# **Coordination, Ellipsis, and Information Structure**

Mark Steedman, University of Edinburgh

July 19, 2012





# Outline

- 1. Introduction: The Problem of Ellipsis
- 2. Grammatical Ellipsis and the Problem of Unbounded Dependency.
- 3. Intonation, Information, and Ellipsis in CCG.
- 4. Conclusion: Can You Do All This in HPSG?



# **The Problem of Ellipsis**

- There are two varieties of ellipsis:
  - Grammatical (syntactic/semantic): e.g. RNR, Gapping, Argument Cluster Coordination, etc.
  - Anaphoric: e.g. VP Anaphora/Ellipsis, *Do So* Anaphora, Sluicing, etc.
- This distinction is related to Hankamer and Sag 1976 deep vs. surface.
- However, they differ on detailed assumptions about which constructions belong where.
- We shall consider only grammatical ellipsis here.



# **Grammatical Ellipsis and Unbounded Dependency**

- Natural Language Grammars appear not to conform to the subsumption condition, a.k.a. the Constituent Condition on Rules (Chomsky 1955/1975, LSLT; Steedman 2000b)
  - The residue of relativization appears to be a non-constituent:
     Articles which I filed without reading
  - Coordination appears to apply to non-constituents:
     I introduced Anna to Manny, and Tom to Sue
  - Intonational phrases appear to be non-constituents:

(You LIKE) (the doggies !) H\* LL%



# What Is To Be Done?

- One (LSLT) response is to make *I filed without reading*, *Tom to Sue*, and *You* LIKE be constituents of type S, via otherwise unmotivated nonmonotonic operations of movement and/or deletion and/or focus projection.
- An alternative (Gazdar 1981; Ades and Steedman 1982; Szabolcsi 1983; Joshi 1988):
  - Make I filed without reading, Tom to Sue, and You LIKE constituents in their own right.
  - Construct all such residues as constituents by near-context-free derivation.
  - Parse with standard divide-and-conquer algorithms and standard statistical (head-dependency) parsing models that run like a bat out of hell.
  - ... with the added-value of capturing long-range dependencies (Hockenmaier and Steedman 2002; Clark and Curran 2004).



# **Categorial Grammar**

• Categorial Grammar replaces PS rules by lexical categories and general combinatory rules (Lexicalization):



- Categories:
  - (2) proved :=  $(S \setminus NP)/NP$
  - (3) think :=  $(S \setminus NP) / S$



# **Categorial Grammar**

• Categorial Grammar replaces PS rules by lexical categories and general combinatory rules (Lexicalization):



- Categories:
  - (2) proved :=  $(S \setminus NP)/NP$  : *prove'*
  - (3) think :=  $(S \setminus NP) / S$  : *think'*



# **Applicative Derivation**

• Functional Application

$$rac{X/_{\star}Y}{X} > rac{Y}{X} < rac{Y X \setminus _{\star}Y}{X} <$$



# **Applicative Derivation**

• Functional Application

$$\frac{X/_{\star}Y:f \quad Y:g}{X:f(g)} > \quad \frac{Y:g \quad X\setminus_{\star}Y:f}{X:f(g)} <$$



# **Applicative Derivation**



# 10 informatics

# **Applicative Derivation**





# **Combinatory Categorial Grammar (CCG)**

• Combinatory Rules:



• All arguments are type-raised in the lexicon, as if they had morphological case:



11 informatics

# **Combinatory Categorial Grammar (CCG)**

• Combinatory Rules:

$$\frac{X/_{\star}Y:f}{X:f(g)} > \frac{Y:g}{X:f(g)} < X \setminus_{\star}Y:f}{X:f(g)} < \frac{X/_{\diamond}Y:f}{X:f(g)} < \mathbf{B}$$

$$\frac{X/_{\diamond}Y:f}{X/_{\diamond}Z:\lambda z.f(g(z))} > \mathbf{B} \quad \frac{Y \setminus_{\diamond}Z:g}{X \setminus_{\diamond}Z:\lambda z.f(g(z))} < \mathbf{B}$$

$$\frac{X/_{\times}Y:f}{X \setminus_{\times}Z:\lambda z.f(g(z))} > \mathbf{B}_{\times} \quad \frac{Y/_{\times}Z:g}{X/_{\times}Z:\lambda z.f(g(z))} < \mathbf{B}_{\times}$$

• All arguments are type-raised in the lexicon, as if they had morphological case:

$$\frac{X:x}{\mathrm{T}/(\mathrm{T}\backslash X)\lambda f.f(x)} > \mathsf{T} \quad \frac{X:x}{\mathrm{T}\backslash(\mathrm{T}/X):\lambda f.f(x)} < \mathsf{T}$$

**Informatics** 

12



# **Combinatory Derivation**





# **Combinatory Derivation**





# Linguistic Predictions: Unbounded "Movement"

• The combination of type-raising and composition allows derivation to project lexical function-argument relations onto "unbounded" constructions such as relative clauses and coordinate structures, without transformational rules:



 $\otimes$  MOVE = MERGE



#### **Predictions: Coordinate Structure Constraint and ATB Exception**

- Because *S*/*NP* is typable, and distinct from *S* we predict the Coordinate Structure Constraint (a,b), the Across-the-Board exception to CSC (c), and the Same Case Condition on the exception to the constraint (d,e) (cf. Gazdar 1981):
  - (9) a. \*a man who I like and you hate him
    - b. \*a man who walks and he talks
    - c. a man who I like and you hate
    - d. \*a man who I like and hates dogs
    - e. ?\*a man who hates dogs and I like

 $(\mathfrak{S})$  (9e) is marginally acceptable because of the possibility of regarding *I like* as a reduced relative clause of the same type  $N \setminus N$  as *who hates dogs*.

# 17 informatics

# **Predictions: Argument-Cluster Coordination**



- $VP = S \setminus NP$ ;  $TV = (S \setminus NP) / NP$ ;  $DTV = ((S \setminus NP) / NP) / NP$
- COPY/DELETE = MERGE



# Syntax = Type-Raising and Composition

- The argument cluster coordination construction (10) is an example of a universal tendency for "deletion under coordination" to respect basic word order: in all constructions in all languages, if arguments are on the left of the verb then argument clusters coordinate on the left, if arguments are to the right of the verb then argument clusters coordinate to the right of the verb (Ross 1970):
  - (11) SVO: \*SO and SVO SVO and SOVSO: \*SO and VSO VSO and SOSOV: SO and SOV \*SOV and SO
- We'll consider some putative examples of exceptions to these generalizations including the ATB condition at the end of the talk.



# SVO Gapping: An Open Problem

♦ Why do SVO languages pattern with VSO and not with SOV?

- Steedman 1990, 2000b show that SVO types force this fact, but do not say how the gap semantics is recovered.
- The strong constraints on Intonation associated with SVO gapping suggest that Information Structure plays a role in ellipsis.



# **Another Prediction: Intonation Phrases**

- (12) Q: I know ARNIM proved SOUNDNESS. But who proved COMPLETENESS?
   A: (MARCEL)(proved COMPLETENESS).
   H\*
   L+H\*
   LH%
  - (13) Q: I know Marcel conjectured SOUNDNESS. But what did he PROVE?
    A: (MARcel PROVED )( COMPLETENESS).
    L+H\* LH% H\* LL%
- Exchanging the A(nswer)s to the same questions is highly unacceptable.



#### Four Dimensions of Information-Structural Meaning

- Steedman 2007a—cf. Pierrehumbert and Hirschberg 1990:
  - 1. Contrast/Background: all accents mark contrast with evoked or inferrable alternative terms;
  - Theme/Rheme: the L+H accents mark Theme (topic); the others mark Rheme (comment). Theme/Rheme marking is "projected" onto constituents by surface derivation;
  - Common Ground: The H\* accents indicate presence in/introduction to Common Ground, whereas the L\* accents indicate absence from/nonaddition to Common Ground;
  - Speaker/Hearer Agency: The L% boundaries indicate Speaker agency; the H% boundaries indicate (claims of) Hearer agency.



# What We Talk About when We Talk about Accent

- When I talk about accent in terms of categories like H\*, I am not talking about pitch as such.
- There is enormous personal variation in the realization of accent in English (Calhoun 2010; Calhoun *et al.* 2010).
- The Pierrehumbert tones are abstract phonological categories, which some speakers like myself realize as pitch.
- The theory is not falsified by the existence of speakers like Glaswegians, who goes up where I go down, and vice versa, or by Finns, who often show little or no pitch range at all.



# **Alternative Semantics for CCG: Accents**

- All categories have two logical forms, the "ordinary" LF  $\Lambda^o$ , and the "alternative"  $\Lambda^a$  in which accented elements are replaced by free variables.
- The proper name Anna bearing an H\* pitch-accent has the following nominative category, among other case-like type-raised categories:

(14) 
$$\begin{array}{l} \operatorname{ANNA} := S_{\top,\rho} / (S_{\top,\rho} \setminus NP_{\top,\rho}) : \left\{ \begin{array}{l} \lambda p. p \ anna \\ \lambda p. p \ v_{\tau_{anna}} \end{array} \right\} \\ \operatorname{H*} \end{array}$$

• A subject bearing no accent has the following category:

(15) Anna:=
$$S_{\pi,\eta}/(S_{\pi,\eta}\setminus NP_{\pi,\eta}): \left\{ \begin{array}{l} \lambda p.p \ anna \\ \lambda p.p \ anna \end{array} \right\}$$

(Where logical forms are identical as here we will write them as one  $\lambda p.p$  anna.)



# **Alternative Semantics for CCG: Boundaries**

- Boundaries are not properties of words or phrases, but independent string elements in their own right.
- They bear a category which "freezes"  $\pm$ ,  $\theta/\pm$ ,  $\rho$ -marked constituents as complete information-/intonation- structural units, making them unable to combine further with anything except similarly complete prosodic units.
- For example, the speaker-supposition- signaling LL% boundary bears the following category:

(16) LL%:= 
$$S_{\phi} \setminus S_{\pi,\eta} : \lambda f.\pi(\eta fS)$$

# A Derivation



"You do not suppose the question of who Anna (as opposed to anyone else) married to be common ground, I make it common ground that she married Manny (as opposed to anyone else)"

nformation

25



# Remarks

- Theme/Rheme marking is projected onto phrasal constituents by syntactic derivation alone.
- It is bounded by combination of the phrase with a boundary tone.
- No independent extrasyntactic mechanism of "Focus Projection" is needed to achieve the semantics of "broad focus"

# The Focusing Particle "only"

### (18) only := $NP^{\uparrow}/NP^{\uparrow}$ : $\lambda np \lambda p \lambda \dots np^{o} p \dots \wedge \forall a \in \{np^{a}\}[a p \dots \rightarrow (a = np^{o})]$



"I suppose the question of who Anna married to be common ground, I make it common ground she married Manny and none of the alternatives."

School of \_.

27



# A Problem Solved

- Unlike Rooth 1992, this locks together derivation and the scope of accent.
- Wold (1996) notes that Rooth's nonsyntactic focus-capture mechanism gets only the infelicitous interpretation (b i), and excludes the felicitous (b ii), for "nested focus" (that is, nested rheme) examples like the following elaborated answers to the question "Who did John introduce to Bill?":
  - (20) a. Anna only introduced  $\operatorname{Sue}$  to Bill.
    - b. i. #Anna also ((only introduced Sue to TOM))
      - ii. Anna also ((only introduced Sue) to TOM)
- We get both. (The derivation for (b ii) is too complex to take in from a slide!)
- We exclude a third crossed dependency analysis (= Anna also introduced Sue to only TOM) apparently allowed by Rooth 2010's "structured meanings" analysis.



### **Intonational Phrases are Constituents**

- The present theory makes intonation structure as defined by intonational boundaries isomorphic with the top-level constituency of surface syntactic derivational structure.
- Surface derivational structure is also, as we have seen, isomorphic to coordinate structure and the domain of relativization.
- It follows that this theory predicts the strongest possible relation between intonation structure, information structure, coordination, and movement, as follows (cf. Steedman 1991, 2000a):
  - All and only those substrings that can either undergo coordination or be extracted over can be intonational phrases and elements of information structure, and *vice versa*.

Steedman, U. Edinburgh



- It has sometimes been suggested on the basis of examples like the following that the Coordinate Structure Constraint and the ATB exception are an illusion (Munn 1993; Yatabe 2003):
  - (21) a. What did you go to the store and buy?
    - b. How much beer can you drink and not get sick?
    - c. This is the stuff that people in the Caucasus drink every day and live to be a hundred.
- Ross 1967; Goldsmith 1985 argued that these extractions involve a distinct, noncoordinate, subordinating lexical category for *and*.
- They note the presuppositional and volitional semantics of the sentences in question (and the absence of such overtones from true coordinates), as well as the fact that no *other* conjunctions support such extractions.



- Nor are such leftward extractions mirrored by equivalent right-node raising, as in (22a), unlike the across-the-board cases like (22b):
  - (22) a. \*Those guys in the Caucasus drink every day, and live to be a hundred, a kind of fermented mare's milk.
    - b. Harry admires. and Louise says she detests, some saxophonist.
- These examples are discussed in Steedman 2007b, as well as by Cormack and Smith (2005). I pass over them here.



- Like many others, Beavers and Sag (2004) and Cann *et al.* (2005:222) overlook the intonational requirements of rightward movement in their respective critiques of the CCG account of extraction, and the similar effect of intonation in facilitating leftward movement.
- It is widely claimed (Munn 1993; Beavers and Sag 2004, *passim*) that rightnode raising is less sensitive to islands than leftward extraction.
- It is less often noticed that examples like (23a) require "focal" stress or accent on the stranded prepositions, as indicated by capitals, and that similar intonated conjoined fragments also seem to license wh-extraction, as in (23b):



- (23) a. CHESTER likes the person who visited us FROM, and LESTER likes the person who gave us a ticket TO, the beautiful island of Capri.
  - b. A place that CHESTER likes the person who visited us FROM, and LESTER likes the person who gave us a ticket TO.
- Any overall stipulation of "barrier" status for relatives (say by stipulating categories like  $(N \setminus N)/(S/NP)$  for relative pronouns in placef of the standard category) will wrongly exclude (23a,b).



- Cluster Coordination examples like the following, in which the types of the cluster conjuncts require different diathesis alternates of the verb *show*, seem not too bad (Beavers and Sag 2004):
  - (24) I showed [three boys a movie] $(S \setminus NP) \setminus (((S \setminus NP)/NP)/NP))$ , and [a video to two girls] $(S \setminus NP) \setminus (((S \setminus NP)/PP)/NP))$ .
- On the assumption that diathesis alternates share the same logical form, the gapping mechanism of Steedman 1990, 2000b with the addition sketched below offers a second, gapping route for such clusters (but see Oehrle 1975).
- Such examples therefore do not necessarily controvert the generalization that coordination is essentially an operation over like types, contrary to their claim.

#### 

# Some Putative Exceptions to the Generalization

- Siegel 1987; Oehrle 1987; Kubota and Levine 2012:
  - (25) a. Harry can't live in Paris and his wife in San Francisco. (¬>&/&>¬)
    b. \*Harry doesn't want to live in Paris and his wife in San Francisco.

 $(*\neg > \&/\& > \neg)$ 

c. \*I won't help Harry live in Paris and his wife in San Francisco.

 $(*\neg>\&/\&>\neg)$ 

- Like many of the above putative counterexamples, (25a) has a feeling of "I've started this sentence, and I'm damned well going to finish it".
- These are not strong facts.


## **Future Work: Information Structure and Gapping**

• I conjecture that the Alternative Logical Form defined above is the locus of the Gap information in the English gapped construction.

(26) ANNA married MANNY and TOM SUE  

$$S: \left\{ \begin{array}{c} married'manny'anna' \\ married'v_{\tau_{manny'}}v_{\tau_{anna'}} \end{array} \right\} \left( X \setminus X / X \right) / X S \setminus ((S/NP)/NP_{SG}): \left\{ \begin{array}{c} \lambda tv.tv \ sue'tom' \\ \lambda tv.tv \ v_{\tau_{sue'}}v_{\tau_{tom'}} \end{array} \right\}$$

• This would fill a hole in the account of gapping as constituent coordination in Steedman (1990).



## Conclusion

- Intonation Structure subsumes Surface Syntactic Derivation, a.k.a. PF.
- Information Structure subsumes Logical Form, a.k.a. LF.
- PF and LF are the only "interface levels"
- LF is the only structural representational level.

 $\bigotimes$  MOVE = MERGE = COPY/DELETE = LF  $\lambda$ -reduction/unification.

• SO . . . ?



## Can you Do the Same Thing in HPSG?

- Yes, in principle (see Klein 2000, who argues that *S/NP*, *S/VP*, etc., should be HPSG-typable to account for prosodic structure in HPSG along lines similar to Steedman 1991).
- However, to account for universal phenomena of cluster coordination and gapping, it looks (*contra* Beavers and Sag 2004) as though HPSG will require the structural equivalent of raised argument types such as  $S \setminus (S/NP)$ , and rules of Composition, as in Karttunen's 1989 "Radical Lexicalism", on top of HPSG slash-feature inheritance.
- If so, you might well want to consider entirely eliminate slash features (as distinct from subcategorization) from HPSG, as we argued was the case for GPSG at the dawn of CCG (Ades and Steedman 1982; Steedman 1985).



## References

Ades, Anthony and Steedman, Mark, 1982. "On the Order of Words." *Linguistics* and *Philosophy* 4:517–558.

 Beavers, John and Sag, Ivan, 2004. "Some Arguments for Coordinate Ellipsis."
 In Proceedings of the 11th International Conference on Head-Driven Phrase Structure Grammar. Leuven: Katholiek Universiteit.

Calhoun, Sasha, 2010. "The Centrality of Metrical Structure in Signaling Information Structure: A Probabilistic Perspective." *Language* 86:1–42.

Calhoun, Sasha, Carletta, Jean, Brenier, Jason, Mayo, Neil, Jurafsky, Dan,



Steedman, Mark, and Beaver, David, 2010. "The NXT-Format Switchboard Corpus: A Rich Resource for Investigating the Syntax, Semantics, Pragmatics, and Prosody of Dialog." *Language Resources and Evaluation* 44:387–419.

- Cann, Ronnie, Kempson, Ruth, and Marten, Lutz, 2005. *The Dynamics of Language*. Syntax and Semantics 35. San Diego: Elsevier Academic Press.
- Chomsky, Noam, 1955/1975. *The Logical Structure of Linguistic Theory*. Chicago: University of Chicago Press.
- Clark, Stephen and Curran, James R., 2004. "Parsing the WSJ using CCG and Log-Linear Models." In *Proceedings of the 42nd Annual Meeting of the Association for Computational Linguistics*. Barcelona, Spain: ACL, 104–111.
- Cormack, Annabel and Smith, Neil, 2005. "What is Coordination?" *Lingua* 115:395–418.

informatics

Gazdar, Gerald, 1981. "Unbounded Dependencies and Coordinate Structure." *Linguistic Inquiry* 12:155–184.

- Goldsmith, John, 1985. "A Principled Exception to the Coordinate Structure Constraint." In *Proceedings of the 21st Regional Meeting of the Chicago Linguistic Society*. Chicago: CLS.
- Hankamer, Jorge and Sag, Ivan, 1976. "Deep and Surface Anaphora." *Linguistic Inquiry* 7:391–428.
- Hockenmaier, Julia and Steedman, Mark, 2002. "Generative Models for Statistical Parsing with Combinatory Categorial Grammar." In *Proceedings of the 40th Meeting of the Association for Computational Linguistics*. Philadelphia, 335– 342.

Joshi, Aravind, 1988. "Tree-Adjoining Grammars." In David Dowty, Lauri

<sup>42</sup> informatics

Karttunen, and Arnold Zwicky (eds.), *Natural Language Parsing*, Cambridge: Cambridge University Press. 206–250.

Karttunen, Lauri, 1989. "Radical Lexicalism." In Mark Baltin and Anthony Kroch (eds.), *Alternative Conceptions of Phrase Structure*, Chicago: University of Chicago Press. 43–65.

Klein, Ewan, 2000. "Prosodic Constituency in HPSG." In Ronnie Cann, Claire Grover, and Philip Miller. (eds.), *Grammatical Interfaces in HPSG: Studies in Constraint-Based Lexicalism*, CSLI Publications. 171–203.

Kubota, Yusuke and Levine, Robert, 2012. "Gapping as like-category coordination." In *Proceedings of the Conference on Logic and Computational Linguistics*. Nantes: Springer, 135–150.

Steedman, U. Edinburgh

Munn, Alan, 1993. *Topics in the Syntax and Semantics of Coordinate Structures*. Ph.D. thesis, University of Maryland.

Oehrle, Richard, 1975. *The Grammatical Status of the English Dative Alternation*. Ph.D. thesis, MIT.

Oehrle, Richard, 1987. "Boolean Properties in the Analysis of Gapping." In Geoffrey Huck and Almerindo Ojeda (eds.), *Syntax and Semantics, Vol. 20: Discontinuous Constituency*, Orlando, FL: Academic Press. 201–240.

Pierrehumbert, Janet and Hirschberg, Julia, 1990. "The Meaning of Intonational Contours in the Interpretation of Discourse." In Philip Cohen, Jerry Morgan, and Martha Pollack (eds.), *Intentions in Communication*, Cambridge, MA: MIT Press. 271–312.

nformati



Rooth, Mats, 1992. "A Theory of Focus Interpretation." *Natural Language Semantics* 1:75–116.

Rooth, Mats, 2010. "Second Occurrence Focus and *Relativized Stress F*." In Malte Zimmerman and Féry (eds.), *Information Structure: Theoretical, Typological, and Experimental Perspectives*, Oxford: Oxford University Press. 14–35.

Ross, John Robert, 1967. *Constraints on Variables in Syntax*. Ph.D. thesis, MIT. Published as Ross 1986.

Ross, John Robert, 1970. "Gapping and the Order of Constituents." In Manfred Bierwisch and Karl Heidolph (eds.), *Progress in Linguistics*, The Hague: Mouton. 249–259.

Ross, John Robert, 1986. Infinite Syntax! Norton, NJ: Ablex.

Steedman, U. Edinburgh

Siegel, Muffy, 1987. "Compositionality, Case, and the Scope of Auxiliaries." *Linguistics and Philosophy* 10:53–75.

Steedman, Mark, 1985. "Dependency and Coordination in the Grammar of Dutch and English." *Language* 61:523–568.

Steedman, Mark, 1990. "Gapping as Constituent Coordination." *Linguistics and Philosophy* 13:207–263.

Steedman, Mark, 1991. "Structure and Intonation." Language 67:262–296.

Steedman, Mark, 2000a. "Information Structure and the Syntax-Phonology Interface." *Linguistic Inquiry* 34:649–689.

Steedman, Mark, 2000b. The Syntactic Process. Cambridge, MA: MIT Press.

Steedman, U. Edinburgh

nformatics



- Steedman, Mark, 2007a. "Information-Structural Semantics for English Intonation." In Chungmin Lee, Matthew Gordon, and Daniel Büring (eds.), *Topic and Focus:Cross-linguistic Perspectives on Meaning and Intonation*, Dordrecht: Kluwer, number 82 in Studies in Linguistics and Philosophy. 245– 264. Proceedings of the LSA Workshop on Topic and Focus, Santa Barbara CA July, 2001.
- Steedman, Mark, 2007b. "On "The Computation"." In Gillian Ramchand and Charles Reiss (eds.), *The Oxford Handbook of Linguistic Interfaces*, Oxford: Oxford University Press. 575–611.
- Steedman, Mark, 2012. *Taking Scope: The Natural Semantics of Quantifiers*. Cambridge, MA: MIT Press.
- Szabolcsi, Anna, 1983. "ECP in Categorial Grammar." Ms., Max Planck Institute, Nijmegen.



Wold, Dag, 1996. "Long Distance Selective Binding: The Case of Focus." In *Proceedings of SALT*. volume 6, 311–328.

Yatabe, Shûichi, 2003. "A Linearization-Based Theory of Summative Agreement in Peripheral-Node Raising Constructions." In *Proceedings of the 9th International Conference on Head-Driven Phrase-Structure Grammar*. Stanford: CSLI Publications, 391–411.