# The Tale of the Great Green (not Green Great) Dragon: How to puzzle about adjectives... for science!

Meaghan Fowlie
mfowlie@ucla.edu – meaghanfowlie.com

UCLA Linguistics

February 2, 2015
SFSU

"I first tried to write a story when I was about seven. It was about a dragon. I remember nothing about it except a philological fact. My mother said nothing about the dragon, but pointed out that one could not say 'a **green** **great** dragon', but had to say 'a **great** **green** dragon'. I wondered why, and still do. The fact that I remember this is possibly significant, as I do not think I ever tried to write a story again for many years, and was taken up with language."

– J.R.R. Tolkien in a letter to W.H Auden (Carpenter, 1981)

http://img2.wikia.nocookie.net/__cb20070313193828/forgottenrealms/images/0/01/Green__dragon.JPG

## Roadmap

1. Adjective data
2. Grammars
3. 3 grammars for adjectives
4. A bit of formalisation

# Adjective ordering

(1) a. a **great** **green** dragon

b. *a **green** **great** dragon

(2) a. The big bad wolf

b. *The bad big wolf

(3) a. a good red heavy table

b. *a red heavy good table

(4) a. The small ancient triangular green Irish pagan metal artifact

b. *The metal green ancient triangular pagan Irish small artifact

# What is linguistics?

- A natural science
- Object of study: human language

**Mystery:** This baby is learning things about her native language that generations of linguists haven't been able to work out yet.

- What is it about the human brain that makes this possible?
- What is it about babies that makes this possible?
- **What is it about human language that makes this possible?**

# Adjectives across languages

(5) a. dugacka uska ulica *uska dugacka ulica
long narrow street *narrow long street (Serbo-Croat)

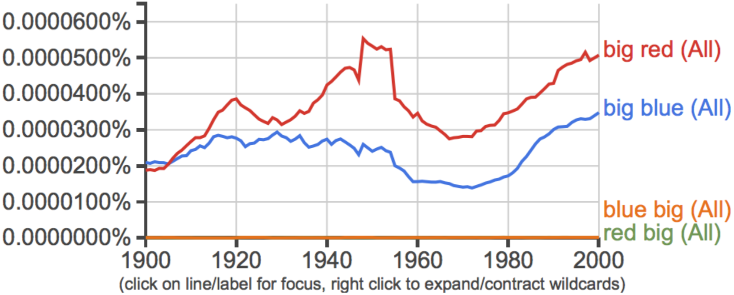
b. lyhyt ohut terä *ohut lyhyt terä
short thin blade *thin short blade (Finnish)

c. en lang vid kjol *en vid lang kjol
a long wide skirt *a wide long skirt (Swedish)

d. haf hir poeth *haf poeth hir
summer long hot *summer hot long (Welsh)

Data from Scott (2002)

# Corpus data



Google N-grams

# Adjective classes

Tendency across languages (partial):

(6)

| D | Size | Age | Shape | Color | Origin | Material | N |
|---|---|---|---|---|---|---|---|
| a | big | old | round | red | indian | rubber | ball |

- **Why?**
  - We don't know. Some attempts to account for it on general semantic grounds, but not entirely convincing
- **How?**
  - What kind of grammar do we need to model this?
  - Can/should we build this general ordering into the universal structure of grammars to account for the cross-linguistic parallels?

## Things to account for

1. You can have lots of modifiers on one noun (*good heavy red table*)
2. Some orders just sound weird (**pine old floorboards*)
3. Sometimes switching the order changes nothing (*beautiful cute baby*, *cute beautiful baby*)
4. Some orders change the meaning (*blistered varnished* wood = varnished wood that got blistered, *varnished blistered* wood = wood that was blistered and then varnished
5. Semantic adjective classes are similarly ordered across languages
6. Prepositional phrases and relative clauses come after the noun (dog *with long ears*, baby *who cried the whole flight*)

## More things to account for

1. Some languages allow adjectives to repeat (*my love is like a red red rose*)
2. Modifiers are nearly always optional (*my love is like a rose*)
3. Occasionally they're required (*John makes a good father, *John makes a father*)
4. Prepositional phrases and relative clauses are *not* strictly ordered. (*A linguist with tin ear at a conference, A linguist at a conference with a tin ear*)
5. Modifiers can themselves be modified (*a surprisingly short basketball player* )
6. Some things can be both modifiers and "arguments" (*They seem nice*, *He put it on the table*)

# What's a Grammar?

Definition (Grammar)

A *grammar* is just a set of rules for making words or sentences

Example: Hawai'ian words

1. The sounds available are: i,e,a,o,u,m,n,p,k,ʔ,h,w,l
2. Divide the sounds into two sets, *vowels* and *consonants*, as follows:

| | |
|---|---|
| Vowels | i,e,a,o,u |
| Consonants | m,n,p,k,ʔ,h,w,l |

3. You can make a word by putting together a sequence of vowels and consonants, but you can never put two consonants next to each other, and every word ends in a vowel.

ahi, kahi, uahi, uahi, aloha, huali, kakahi, uai, uhai, kuai, wawai, hawaiʔi, mele, kalikimaka, *klikmak, *krɪsməs, aaaaaaa

# A Grammar for Modifiers

Suppose these are our words: *a, big, great, bad, red, green, Canadian, wolf, rose, dragon*

Divide them into sets:

| | |
|---|---|
| D | a |
| A | big, great, bad, red, green, Canadian |
| N | wolf, rose, dragon |

**Rules for making NPs:**

1. an N is an NP
2. if you have an NP, you can put an A before it, and that's an NP too.

*wolf* is an N so *wolf* is an NP
*big* is an A and *wolf* is an NP, so *big wolf* is an NP

# A Grammar for Modifiers

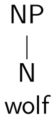
**Rules for making DPs:**

1. if you have a D, you can follow it with an NP, and you have a DP

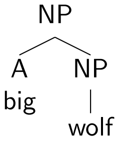
*big wolf* is an NP and *a* is a D, so *a big wolf* is a DP.

# Trees

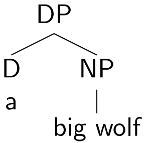
*wolf* is an N so *wolf* is an NP →

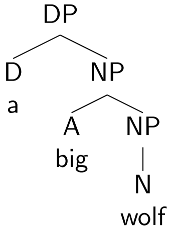


*big* is an A and *wolf* is an NP, so *big wolf* is an NP →



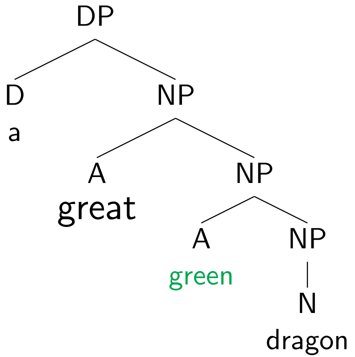
*big wolf* is an NP and *a* is a D, so *a big wolf* is a DP. →

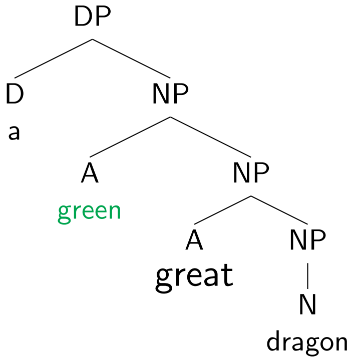




# A Grammar for Modifiers

**Tolkien's problem:**





# A Grammar for Modifiers

**Problems we've solved:**

1. You can have lots of modifiers on one noun (*good red heavy* table)
2. Sometimes switching the order changes nothing (*a beautiful cute baby, a cute beautiful baby* )
3. Some languages allow adjectives to repeat (*my love is like a red red rose*)
4. Modifiers are nearly always optional (*my love is like a rose*)

# A Grammar for Modifiers II: encoded order

Suppose these are our words: *a, big, great, bad, good, red, green, Canadian, wolf, rose, dragon*

Divide them into sets:

| | |
|---|---|
| D | a |
| SIZE | big, great |
| GOOD | bad, great |
| COL | red, green |
| NAT | Canadian |
| N | wolf, rose, dragon |

# A Grammar for Modifiers II: encoded order

**Rules for making natPs:**

1. If you have an N, you can put a NAT before it to make a NATP
2. If you have a NATP, you can put a NAT before it to make a NATP

eg: *wolf* is an N and *Canadian* is a NAT, so *Canadian wolf* is a NATP

**Rules for making colPs:**

1. If you have an N, you can put a COL before it to make a COLP
2. If you have a NATP, you can put a COL before it to make a COLP
3. If you have a COLP, you can put a COL before it to make a COLP

eg: *wolf* is an N and *red* is a COL, so *red wolf* is a COLP
eg: *Canadian wolf* is a NATP and *red* is a COL, so *red Canadian wolf* is a COLP

# A Grammar for Modifiers II: encoded order

**Rules for making goodPs:**

1. If you have an N, you can put a GOOD before it to make a GOODP
2. If you have a NATP, you can put a GOOD before it to make a GOODP
3. If you have a COLP, you can put a GOOD before it to make a GOODP
4. If you have a GOODP, you can put a GOOD before it to make a GOODP

eg: *red wolf* is a COLP and *bad* is a GOOD, so *bad red wolf* is a GOODP

**Rules for making sizePs:**

1. If you have an N, you can put a SIZE before it to make a SIZEP
2. If you have a NATP, you can put a SIZE before it to make a SIZEP
3. If you have a COLP, you can put a SIZE before it to make a SIZEP
4. If you have a GOODP, you can put a SIZE before it to make a SIZEP
5. If you have a SIZEP, you can put a SIZE before it to make a SIZEP

eg: *wolf* is an N and *small* is a SIZE, so *small wolf* is a SIZEP

# A Grammar for Modifiers II: encoded order

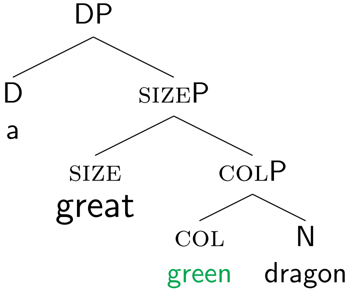
**Rules for making DPs:**

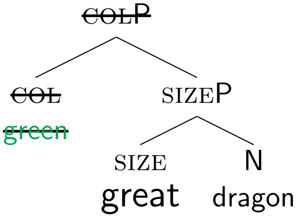
1. If you have an N, you can put a D before it to make a DP
2. If you have a NATP, you can put a D before it to make a DP
3. If you have a COLP, you can put a D before it to make a DP
4. If you have a GOODP, you can put a D before it to make a DP
5. If you have a SIZEP, you can put a D before it to make a DP

*small wolf* is a SIZEP and *a* is a D so *a small wolf* is a DP

# A Grammar for Modifiers II: encoded order

**Tolkien's problem:**





**All the "if you have a sizeP..." rules:**

- If you have a SIZEP, you can put a D before it to make a DP
- If you have a SIZEP, you can put a SIZE before it to make a SIZEP

# A Grammar for Modifiers II: encoded order

**Problems we've solved:**

1. You can have lots of modifiers on one noun (*good heavy red* table)
2. Sometimes switching the order changes nothing (*beautiful cute, cute beautiful* baby)
3. Some languages allow adjectives to repeat (*my love is like a red red rose*)
4. Modifiers are nearly always optional (*my love is like a rose*)
5. **Some orders just sound weird (*pine old floorboards)**

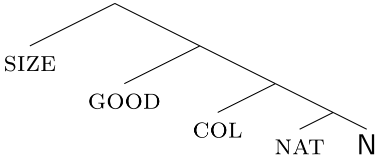
# A Grammar for Modifiers II: encoded order

**This grammar has 19 rules! Try I only had 3.**

**It seems like everything we need to know is contained in the sizeP rules:**

1. If you have an N, you can put a SIZE before it to make a SIZEP
2. If you have a NATP, you can put a SIZE before it to make a SIZEP
3. If you have a COLP, you can put a SIZE before it to make a SIZEP
4. If you have a GOODP, you can put a SIZE before it to make a SIZEP
5. If you have a SIZEP, you can put a SIZE before it to make a SIZEP

# Hierarchy



# A Grammar for Modifiers III: encoded hierarchy

Suppose these are our words: *a, big, great, bad, good, red, green, Canadian, wolf, rose, dragon*
Divide them into sets:

| | |
|---|---|
| D | a |
| A | big, great, bad, good, red, green, Canadian |
| N | wolf, rose, dragon |

**Give the A's numbers:**

| | | |
|---|---|---|
| A-4 | big, great | SIZE |
| A-3 | bad, good | GOODNESS |
| A-2 | red, green | COLOUR |
| A-1 | Canadian | NATIONALITY |

# A Grammar for Modifiers III: encoded hierarchy

**Rules for making NPs:**

1. an N is also an NP-0
2. Suppose we have a number *i*. If you have an NP-i, you can put an A-j before it to make an NP-j, but only if *j* is a least as high a number as *i*

*wolf* is an N so *wolf* is an NP-0
*big* is an A-4 and *wolf* is an NP-0, so *big wolf* is an NP-4, since 4 is higher than 0.

# A Grammar for Modifiers III: encoded hierarchy

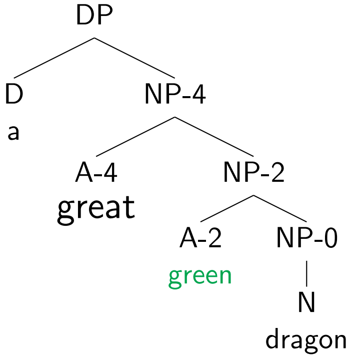
**Rules for making DPs:**

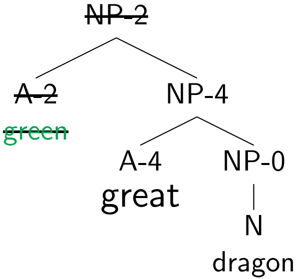
1. Suppose we have a number *i*. If you have an NP-i, you can put a D before it to make a DP

*big wolf* is an NP-4 and *a* is a D, so *a big wolf* is a DP.

# A Grammar for Modifiers III: encoded hierarchy

**Tolkien's problem:**





# A Grammar for Modifiers III: encoded hierarchy

**Problems we've solved:**

1. You can have lots of modifiers on one noun (*good heavy red* table)
2. Some orders just sound weird (**pine old* floorboards)
3. Sometimes switching the order changes nothing (*beautiful cute, cute beautiful* baby)
4. Some languages allow adjectives to repeat (*my love is like a red red rose*)
5. Modifiers are nearly always optional (*my love is like a rose*)
6. Modifiers can themselves be modified (*a surprisingly short basketball player* )
7. Some things can be both modifiers and "arguments" (*They seem nice*, *He put it on the table*)

But we only have 3 rules, just like Take 1!

# Computational complexity

Our third grammar is significantly smaller than our second grammar. For $k$ hierarchies, the size of the lexicon is bounded below by a polynomial function of the depths of the hierarchies as follows. For $k$ = number of hierarchies in Lex, $l_i$ = number of levels in hierarchy $i$, $x_i$ = number of LIs at level 0 of hierarchy $i$ (by Gauss's function for adding sequences):

$$|\mathrm{Lex}'| \geq \sum_{i=1}^{k} 1/2(l_i^2 + l_i) + x_i$$

- That is, the size of the lexicon in encoded order is a polynomial function of the depths of the hierarchies.
- The size of the lexicon in encoded hierarchy is linear in the depths of the hierarchies.

# A Grammar for Modifiers III: encoded hierarchy

**Problems we can extend this approach to solve:**

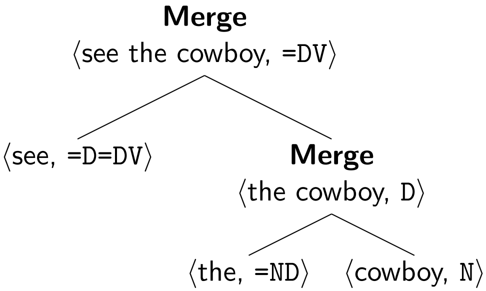
1. Some orders change the meaning (*blistered varnished wood* = varnished wood that got blistered, *varnished blistered wood* = wood that was blistered and then varnished
2. Prepositional phrases and relative clauses come after the noun (dog *with long ears*, baby *who cried the whole flight*)
3. Prepositional phrases and relative clauses are *not* strictly ordered. (*A linguist with tin ear at a conference, A linguist at a conference with a tin ear*)

# De-simplifying

- These are *context-free grammars*
- Context free isn't powerful enough to model human languages
- →Embed this solution in a *Minimalist Grammar*

# Minimalist Grammars

## Example: Merge



**Merge**

⟨see the cowboy, =DV⟩

⟨see, =D=DV⟩

**Merge**

⟨the cowboy, D⟩

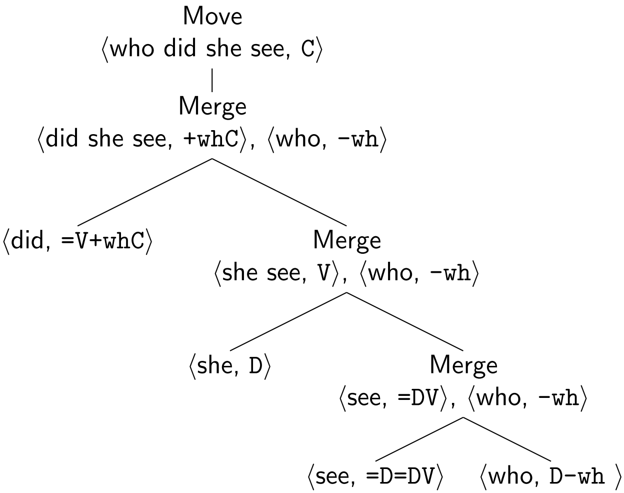
⟨the, =ND⟩

⟨cowboy, N⟩

(Stabler, 1997; Chomsky, 1995)

# Minimalist Grammars

## Example: Move



Move

⟨who did she see, C⟩

Merge

⟨did she see, +whC⟩, ⟨who, -wh⟩

⟨did, =V+whC⟩

Merge

⟨she see, V⟩, ⟨who, -wh⟩

⟨she, D⟩

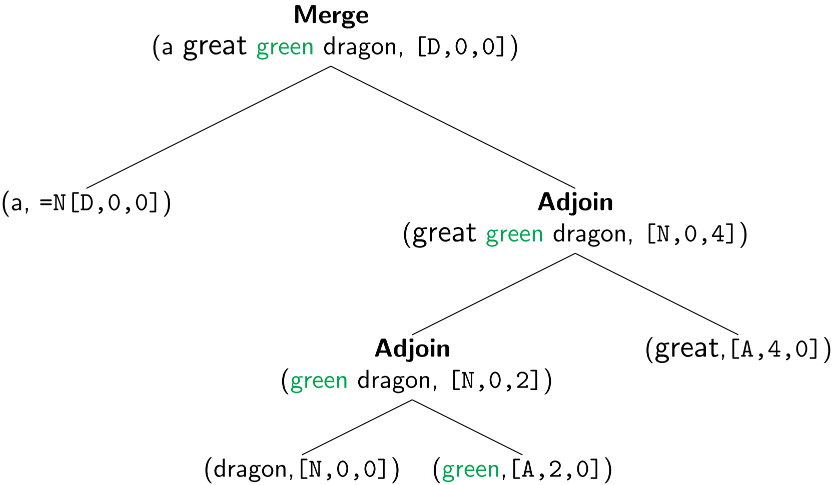
Merge

⟨see, =DV⟩, ⟨who, -wh⟩

⟨see, =D=DV⟩

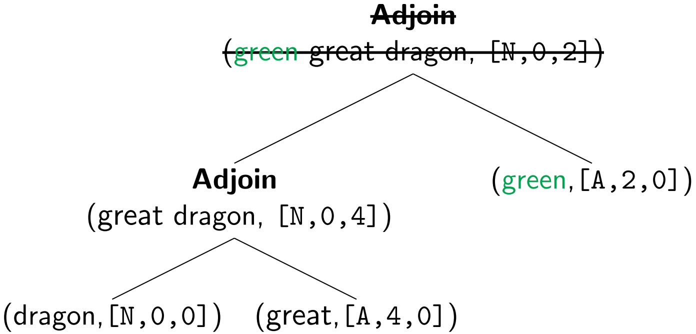
⟨who, D-wh ⟩

# Minimalist Grammars with Adjunction: Tolkien's problem



Fowlie (2014)

# Minimalist Grammars with Adjunction: Tolkien's problem



~~**Adjoin**~~

~~(green great dragon, [N,0,2])~~

**Adjoin**

(great dragon, [N,0,4])

(green, [A,2,0])

(dragon, [N,0,0])

(great, [A,4,0])

Fowlie (2014)

# Minimalist Grammars with Adjunction

Code interlude...

# Computation

**Advantages of a computational approach:**

- Telling a computer our theory forces us to think carefully about our claims
- We can use the computer to check our work
- Once we've defined our ideas mathematically, we can use what is already known about math to learn new things about our ideas
- Universals of human language almost certainly include mathematical properties of grammars – in fact, these could end up being the only universals
- With a carefully defined grammar, we can set a computer to parse for us
- We can define learning algorithms to try to model how babies might learn language

# Future/ongoing research

- *How do people learn adjuncts?* → Ongoing artificial language learning experiments
- *What are the mathematical properties of MGAs?* → Not all the same as MGs!
- *Can songbirds have human-like grammars?* → Collaborative work with linguists, CS, biologists

# Conclusion

- Adjectives have a default ordering
- Semantic classes of adjectives are ordered (fairly) universally across languages
- Proposal: this order is built directly into the grammar – perhaps it is even somehow built into our brains, so that babies don't even have to learn it
- Modelling adjuncts accurately in Minimalist Grammars allows for tractable parsing, regular derivation tree languages
- Thinking computationally forces clarity and explicitness in syntactic research; syntax will very soon become more useful in industry as the tasks being tackled become harder and more localised

# References

Carpenter, Humphrey. 1981. *The letters of JRR Tolkien*, volume 144. Boston: Houghton Mifflin.

Chomsky, Noam. 1995. *The minimalist program*. Cambridge, MA: MIT Press.

Cinque, Gugliemo. 1999. *Adverbs and functional heads: a cross-linguistic perspective*. Oxford studies in comparative syntax. Oxford: Oxford University Press.

Fowlie, Meaghan. 2014. Adjuncts and minimalist grammars. In *Proceedings of Formal Grammar 2014*, ed. Glyn Morrill, Reinhard Muskens, Rainer Osswald, and Frank Richter, volume 8612 of *Lecture Notes in Computer Science*.

Scott, Gary-John. 2002. Stacked adjectival modification and the structure of nominal phrases. In *Functional structure in DP and IP: The cartography of syntactic structures*, volume 1, 91–120. Oxford University Press.

Stabler, Edward. 1997. Derivational minimalism. *Logical Aspects of Computational Linguistics* 68–95.

# Thank you!