

LING5702: Lecture Notes 9

Syntactic Competence

Until now we've only seen processes for recognizing phonemes, words or flat sequences of words. But that won't help us assemble meanings, which need to distinguish restrictors and nuclear scopes.

This lecture will describe **syntactic** competence (knowledge about **structure** in language).

Contents

9.1	We can identify syntactic categories in human languages like English	1
9.2	Categories compose by rules [Ajdukiewicz, 1935, Bar-Hillel, 1953]	1
9.3	Some rules are non-local [Gazdar et al., 1985, Pollard & Sag, 1994]	2

9.1 We can identify syntactic categories in human languages like English

Human languages seem to have rich category types we can see in conjunctions [Sag et al., 1985].

First, observe that natural languages use different argument structures for different verbs:

- They sleep. – one argument ahead (intransitive)
- They find pets. – one argument ahead and one argument behind (transitive)
- They give people pets. – one argument ahead, two arguments behind (ditransitive)

Next, observe natural languages *coordinate* conjunctions (combine like types): $\langle \alpha \rangle \rightarrow \langle \alpha \rangle$ and $\langle \alpha \rangle$.

- $[\beta$ $[\beta$ They sleep] and $[\beta$ they find pets]]. – sounds ok (β is sentence)
- They find $[\gamma$ $[\gamma$ people] and $[\gamma$ pets]]. – sounds ok (γ is noun phrase)
- *They find $[\gamma$ $[\beta$ they sleep] and $[\gamma$ pets]]. – sounds **wrong**; conjuncts must match

Now, allowable conjunctions give us insight into the category structure of language:

- They $[\delta$ $[\delta$ sleep] and $[\delta$ find pets]]. – sounds ok (δ is verb phrase)
- They $[\eta$ $[\eta$ find] and $[\eta$ give people]] pets. – sounds ok (but what's η ?)

Transitive verbs (find) match type with ditransitive verb + indirect object (give people)!

Both lack argument ahead and behind – it seems types are defined by missing arguments!

9.2 Categories compose by rules [Ajdukiewicz, 1935, Bar-Hillel, 1953]

Formalize set of categories C as follows – clauses with various unmet requirements:

1. every U is in C , for some set U of primitive categories;

2. every $C \times O \times C$ is in C , for some set O of type-combining operators;
3. nothing else is in C

Define **primitive categories** $U = \{\mathbf{N}, \mathbf{V}\}$:

- **N**: noun-headed category with no missing arguments (noun phrase)
- **V**: verb-headed category with no missing arguments (sentence)

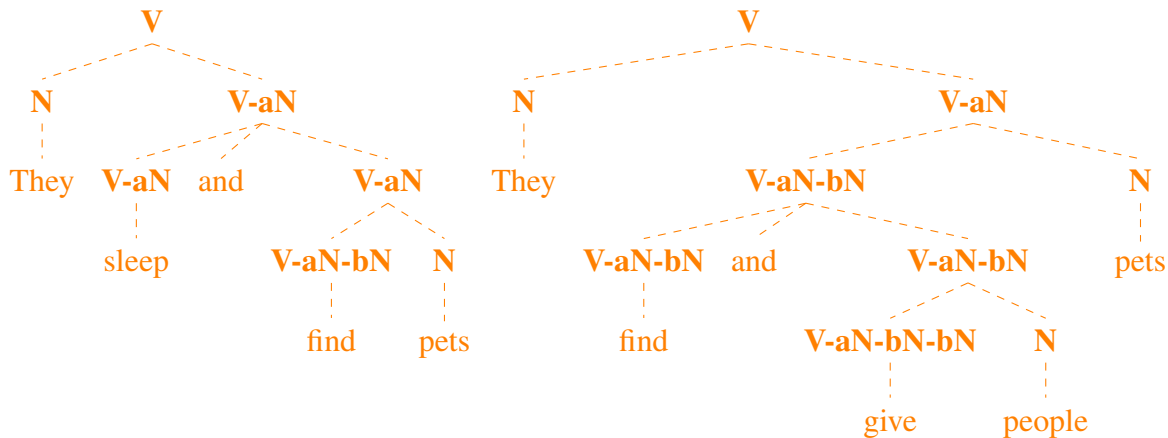
Define **type-combining operators** $O = \{-\mathbf{a}, -\mathbf{b}\}$:

- $\langle \alpha - \mathbf{a} \beta \rangle$: α lacking β argument ahead (e.g. **V-aN** for intransitive δ above)
- $\langle \alpha - \mathbf{b} \beta \rangle$: α lacking β argument behind (e.g. **V-aN-bN** for transitive η above)

Now we can define ‘**context-free**’ rules R over these categories:

- $\langle \alpha \rangle \rightarrow \langle \beta \rangle \langle \alpha - \mathbf{a} \beta \rangle$: argument attachment ahead
- $\langle \alpha \rangle \rightarrow \langle \alpha - \mathbf{b} \beta \rangle \langle \beta \rangle$: argument attachment behind
- $\langle \alpha \rangle \rightarrow \langle \alpha \rangle$ and $\langle \alpha \rangle$: conjunction

These three rules model all of the above sentences:



Also note that the parents in these rules all have simpler types than the children.

This means for any lexicon (constraining types at tree leaves), the set of categories C is finite.

9.3 Some rules are non-local [Gazdar et al., 1985, Pollard & Sag, 1994]

Natural languages may also use non-local dependencies.

In English, these show up in topicalization, which seem to use a gap ‘_’ at one argument:

- *These pets, you say they found _.*

These coordinate as well, but our test shows categories with gaps differ from those without:

- *These pets, you [δ [δ say they found _] and [δ think _ gave people joy]].* – sounds ok

- *These pets, you [δ [v-aN say they found pets] and [δ think _ gave people joy]]. – wrong

We can model this by adding a **new type-combining operator** for non-local dependencies:

- $\langle \alpha\text{-g}\beta \rangle$: α lacking non-local β argument (e.g. **V-aN-gN** for intransitive δ above)

and adding rules to **introduce** non-local dependencies:

- $\langle \alpha\text{-g}\beta \rangle \rightarrow \langle \alpha\text{-a}\beta \rangle$: introduce non-local dependency to argument ahead
- $\langle \alpha\text{-g}\beta \rangle \rightarrow \langle \alpha\text{-b}\beta \rangle$: introduce non-local dependency to argument behind

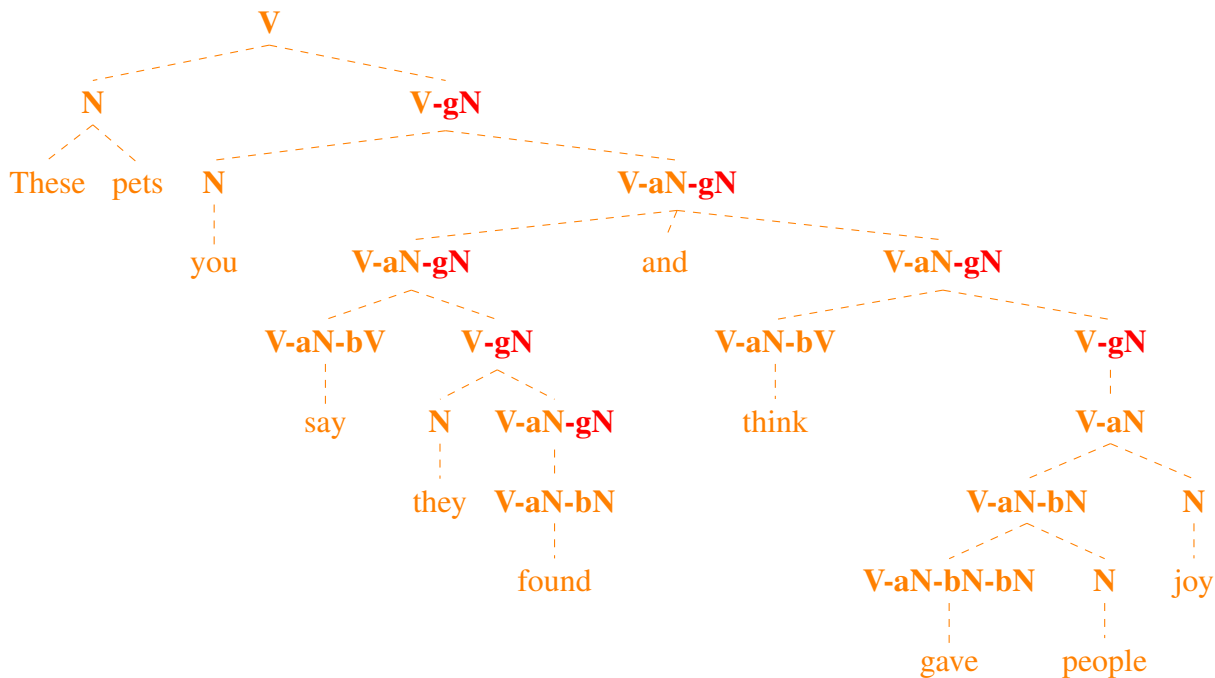
and adding rules to **attach** non-local dependencies:

- $\langle \alpha \rangle \rightarrow \langle \beta \rangle \langle \alpha\text{-g}\beta \rangle$: non-local dependency attachment

and modifying existing rules to **propagate** non-local dependencies $\psi_m \in \{-g\} \times C$:

- $\langle \alpha\psi_{1..M} \rangle \rightarrow \langle \beta\psi_{1..m} \rangle \langle \alpha\text{-a}\beta\psi_{m+1..M} \rangle$: argument attachment ahead, with propagation
- $\langle \alpha\psi_{1..M} \rangle \rightarrow \langle \alpha\text{-b}\beta\psi_{1..m} \rangle \langle \beta\psi_{m+1..M} \rangle$: argument attachment behind, with propagation

Here's the analysis:



Note that M above is unbounded, so our rules no longer guarantee a finite set of categories.

(Any number of arguments may be extracted and propagated up from children.)

Some use evidence like this to argue language isn't context-free but mildly context-sensitive [Shieber, 1985, Joshi, 1985, Steedman, 2000].

In practice, though, we can just constrain category sets to combinations seen in training data.

References

- [Ajdukiewicz, 1935] Ajdukiewicz, K. (1935). Die syntaktische konnexitat. In S. McCall (Ed.), *Polish Logic 1920-1939* (pp. 207–231). Oxford University Press. Translated from *Studia Philologica* 1: 1–27.
- [Bar-Hillel, 1953] Bar-Hillel, Y. (1953). A quasi-arithmetical notation for syntactic description. *Language*, 29, 47–58.
- [Gazdar et al., 1985] Gazdar, G., Klein, E., Pullum, G., & Sag, I. (1985). *Generalized Phrase Structure Grammar*. Cambridge, MA: Harvard University Press.
- [Joshi, 1985] Joshi, A. K. (1985). How much context sensitivity is necessary for characterizing structural descriptions: Tree adjoining grammars. In L. K. D. Dowty & A. Zwicky (Eds.), *Natural language parsing: Psychological, computational and theoretical perspectives* (pp. 206–250). Cambridge, U.K.: Cambridge University Press.
- [Pollard & Sag, 1994] Pollard, C. & Sag, I. (1994). *Head-driven Phrase Structure Grammar*. Chicago: University of Chicago Press.
- [Sag et al., 1985] Sag, I. A., Gazdar, G., Wasow, T., & Weisler, S. (1985). Coordination and how to distinguish categories. *Natural Language & Linguistic Theory*, 3, 117–171.
- [Shieber, 1985] Shieber, S. (1985). Evidence against the context-freeness of natural language. *Linguistics and Philosophy*, 8, 333–343.
- [Steedman, 2000] Steedman, M. (2000). *The syntactic process*. Cambridge, MA: MIT Press/Bradford Books.