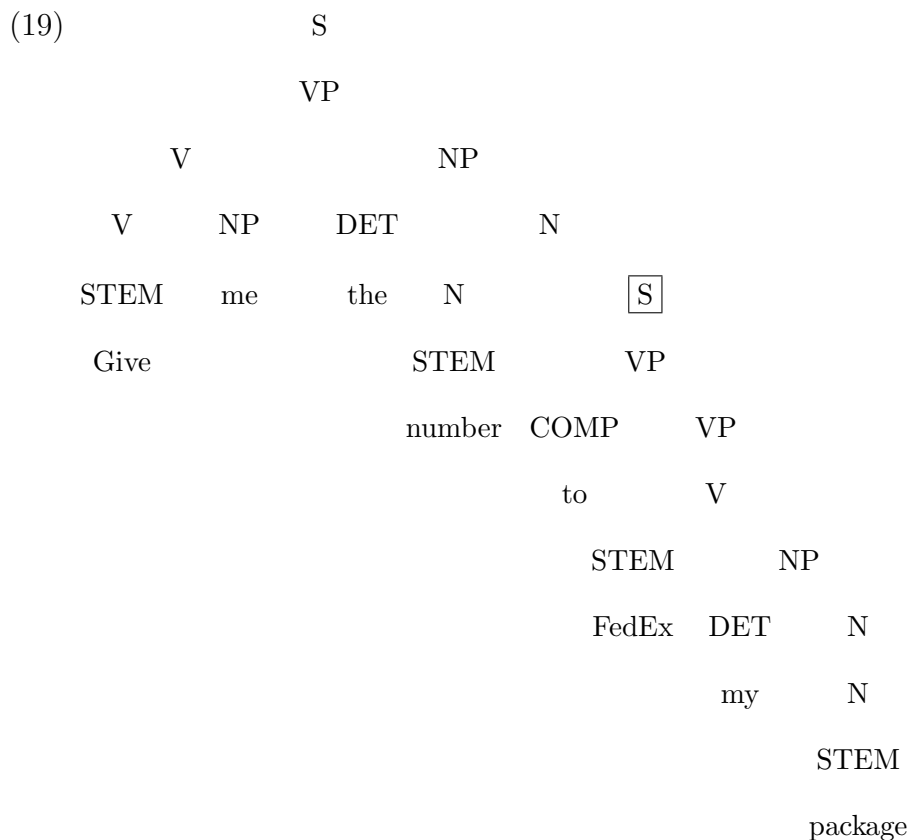
- To note about this construction:
 - It takes one daughter (singleton ARGS list).
 - The daughter is an infinitival CP headed by *to* with a non-overt subject.
 - The daughter must have an empty SLASH list, that is, must be gap-free.
 - The construction adds semantic content (C-CONT): the proposition relation (illocutionary force, cf. Ginzburg and Sag 2000) and the *means* relation.
 - The *means* relation takes as its arguments the event index of the relative clause (Ⓔ) and the referential index of the modified noun (Ⓘ).
 - The *means* relation can be specialized or interpreted differently depending on whether its argument is abstract or concrete (see (18)).
 - Also rules out other roles as in (10) which aren't related to means/purpose.
 - Constraints inherited from *basic_non_wh_rel_cl* and its supertypes include:
 - * MOD value is an \bar{N} .
 - * The relative clause is an intersective (rather than scopal) modifier.
 - * The relative clause is a post-head modifier.
 - * Constraints implementing semantic compositionality.

- Hierarchy of semantic types:

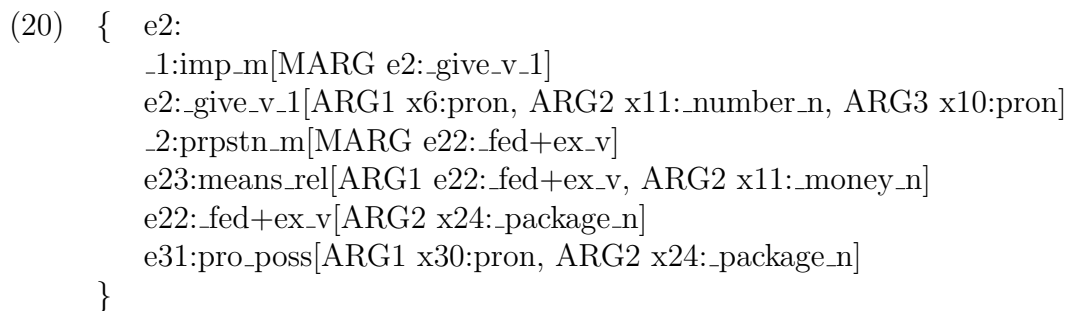
(18) *means_rel*

with_rel *enable_rel*

- Tree illustrating use of *instr_inf_non_wh_rel_cl*:



- Minimal Recursion Semantics (Copestake et al. 2003) representation generated for this sentence:



- The coindexation of the ARG1 of the **means_rel** with the relative clause's index, necessarily that of the event predicated of by the head verb, prevents long distance dependencies (see (11)).

6 Conclusions

- Grammatical:
 - Perhaps similar to gap-less relative clauses in Japanese (Matsumoto 1989, Matsumoto 1997), although much more constrained.
 - The general means/purpose semantics explains why of all role types instruments and enablers (but not comitatives, goals, benefactives, etc.) are the only ones allowed.
 - The semantics and surface level similarities also speak of possible relationships to verbal purposives, an avenue we haven't fully explored here.
- Methodological:
 - Joint human-computer processing of corpora can bring valuable and new data to the attention of linguists.
 - These techniques also allow for automatic and efficient ways of isolating relevant data.
 - Quirky constructions like the one discussed here support grammatical frameworks capable of capturing generalizations at varying degrees of granularity, relating the 'ordinary' aspects of 'peripheral' constructions to the rest of the grammar.

Acknowledgments

We'd like to thank Tim Baldwin, Dan Flickinger, and Ivan Sag for the helpful comments. Any errors and omissions are our own responsibility.

REFERENCES

- Bender, Emily M., and Andreas Kathol. 2001. Constructional effects of *Just Because ... Doesn't Mean ...*. In *Proceedings of the 27th Annual Meeting of the Berkeley Linguistics Society*, Berkeley, CA.
- Briscoe, Ted, and John Carroll. 2002. Robust accurate statistical annotation of general text. In *Proceedings of the 3rd International Conference on Language Resources and Evaluation*, 1499–1504, Las Palmas, Gran Canaria.
- Burnard, L. 1995. *The Users Reference Guide for the British National Corpus*. Oxford: Oxford University Computing Services.
- Copestake, Ann, Dan Flickinger, Carl Pollard, and Ivan A. Sag. 2003. Minimal Recursion Semantics: An introduction. Unpublished ms.
- Flickinger, Dan. 2000. On building a more efficient grammar by exploiting types. *Natural Language Engineering* 6 (1) (Special Issue on Efficient Processing with HPSG):15–28.

- Ginzburg, Jonathan, and Ivan A. Sag. 2000. *Interrogative Investigations: The Form, Meaning and Use of English Interrogatives*. Stanford, CA: CSLI.
- Huddleston, Rodney, and Geoffrey K. Pullum. 2002. *The Cambridge Grammar of the English Language*. Cambridge, UK: Cambridge University Press.
- Matsumoto, Yoshiko. 1989. Japanese-style noun modification ... in English. In K. Hall, M. Meacham, and R. Shapiro (Eds.), *Proceedings of the Fifteenth Annual Meeting of the Berkeley Linguistics Society*, 226–237.
- Matsumoto, Yoshiko. 1997. *Noun-Modifying Constructions in Japanese: A Frame-Semantic Approach*. Amsterdam: John Benjamins.
- McCawley, James D. 1988. *The Syntactic Phenomena of English*. Chicago: University of Chicago Press.
- Oepen, Stephan, Kristina Toutanova, Stuart Shieber, Chris Manning, Dan Flickinger, and Thorsten Brants. 2002. The LinGO Redwoods treebank. Motivation and preliminary applications. In *Proceedings of the 19th International Conference on Computational Linguistics*, Taipei, Taiwan.
- Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech, and Jan Svartvik. 1985. *A Comprehensive Grammar of the English Language*. London: Longman.
- Sag, Ivan A. 1997. English relative clause constructions. *Journal of Linguistics* 33(2):431–483.