Scope Ambiguities and Lexical Resource Semantics: The case of Gapping

Sang-Hee Park, Jean-Pierre Koenig and Rui P. Chaves
University at Buffalo

Gapping can interact with scopal operators to create systematic semantic ambiguities (Siegel 1984). For example, in *John can't live in L.A. and Mary in New York*, the negated modal can be part of each conjunct's interpretation, or it can be interpreted outside the entire coordination. In this paper, we show how Lexical Resource Semantics (Richter & Sailer 2004) (hereafter, LRS) can model these ambiguities purely semantically, without the need to assume syntactic ambiguity (Kubota & Levine 2016; Potter *et al.* 2017). We provide an analysis in which the scope ambiguity in gapping results from independently motivated semantic types for clauses (Champollion 2015) and the underspecification of what is being conjoined. Our paper thus provides a new solution to a long-standing puzzle (why the interaction between gapping and scopal operators allows for readings unavailable to ungapped sentences) and serves as a case study in the benefits of HPSG and LRS in modeling the interface between ellipsis resolution and scope determination.

1 The problem

Gapping is an elliptical construction in which a finite verb, and optionally, other material are omitted from, typically, a non-initial conjunct. As is well-known, gapping can interact with scopal operators such as negation and modals in unexpected ways (Siegel 1984). This is illustrated by the example in (1a), where a main verb as well as an auxiliary verb are missing:

(1) a. John can't live in L.A. and Mary in New York.
    b. John can't live in L.A. and Mary can't live in New York.

This sentence and its supposed ungapped counterpart in (1b) both admit a *distributive-scope reading* (hereafter, DSR), in which the negated modal contributed by *can't* is part of each conjunct's interpretation ($[\neg\Diamond p]\wedge[\neg\Diamond q]$). But (1a) also admits a *wide-scope reading* (hereafter, WSR) that is *not* available in the ungapped counterpart to (1a), i.e. (1b), which is roughly paraphrasable as 'it is not possible for John and Mary to live in two different places' ($\neg\Diamond[p\wedge q]$).

Recent analyses of the ambiguity in (1a) posit two distinct syntactic derivations for (1a) with each derivation corresponding to one of the two readings (Kubota & Levine 2016; Potter *et al.* 2017). We show that positing this syntactic ambiguity is unnecessary, if we assume that scope ambiguities result from various ways of resolving a single, underspecified meaning, an approach to scope ambiguity that is widely accepted within the HPSG community (see Richter & Sailer (2004) and Copestake *et al.* (2005) among others).

2 Syntactic analyses

Potter *et al.* (2017) offer a move-and-delete analysis in which the ambiguity in (1a) is reduced to an ambiguity between coordination below and above T:

(2) [$_{\mathsf{CP}}$ [$_{\mathsf{CP}}$ John can't live in LA] and [$_{\mathsf{CP}}$ Mary$_x$ [in New York]$_y$ t_x ~~can't live~~ t_y]] (DSR)

(3) John$_j$ [$_{\mathsf{T}}$ can't [$_{\mathsf{vP}}$ [$_{\mathsf{vP}}$ t_j live in LA] and [$_{\mathsf{vP}}$ Mary$_x$ [in New York]$_y$ t_x ~~live~~ t_y]]] (WSR)

This analysis predicts a tight connection between the site of coordination (vP/CP) and interpretation (wide/distributive-scope readings): A DSR results if CP coordination is involved, as each conjunct contains its own T (as in (2)); a WSR results if vP coordination is involved, as there is only a single T above the entire coordinate structure (as in (3)). These predictions do not always hold, however. For example, assuming that topicalization in English exclusively targets the left edge of CP, a vP-coordination parse is ruled out for (4). But this makes the incorrect prediction that a WSR would *not* be available in (4), contrary to fact.

(4) (Was your father in a bad mood last night?) During dinner he didn't address his colleagues from Stuttgart or at any time his boss, for that matter. (López & Winkler 2003)

Kubota & Levine (2016), working within Type Logical Categorial Grammar, propose an analysis in which gapping results from conjoining clauses with medial gaps, for example (5):

(5) $\lambda\varphi.\mathsf{john} \circ \varphi \circ \mathsf{steak};\ \lambda Q.Q(\mathrm{s})(\mathrm{j});\ \mathrm{S}|((\mathrm{S}\backslash\mathrm{NP})/\mathrm{NP})$

This analysis relies on the gapping-specific conjunction in (6), which captures the surface asymmetry between the conjuncts in gapping: the first conjunct contains a bound variable in the phonology (φ) and the second conjunct an empty string (ε) that fills in the position of the gap.

(6) $\lambda\sigma_2\lambda\sigma_1\lambda\varphi.[\sigma_1(\varphi) \circ \mathsf{and} \circ \sigma_2(\varepsilon)];\ \lambda W \lambda V.V \sqcap W;\ (\mathrm{S}|X)|(\mathrm{S}|X)|(\mathrm{S}|X)$

In cases where the auxiliary and the main verb are both missing, a constituent consisting of the main verb and an unbound variable representing the auxiliary is derived and then lowered to the conjoined gapped clause, producing a linguistic sign like (7) (see Kubota & Levine (2016) for details):

(7) $\lambda\varphi_0.\mathsf{john} \circ \varphi_0 \circ \mathsf{eat} \circ \mathsf{steak} \circ \mathsf{and} \circ \mathsf{mary} \circ \varepsilon \circ \mathsf{pizza};\ \lambda f.[f(\mathrm{eat(s)})(\mathrm{j}) \wedge f(\mathrm{eat(p)})(\mathrm{m})];\ \mathrm{S}|(\mathrm{VP}/\mathrm{VP})$

This sign can then be given as an argument to the auxiliary, as in (8); as a result, a wide-scope reading of the auxiliary is obtained.

(8)
$$\frac{\begin{array}{ll}\lambda\sigma_0.\sigma_0(\mathsf{can't}); & \lambda\varphi_0.\mathsf{john} \circ \varphi_0 \circ \mathsf{eat} \circ \mathsf{steak} \circ \mathsf{and} \circ \mathsf{mary} \circ \varepsilon \circ \mathsf{pizza}; \\ \lambda\mathscr{F}\neg\Diamond\mathscr{F}(\mathsf{id}_{et});\ \mathrm{S}|(\mathrm{S}|(\mathrm{VP}/\mathrm{VP})) & \lambda f.[f(\mathrm{eat(s)})(\mathrm{j}) \wedge f(\mathrm{eat(p)})(\mathrm{m})];\ \mathrm{S}|(\mathrm{VP}/\mathrm{VP})\end{array}}{\mathsf{john} \circ \mathsf{can't} \circ \mathsf{eat} \circ \mathsf{steak} \circ \mathsf{and} \circ \mathsf{mary} \circ \varepsilon \circ \mathsf{pizza};\ \neg\Diamond[\mathrm{eat(s)(j)} \wedge \mathrm{eat(p)(m)}];\ \mathrm{S}}\ |\mathrm{E}$$

However, this analysis relies on a rather non-standard treatment of modals in *all* contexts (namely, modals are propositional operators that take propositions missing a predicate modifier as argument) which, as we can see, is only needed to derive wide-scope interpretations. Below, we present a novel semantic analysis which does not rely on any gapping-specific assumption.

3 An analysis in HPSG and LRS

We suggest that a more satisfactory solution to the ambiguity in (1a) becomes available if we hypothesize that the various readings of gapping sentences result from different ways of resolving a *single, underspecified meaning*. We show in our paper that the combination of semantic underspecification approaches (Egg *et al.* 2001; Richter & Sailer 2004; Copestake *et al.* 2005) and a surface-based approach to ellipsis (Ginzburg & Sag 2000) leads to a simpler account of the syntax and semantics of gapping sentences.

We present a model of gapping and its scopal interactions with negation and modal operators within HPSG, using LRS for our semantic component. There are two observations any adequate theory of gapping must account for:

- The semantic identity between the gap and the corresponding material in the source clause: The content of the gap must be identical to the content of the corresponding material in the source clause;
- The necessary co-occurrence between gapping and the scope ambiguity in examples like (1a): Wide-scope readings require the presence of the gap (see, for example, (1a) vs. (1b)).

To model the first observation, we propose the constraint in (9). This constraint states that gapped clauses consist of two or more remnants and that each remnant's FOC-CONT value (Hasegawa & Koenig 2011) is identical to its EXCONT value. The constraint also says that the content of the remnant daughters and the MAX-QUD provide the basis for the interpretation of gapped clauses. We assume that the value of MAX-QUD is accommodated on the basis of the focus/nonfocus semantics of the gapped clause and its source, which is consistent with the analysis of focus proposed by Rooth (1992) and Roberts (1996/2012), among others.

(9) *gapped-phrase* ⇒

$$\begin{bmatrix} \text{PHON} & \boxed{1} \oplus \ldots \oplus \boxed{n} \\ \text{CAT} & \begin{bmatrix}\text{VAL} \ \langle \ \rangle\end{bmatrix} \\ \text{LF} & \begin{bmatrix}\text{EXCONT} & \phi \\ \text{PARTS} & \left\langle \ldots,\ \psi, \ldots \right\rangle\end{bmatrix} \\ \text{MAX-QUD} & \psi \\ \text{DTRS} & \left\langle \begin{bmatrix} \text{PHON} \ \boxed{1} & \\ \text{LF} & \begin{bmatrix}\text{EXCONT} & \beta_1 \\ \text{FOC-CONT} & \left\langle \beta_1 \right\rangle\end{bmatrix}\end{bmatrix}, \ldots, \begin{bmatrix} \text{PHON} \ \boxed{n} & \\ \text{LF} & \begin{bmatrix}\text{EXCONT} & \beta_n \\ \text{FOC-CONT} & \left\langle \beta_n \right\rangle\end{bmatrix}\end{bmatrix} \right\rangle \end{bmatrix}$$

(where $n \geq 2$)

To account for the second observation, we assume that the semantics of conjunction is underspecified in that what is being conjoined can be a (proper) part of the first conjunct, provided that the conjoined parts match in semantic type. As Champollion (2015) convincingly argues, tensed and untensed clauses differ in their types: In his approach, tensed clauses are of type t, and untensed clauses are of type $\langle vt, t \rangle$ (predicates of sets of events). This difference in type predicts the ambiguity of (1a): The DSR results when the conjuncts are of type t, i.e., each conjunct includes tense and negated modal operators. The WSR results when the conjuncts are of type $\langle vt, t \rangle$. The tense and negated modal operators in the first conjunct then apply to the result of conjoining the two clauses.

For the semantics of coordination, we assume a standard Boolean conjunction for *and* (Partee & Rooth 1983). Syntactically, we analyze coordinators as a marker which selects the head via SELECT (Chaves 2012). These requirements are specified in the lexical entry for *and* in (10):

(10)
$$\begin{bmatrix} \text{PHON} & \langle \textsf{and} \rangle \\ \text{SS} \mid \text{LOC} & \begin{bmatrix} \text{CAT} \begin{bmatrix} \text{SELECT} & \begin{bmatrix} \text{INDEX} \mid \text{VAR } \boxed{1} \end{bmatrix} \end{bmatrix} \\ \text{CONT} \mid \text{INDEX} \mid \text{VAR } \boxed{1} \end{bmatrix} \\ \text{LF} & \begin{bmatrix} \text{EXCONT} & \boxed{2} \\ \text{PARTS} & \langle \boxed{2} \alpha \wedge \beta \rangle \end{bmatrix} \end{bmatrix}$$

This lexical entry states that *and* bears the same index as the head that it combines with and that its EXCONT value is a conjunction ($\alpha \wedge \beta$).

We propose the following two additional clauses to the LRS Semantic Principles (Richter & Kallmeyer 2009) to constrain the combinatoric semantics of coordination:

(11) The SEMANTICS PRINCIPLE (Clause 4):
If the non-head is a coordinator with an EXCONT value of the form $\alpha \wedge \beta$, then the EXCONT value of the head is a component of β.

(12) The SEMANTIC PRINCIPLE (Clause 5):
(In every non-headed phrase), if the EXCONT value of the non-initial daughter is of the form $\alpha \wedge \beta$, then the INCONT value of the initial daughter is a component of α.

The underspecified representation of (1a) that is licensed by (10)-(12) is provided in Figure 1 (For reasons of the space, we omit the semantic contribution of tense here.):

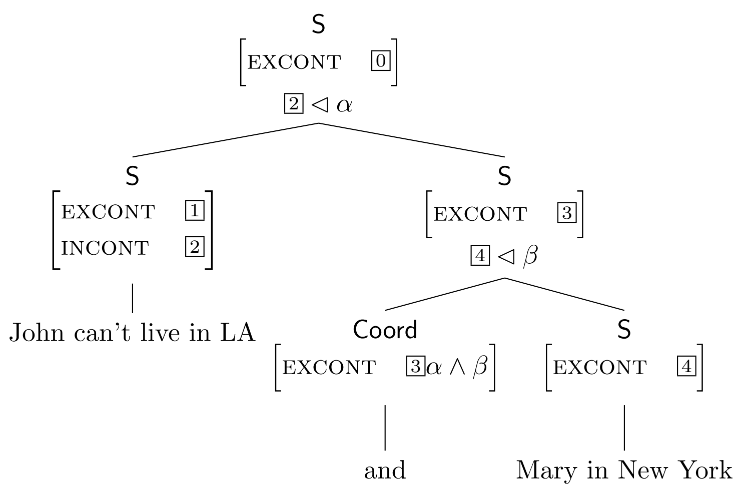

Figure 1: Underspecified representation of (1a)

Given (10)-(12), two different possible values for the EXCONT value of (1a) can be licensed. These are listed in (13): The DSR in (13a) results when the conjunction meaning is identified with the EXCONT value of the entire sentence ($\boxed{3} \equiv \boxed{0}$), and the overall semantics of the first conjunct is identified with the first argument of the conjunction ($\boxed{1} \equiv \alpha$); the WSR in (13b) results when the EXCONT value of the first conjunct is identified with the EXCONT value of the entire sentence ($\boxed{1} \equiv \boxed{0}$), and the first argument of the conjunction is identified with the part of the first conjunct that excludes the negation and modal.

(13) a. $\neg\Diamond\exists e'[\textit{live-in-LA}(e', j)] \wedge \neg\Diamond\exists e''[\textit{live-in-NY}(e'', m)]$ (DSR)
b. $\neg\Diamond[\exists e'[\textit{live-in-LA}(e', j)] \wedge \exists e''[\textit{live-in-NY}(e'', m)]]$ (WSR)

The unavailability of WSR in (1b) follows from the constraint that the semantic type of conjuncts must match (Partee & Rooth 1983) and the fact that tensed clauses are of type t (Champollion 2015).

Semantic underspecification has been shown to enable a simpler interface between syntax and semantics as a single underspecified meaning can cover distinct, fully specified readings (Egg *et al.* 2001; Richter & Sailer 2004; Copestake *et al.* 2005). This paper expands on that research by showing that scope ambiguities in gapping do not require gapping-specific lexical/syntactic constraints or derivations, but result from the underspecification of what is being conjoined and independently motivated semantic types for different types of clauses.[1]

References

Champollion, Lucas. 2015. The interaction of compositional semantics and event semantics. *Linguistics and Philosophy*, **38**, 31–66.

Chaves, Rui P. 2012. On the grammar of extraction and coordination. *Natural Language and Linguistic Theory*, **30**(2), 465–512.

Copestake, Ann, Flickinger, Dan, Pollard, Carl, & Sag, Ivan A. 2005. Minimal recursion semantics: An introduction. *Pages 281–332 of: Research on Language and Computation*, vol. 3. Kluwer Academic Publishers.

Egg, Markus, Koller, Alexander, & Niehren, Joachim. 2001. The constraint language for lambda structures. *Journal of Logic, Language and Information*, **10**, 457–485.

Ginzburg, Jonathan, & Sag, Ivan A. 2000. *Interrogative Investigations: the form, meaning, and use of English interrogatives.* CSLI Publications.

Hasegawa, Akio, & Koenig, Jean-Pierre. 2011. Focus particles, secondary meanings, and Lexical Resource Semantics: The case of Japanese *shika*. *Proceedings of the 18th International Conference on Head-Driven Phrase Structure Grammar.*

Kubota, Yusuke, & Levine, Robert. 2016. Gapping as hypothetical reasoning. *Natural Language and Linguistic Theory*, **34**(1), 107–156.

López, Luis, & Winkler, Susanne. 2003. Variation at the syntax-semantics interface: Evidence from Gapping. *Pages 227–248 of:* Schwabe, K., & Winkler, S. (eds), *The Interfaces: Deriving and interpreting omitted structures.* Linguistk Aktuell 61. Amsterdam; Philadelphia: John Benjamins Publishing Company.

Partee, Barbara H., & Rooth, Mats. 1983. Generalized conjunction and type ambiguity. *Pages 361–383 of:* R. Bauerle, C. Schwartz, & von Stechow, A. (eds), *Meaning, Use, and Interpretation of language.* Berlin: De Gruyter.

Potter, David, Frazier, Michael, & Yoshida, Masaya. 2017. A two-source hypothesis for Gapping. *Natural Language and Linguistic Theory*, **35**(4), 1123–1160.

Richter, Frank, & Kallmeyer, Laura. 2009. Feature logic-based semantic composition: a comparison between LRS and LTAG. *Pages 32–92 of:* Søgaard, A, & Haugereid, P. (eds), *Typed feature structure grammars.* Frankfurt a.M.: Peter Lang.

Richter, Frank, & Sailer, Manfred. 2004. Basic concepts of lexical resource semantics. *Pages 87–143 of: European Summer School in Logic, Language and Information 2003. Course Material I, Volume 5 of Collegium Logicum.* Kurt Gödel Society Wien.

Roberts, Craige. 1996/2012. Information Structure: Towards an integrated formal theory of pragmatics. *Semantics and Pragmatics*, **5**, 1–69.

Rooth, Mats. 1992. A theory of focus interpretation. *Natural Language Semantics*, **1**, 75–116.

Siegel, Muffy E. A. 1984. Gapping and interpretation. *Linguistic Inquiry*, **15**, 523–530.

[1] In contrast to Kubota & Levine, our analysis cannot handle determiner Gapping, e.g. *No dog ate Whiskas or cat Alpo.* But examples like (i) cast doubt on their account since it cannot handle these kinds of examples.

(i) In total, five dogs ate Whiskas and cats Alpo. (the total number of dogs and cats = 5)