

Culbertson & Adger (2014)

'Language learners privilege structured meaning over surface frequency'

Suzana Fong

sznfong@mit.edu

Recitation #2

September 14, 2018

- **Examine key concepts of the hypothesis supported in the paper**
 - ▶ Structure dependency
 - ▶ Scope
 - ▶ Universal 20
- Along the way, I will make occasional comments on the **structure and argumentation** of the paper.

The paper

- The main question that the paper tries to address
 - ▶ What type of knowledge do learners use when acquiring a language?
- Two hypotheses that the paper compares
 - ▶ Structure dependency: learners rely on structure to learn a language
 - ▶ Statistical generalization: learners rely on the frequency of word order
- Experiment 1 to distinguish between the two hypotheses¹

¹Experiment 2 not discussed because of time constraints.

The paper

- The main question that the paper tries to address
 - ▶ What type of knowledge do learners use when acquiring a language?
- Two hypotheses that the paper compares
 - ▶ Structure dependency: learners rely on structure to learn a language
 - ▶ Statistical generalization: learners rely on the frequency of word order
- Experiment 1 to distinguish between the two hypotheses¹

¹Experiment 2 not discussed because of time constraints.

The paper

- The main question that the paper tries to address
 - ▶ What type of knowledge do learners use when acquiring a language?
- Two hypotheses that the paper compares
 - ▶ Structure dependency: learners rely on structure to learn a language
 - ▶ Statistical generalization: learners rely on the frequency of word order
- Experiment 1 to distinguish between the two hypotheses¹

¹Experiment 2 not discussed because of time constraints.

The paper

- The main question that the paper tries to address
 - ▶ What type of knowledge do learners use when acquiring a language?
- Two hypotheses that the paper compares
 - ▶ Structure dependency: learners rely on structure to learn a language
 - ▶ Statistical generalization: learners rely on the frequency of word order
- Experiment 1 to distinguish between the two hypotheses¹

¹Experiment 2 not discussed because of time constraints.

The paper

- The main question that the paper tries to address
 - ▶ What type of knowledge do learners use when acquiring a language?
- Two hypotheses that the paper compares
 - ▶ Structure dependency: learners rely on structure to learn a language
 - ▶ Statistical generalization: learners rely on the frequency of word order
- Experiment 1 to distinguish between the two hypotheses¹

¹Experiment 2 not discussed because of time constraints.

The paper

- The main question that the paper tries to address
 - ▶ What type of knowledge do learners use when acquiring a language?
- Two hypotheses that the paper compares
 - ▶ Structure dependency: learners rely on structure to learn a language
 - ▶ Statistical generalization: learners rely on the frequency of word order
- Experiment 1 to distinguish between the two hypotheses¹

¹Experiment 2 not discussed because of time constraints.

The paper

- The main question that the paper tries to address
 - ▶ What type of knowledge do learners use when acquiring a language?
- Two hypotheses that the paper compares
 - ▶ **Structure dependency**: learners rely on structure to learn a language
 - ▶ Statistical generalization: learners rely on the frequency of word order
- Experiment 1 to distinguish between the two hypotheses¹

¹Experiment 2 not discussed because of time constraints.

Key concepts

Structure dependency

- Yes/No question formation in English:
 - (1) a. Alex **has** been swimming a lot lately.
b. **Has** Alex been swimming recently?
 - (2) a. Celeste **will** bring pie to the party.
b. **Will** Celeste bring pie to the party?
- First-pass description of how these questions are formed:

Structure dependency

- Yes/No question formation in English:
 - (1) a. Alex **has** been swimming a lot lately.
b. **Has** Alex been swimming recently?
 - (2) a. Celeste **will** bring pie to the party.
b. **Will** Celeste bring pie to the party?
- First-pass description of how these questions are formed:

Structure dependency

- Yes/No question formation in English:
 - (1) a. Alex **has** been swimming a lot lately.
b. **Has** Alex been swimming recently?
 - (2) a. Celeste **will** bring pie to the party.
b. **Will** Celeste bring pie to the party?
- First-pass description of how these questions are formed:
 - (3) *Move the first auxiliary to the right of the subject to its left*
Alex **has** been swimming a lot lately

Structure dependency

- Yes/No question formation in English:
 - (1) a. Alex **has** been swimming a lot lately.
b. **Has** Alex been swimming recently?
 - (2) a. Celeste **will** bring pie to the party.
b. **Will** Celeste bring pie to the party?
- First-pass description of how these questions are formed:
 - (3) *Move the first auxiliary to the right of the subject to its left*
Has Alex **has** been swimming a lot lately

- Let's apply this rule to the following sentence:

(4) The student who is running has been swimming a lot lately.

- Result:

- Let's apply this rule to the following sentence:

(4) The student who is running has been swimming a lot lately.

- Result:

(5) the student who **is** running has been swimming a lot lately

- Let's apply this rule to the following sentence:

(4) The student who is running has been swimming a lot lately.

- Result:

(5) **Is** the student who **is** running has been swimming a lot lately

- Let's apply this rule to the following sentence:

(4) The student who is running has been swimming a lot lately.

- Result:

(5) * **Is** the student who **is** running has been swimming a lot lately

In view of this data point, what would be a better alternative?

In view of this data point, what would be a better alternative?

- *Move the auxiliary in the main clause (not in the relative clause) to the left of the subject*

In view of this data point, what would be a better alternative?

- *Move the auxiliary in the main clause (not in the relative clause) to the left of the subject*

(6) [the student who is running] has been swimming a lot lately

In view of this data point, what would be a better alternative?

- *Move the auxiliary in the main clause (not in the relative clause) to the left of the subject*

(6) [the student who is running] **has** been swimming a lot lately

In view of this data point, what would be a better alternative?

- *Move the auxiliary in the main clause (not in the relative clause) to the left of the subject*

(6) **Has** [the student who is running] **has** been swimming a lot lately

- What was C&A's point when opening the paper with this type of data and discussion?
- Useful way to frame the discussion
 - ▶ Auxiliary inversion is not the topic of the paper.
 - ▶ But it gives the reader a broad picture of what structure dependency is.
 - ▶ Furthermore, it helps argue that structure dependency is a general property of language.

- What was C&A's point when opening the paper with this type of data and discussion?
 - ▶ To show that language rules are based on **structure** (*main/subordinate clause*) rather than just linear order (*first auxiliary/left or right*)
- Useful way to frame the discussion
 - ▶ Auxiliary inversion is not the topic of the paper.
 - ▶ But it gives the reader a broad picture of what structure dependency is.
 - ▶ Furthermore, it helps argue that structure dependency is a general property of language.

- What was C&A's point when opening the paper with this type of data and discussion?
 - ▶ To show that language rules are based on **structure** (*main/subordinate clause*) rather than just linear order (*first auxiliary/left or right*)
- Useful way to frame the discussion
 - ▶ Auxiliary inversion is not the topic of the paper.
 - ▶ But it gives the reader a broad picture of what structure dependency is.
 - ▶ Furthermore, it helps argue that structure dependency is a general property of language.

- Assumption: *the order words are combined with each other matters to interpretation.*
- How this can be shown: ambiguous phrases
 - (7) a pretty friendly dog
- This can be represented with brackets or trees:

- Assumption: *the order words are combined with each other matters to interpretation.*
- How this can be shown: ambiguous phrases
 - (7) a pretty friendly dog
- This can be represented with brackets or trees:

- Assumption: *the order words are combined with each other matters to interpretation.*
- How this can be shown: ambiguous phrases
 - (7) a pretty friendly dog
 - a. *a dog who is pretty friendly*
 - b. *a friendly dog who is pretty*
- This can be represented with brackets or trees:

- Assumption: *the order words are combined with each other matters to interpretation.*
- How this can be shown: ambiguous phrases
 - (7) a pretty friendly dog
 - a. *a dog who is pretty friendly*
 - b. *a friendly dog who is pretty*
- This can be represented with brackets or trees:

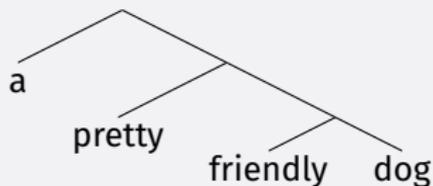
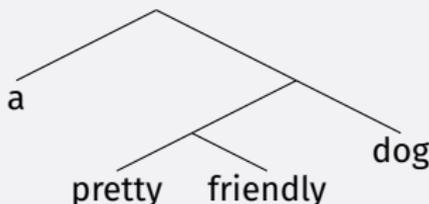
Scope

- Assumption: *the order words are combined with each other matters to interpretation.*
- How this can be shown: ambiguous phrases

- (7) a pretty friendly dog
- a. *a dog who is pretty friendly*
 - b. *a friendly dog who is pretty*

- This can be represented with brackets or trees:

(7')



This is what C&A mean by *scope*:

- It has to do with how one word affects another and how the order of combination affects the interpretation.²
- Recall our example:

- (7) a pretty friendly dog
- a. *a dog who is pretty friendly*
 - b. *a friendly dog who is pretty*

- ▶ *Pretty* takes scope over *friendly* in (7a).
- ▶ *Pretty* takes scope over [*friendly dog*] in (7b).

²Note that you may encounter different uses of the word in the linguistics literature.

This is what C&A mean by *scope*:

- It has to do with how one word affects another and how the order of combination affects the interpretation.²
- Recall our example:

- (7) a pretty friendly dog
- a. *a dog who is pretty friendly*
 - b. *a friendly dog who is pretty*

- ▶ *Pretty* takes scope over *friendly* in (7a).
- ▶ *Pretty* takes scope over [*friendly dog*] in (7b).

²Note that you may encounter different uses of the word in the linguistics literature.

This is what C&A mean by *scope*:

- It has to do with how one word affects another and how the order of combination affects the interpretation.²
- Recall our example:

- (7) a pretty friendly dog
- a. *a dog who is pretty friendly*
 - b. *a friendly dog who is pretty*

- ▶ *Pretty* takes scope over *friendly* in (7a).
- ▶ *Pretty* takes scope over [*friendly dog*] in (7b).

²Note that you may encounter different uses of the word in the linguistics literature.

Universal 20: a generalization about attested word orders in natural languages, among all the logic possibilities

(8) **Prenominal modifiers** (e.g. English)

| | | | |
|---------------|------------|------------|-------------|
| Demonstrative | Numeral | Adjective | Noun |
| <i>these</i> | <i>two</i> | <i>red</i> | <i>cars</i> |

(9) **Postnominal modifiers** (e.g. Thai)

Most frequent order:

| | | | |
|-------------|------------|------------|---------------|
| Noun | Adjective | Numeral | Demonstrative |
| <i>cars</i> | <i>red</i> | <i>two</i> | <i>these</i> |

Why should Universal 20 be true?

- Explanation on p. 5843, second column, paragraph that starts with “There are three converging sources of evidence...”
- This is a very dense paragraph, packed with technical information that may not be understandable by every reader.³
 - ▶ **Suggestion:** make your assumptions explicit and unpack your information.

³Of course, this is not the main focus of the paper and there were probably space limitations.

Why should Universal 20 be true?

There are rules about how demonstratives, numerals, and adjectives are composed together.

- C&A: semantic type constraints (set theory/functions)
- Here: rules of composition (simplification!)
 1. Adjectives, numerals, and demonstratives can each combine with a noun directly.
 2. Adjectives must be the closest element to the noun.
 3. Numerals can also combine with a noun that is already combined with an adjective
 4. Demonstratives can also combine with a noun that is already combined with a numeral

Why should Universal 20 be true?

There are rules about how demonstratives, numerals, and adjectives are composed together.

- C&A: semantic type constraints (set theory/functions)
- Here: rules of composition (simplification!)
 1. Adjectives, numerals, and demonstratives can each combine with a noun directly.
 2. Adjectives must be the closest element to the noun.
 3. Numerals can also combine with a noun that is already combined with an adjective
 4. Demonstratives can also combine with a noun that is already combined with a numeral

→ *Universal 20 is true because the order of modifiers reflects the rules that modifiers obey when combining with a noun and with each other.*

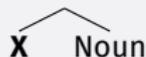
Structure (scope) vs. linear order

- Recall from (7) *a pretty friendly dog*:
 - ▶ the order of combination between words has consequences for scope
 - ▶ this can be represented with syntactic trees
- Let's try to draw trees for our rules.

1. Adjectives, numerals, and demonstratives can each combine with a noun directly.
2. Adjectives must be the closest element to the noun.
3. Numerals can also combine with a noun that is already combined with an adjective
4. Demonstratives can also combine with a noun that is already combined with a numeral

1. Adjectives, numerals, and demonstratives can each combine with a noun directly.

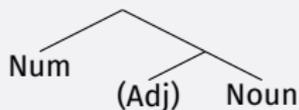
(10)



Where X is: adjective, numeral, or demonstrative

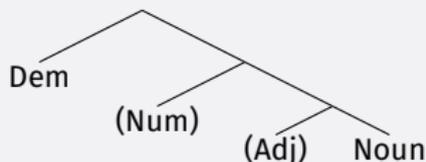
2. Adjectives must be the closest element to the noun.
3. Numerals can also combine with a noun that is already combined with an adjective

(11)

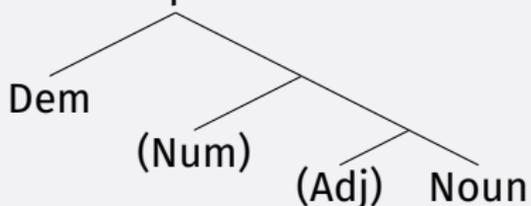


4. Demonstratives can also combine with a noun that is already combined with a numeral

(12)



In terms of C&A's scope:



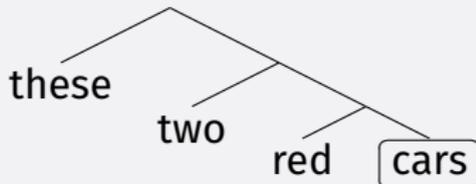
- The adjective scopes over the noun.
- The numeral scopes over the adjective.
- The demonstrative scopes over the numeral.

The rules were illustrated with the first part of Universal 20:

(8) **Prenominal modifiers** (e.g. English)

| | | | |
|---------------|------------|------------|---|
| Demonstrative | Numeral | Adjective | Noun |
| <i>these</i> | <i>two</i> | <i>red</i> | <i>cars</i> |

(13)



■ What about the second part of Universal 20?

(9) **Postnominal modifiers** (e.g. Thai)

Most frequent order:

| Noun | Adjective | Numeral | Demonstrative |
|--|------------|------------|---------------|
| cars | <i>red</i> | <i>two</i> | <i>these</i> |

■ Let's make the following working assumptions:

- ▶ Our rules to compose nouns with Adj, Num, and Dem, are universal.
- ▶ Languages differ in 'directionality':

(14)



- Based on these assumptions, let's draw a tree for (9) that is analogous to the English tree (13) in the previous slide.

■ What about the second part of Universal 20?

(9) **Postnominal modifiers** (e.g. Thai)

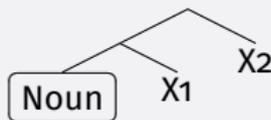
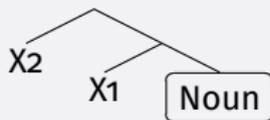
Most frequent order:

| | | | |
|--|------------|------------|---------------|
| Noun | Adjective | Numeral | Demonstrative |
| cars | <i>red</i> | <i>two</i> | <i>these</i> |

■ Let's make the following working assumptions:

- ▶ Our rules to compose nouns with Adj, Num, and Dem, are universal.
- ▶ Languages differ in 'directionality':

(14)



- Based on these assumptions, let's draw a tree for (9) that is analogous to the English tree (13) in the previous slide.

- What about the second part of Universal 20?

(9) **Postnominal modifiers** (e.g. Thai)

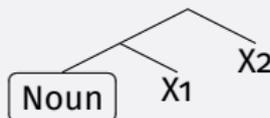
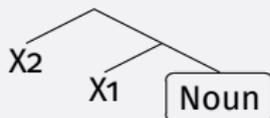
Most frequent order:

| | | | |
|--|------------|------------|---------------|
| Noun | Adjective | Numeral | Demonstrative |
| cars | <i>red</i> | <i>two</i> | <i>these</i> |

- Let's make the following working assumptions:

- ▶ Our rules to compose nouns with Adj, Num, and Dem, are universal.
- ▶ Languages differ in 'directionality':

(14)



- Based on these assumptions, let's draw a tree for (9) that is analogous to the English tree (13) in the previous slide.

1. Adjectives, numerals, and demonstratives can each combine with a noun directly.
2. Adjectives must be the closest element to the noun.
3. Numerals can also combine with a noun that is already combined with an adjective
4. Demonstratives can also combine with a noun that is already combined with a numeral

1. Adjectives, numerals, and demonstratives can each combine with a noun directly.

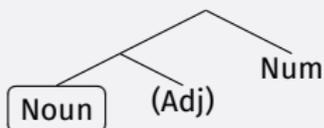
(15)



Where X is: adjective, numeral, or demonstrative

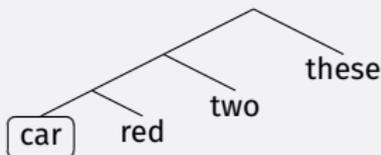
2. Adjectives must be the closest element to the noun.
3. Numerals can also combine with a noun that is already combined with an adjective

(16)

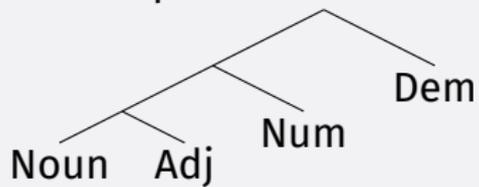


4. Demonstratives can also combine with a noun that is already combined with a numeral

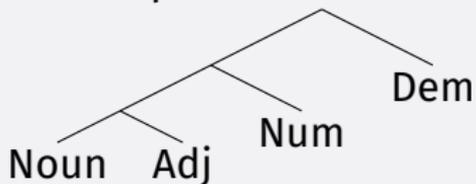
(17)



In terms of C&A's scope:



In terms of C&A's scope:



- The adjective scopes over the noun.
- The numeral scopes over the adjective.
- The demonstrative scopes over the numeral.

Upshot

Universal 20 can be explained with *(i)* a set of rules to combine nouns and their modifiers, coupled with *(ii)* a way to encode differences between languages that is based on directionality

Experiment 1

- If this is true, then could these rules also guide language learning?
- Alternatively, maybe learners just rely on the frequency of word orders.

- **Structure dependency hypothesis:** If this is true, then could these rules also guide language learning?
- Alternatively, maybe learners just rely on the frequency of word orders.

- **Structure dependency hypothesis:** If this is true, then could these rules also guide language learning?
- **Statistical generalization hypothesis:** Alternatively, maybe learners just rely on the frequency of word orders.

- The subjects that participated in the experiment were shown data from a made-up language, with English glosses.
- In this language, modifiers follow nouns, unlike what happens in English.
- They were shown simplex *Noun + Modifier* combinations.

| | | |
|------|---------------|-------------------|
| (18) | a. cars red | <i>Noun + Adj</i> |
| | b. cars these | <i>Noun + Dem</i> |

- **Task:** infer complex *Noun + Modifier* combinations from the simplex data.

- The subjects that participated in the experiment were shown data from a made-up language, with English glosses.
- In this language, modifiers follow nouns, unlike what happens in English.
- They were shown simplex *Noun + Modifier* combinations.

(18) a. cars red *Noun + Adj*
b. cars these *Noun + Dem*

- **Task:** infer complex *Noun + Modifier* combinations from the simplex data.

(19) i. cars these red
ii. cars red these

Structure dependency hypothesis:

- If learners are guided by a rule of composition that says that Dem (*these*) scope over Adj (*red*), what is the predicted order between Dem (*these*) and Adj *red*?

Structure dependency hypothesis:

- If learners are guided by a rule of composition that says that Dem (*these*) scope over Adj (*red*), what is the predicted order between Dem (*these*) and Adj *red*?

(20) cars red these

Statistical generalization hypothesis:

- Fact needed: statistical data about word order in English
 - ▶ More occurrences of Dem + Adj (*these + red*) than of Adj + Dem (*red + these*)
- If learners rely on statistics, what is the predicted order between Dem (*these*) and Adj (*red*)?

Statistical generalization hypothesis:

- Fact needed: statistical data about word order in English
 - ▶ More occurrences of Dem + Adj (*these* + *red*) than of Adj + Dem (*red* + *these*)
- If learners rely on statistics, what is the predicted order between Dem (*these*) and Adj (*red*)?

Statistical generalization hypothesis:

- Fact needed: statistical data about word order in English
 - ▶ More occurrences of Dem + Adj (*these* + *red*) than of Adj + Dem (*red* + *these*)
- If learners rely on statistics, what is the predicted order between Dem (*these*) and Adj (*red*)?

(21) cars these red

Results

- The prediction made by the structure dependency hypothesis proved to be the correct one.⁴
- Very interesting contribution of the paper: experimental validation of Universal 20.

⁴C&A emphasize though that this doesn't mean that statistical generalization is never used.

The paper

Having covered part of the necessary background to understand the paper, **let's think about its structure.**

1. Who do you think the target audience is? How do you know?
2. Do you think the paper successfully communicates with its target audience? Think about this considering Pinker's 'curse of knowledge' chapter.
3. How well do you think the paper highlights take-home messages (interim summaries) and informs the reader of what is coming up (signposting)?