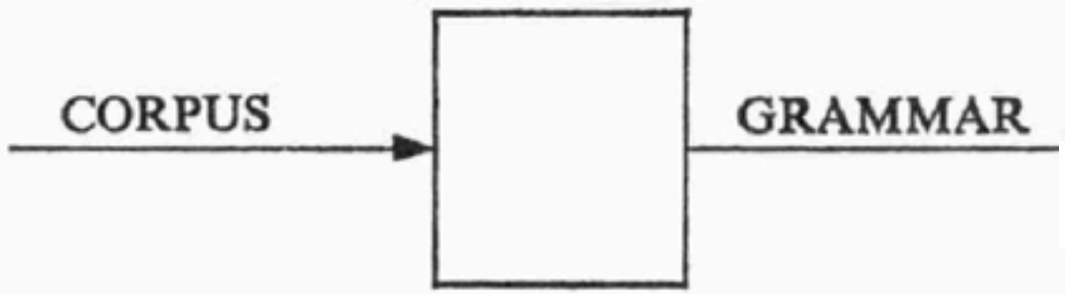
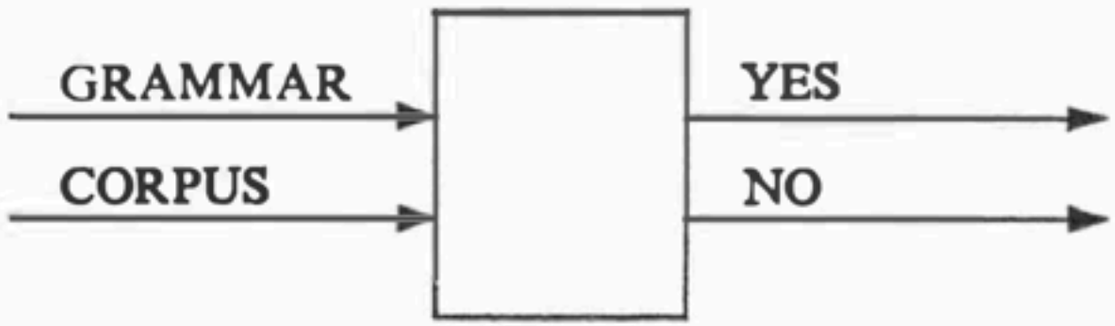
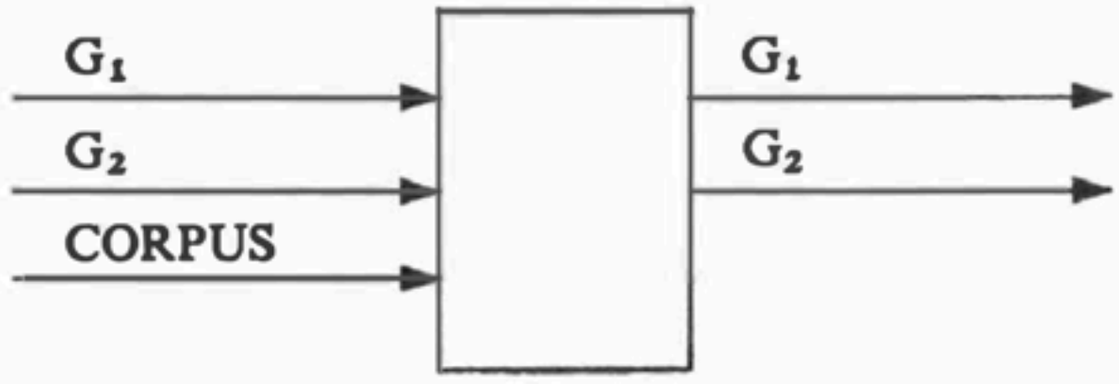
Productivity and the Discovery Procedure

Charles Yang

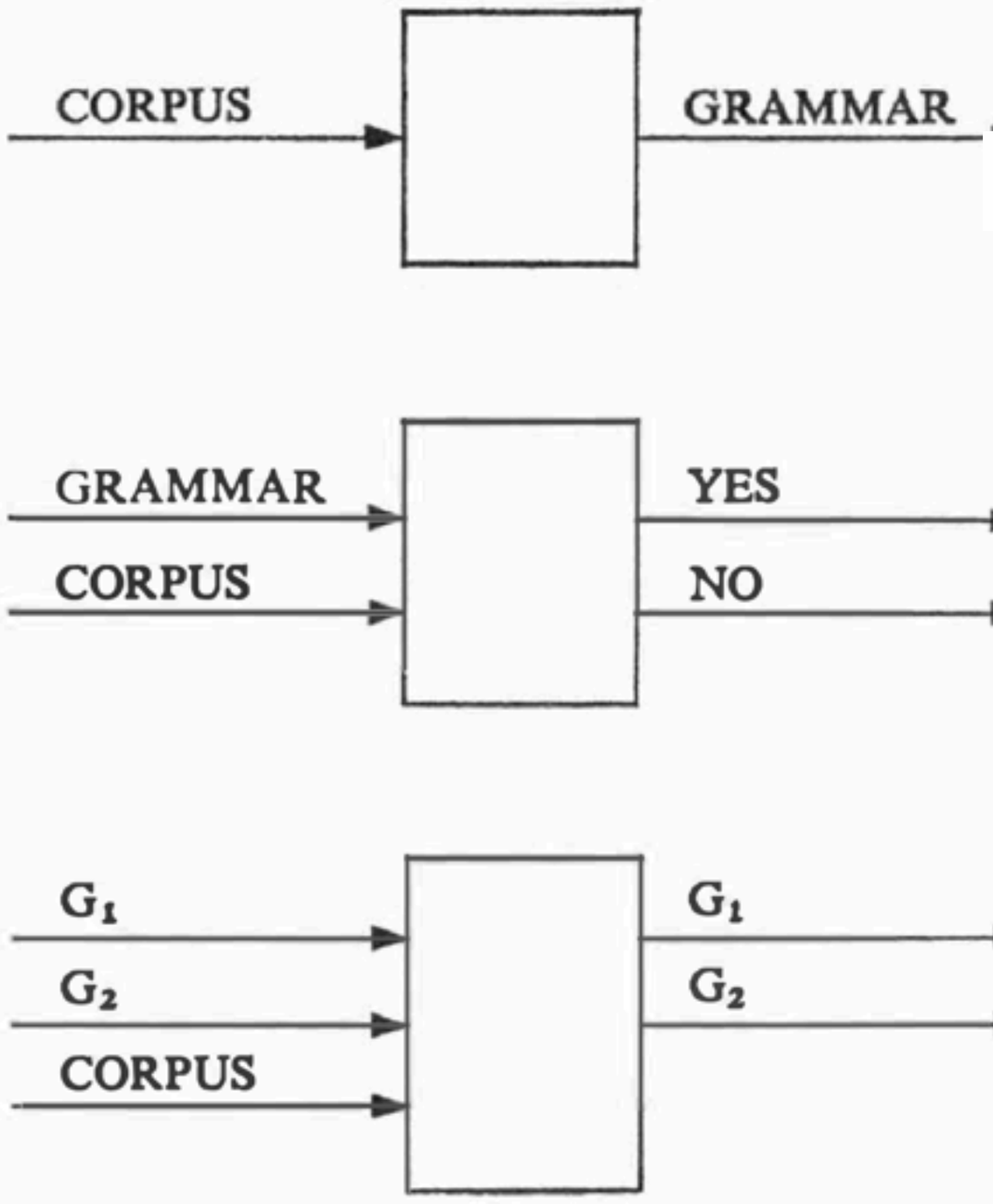
CRISSP 2021





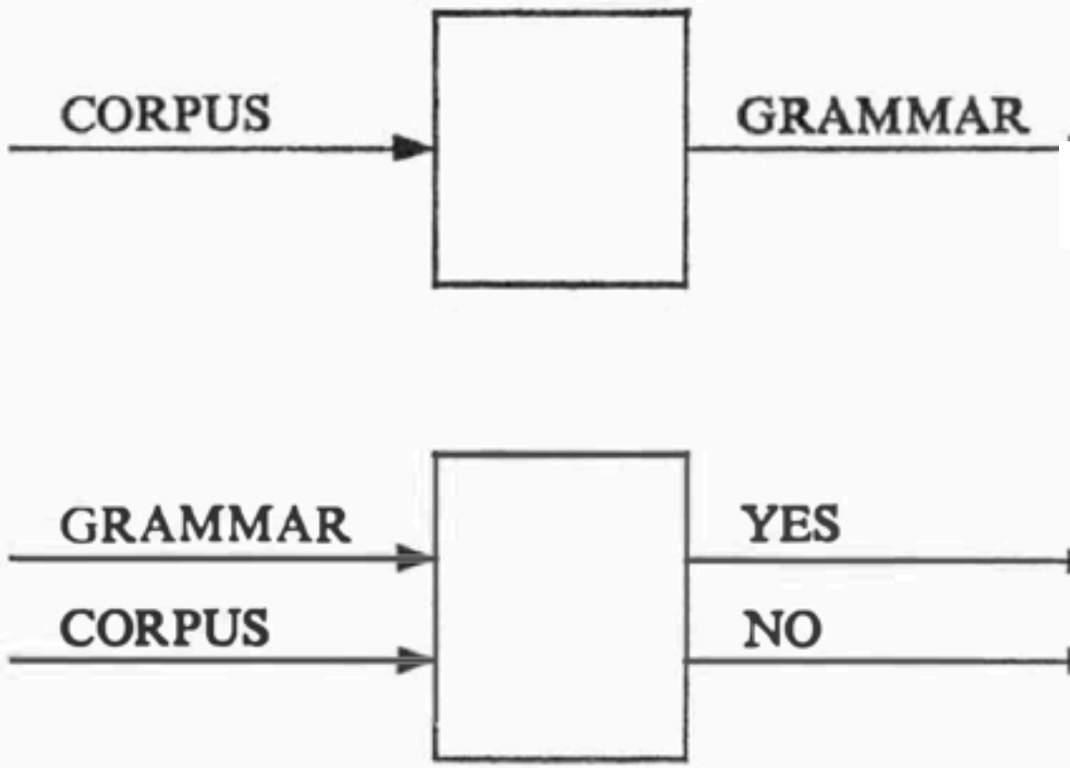


The strongest requirement that could be placed on the relation between a theory of linguistic structure and particular grammars is that the theory must provide a practical and mechanical method for actually constructing the grammar, given a corpus of utterances. Let us say that such a theory provides us with a *discovery procedure*



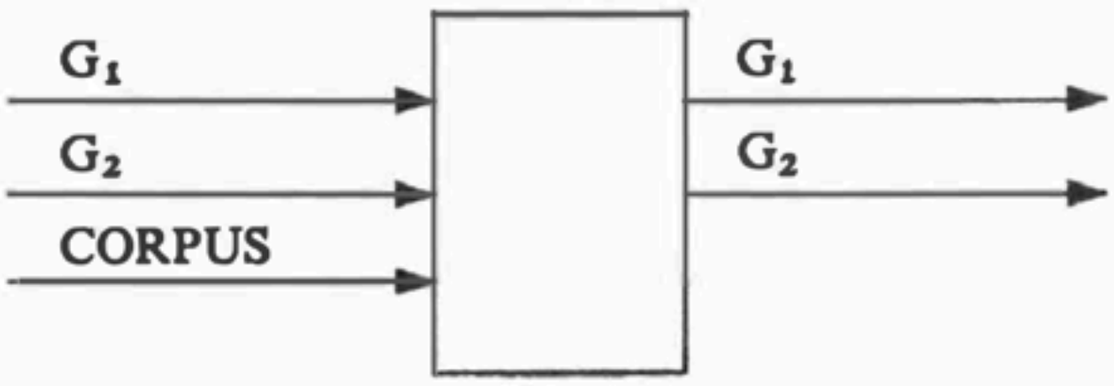
The strongest requirement that could be placed on the relation between a theory of linguistic structure and particular grammars is that the theory must provide a practical and mechanical method for actually constructing the grammar, given a corpus of utterances. Let us say that such a theory provides us with a *discovery procedure*

I think that it is very questionable that this goal is attainable in any interesting way,

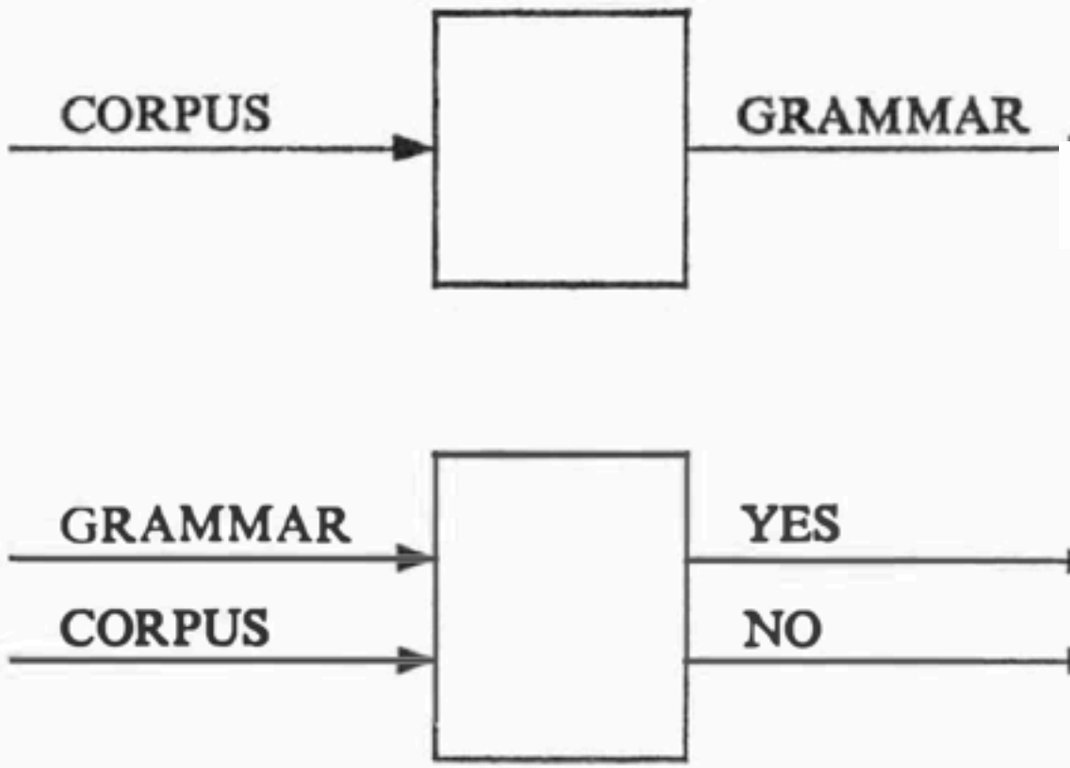


The strongest requirement that could be placed on the relation between a theory of linguistic structure and particular grammars is that the theory must provide a practical and mechanical method for actually constructing the grammar, given a corpus of utterances. Let us say that such a theory provides us with a *discovery procedure*

A weaker requirement would be that the theory must provide a practical and mechanical method for determining whether or not a grammar proposed for a given corpus is, in fact, the best grammar of the language from which this corpus is drawn. Such a theory, which is not concerned with the question of *how* this grammar was constructed, might be said to provide a *decision procedure* for grammars.

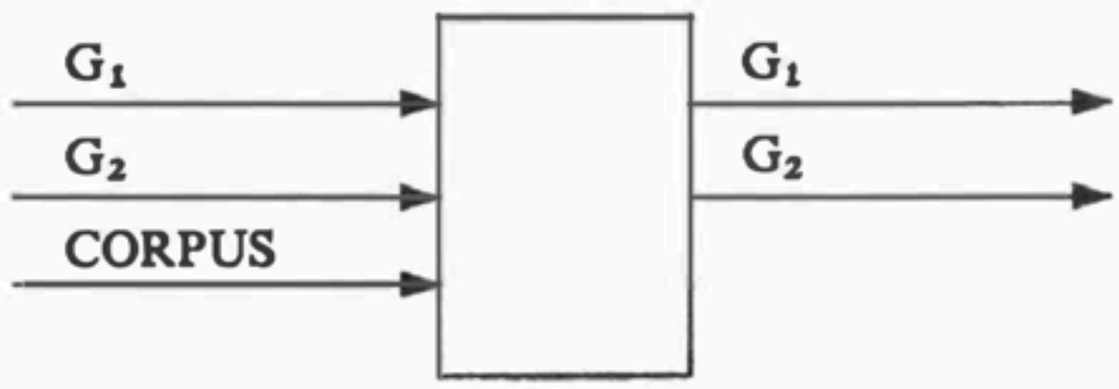


I think that it is very questionable that this goal is attainable in any interesting way,



The strongest requirement that could be placed on the relation between a theory of linguistic structure and particular grammars is that the theory must provide a practical and mechanical method for actually constructing the grammar, given a corpus of utterances. Let us say that such a theory provides us with a *discovery procedure*

A weaker requirement would be that the theory must provide a practical and mechanical method for determining whether or not a grammar proposed for a given corpus is, in fact, the best grammar of the language from which this corpus is drawn. Such a theory, which is not concerned with the question of *how* this grammar was constructed, might be said to provide a *decision procedure* for grammars.



An even weaker requirement would be that given a corpus and given two proposed grammars G_1 and G_2, the theory must tell us which is the better grammar of the language from which the corpus is drawn. In this case we might say that the theory provides an *evaluation procedure* for grammars.

I think that it is very questionable that this goal is attainable in any interesting way,

From Evaluation to Discovery

From Evaluation to Discovery

- Evaluation: UG proposes (from biology), data disposes

From Evaluation to Discovery

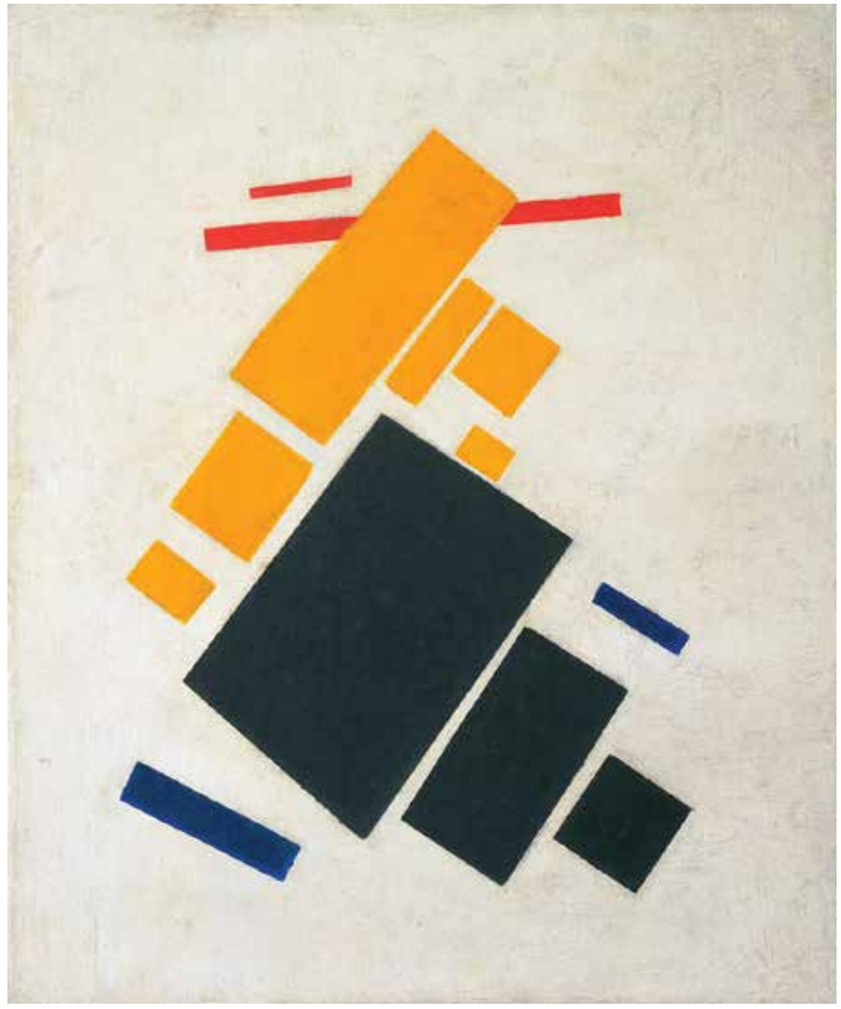
- Evaluation: UG proposes (from biology), data disposes
- Discovery: Child proposes (from data), data disposes

From Evaluation to Discovery

- Evaluation: UG proposes (from biology), data disposes
- Discovery: Child proposes (from data), data disposes
- “Having selected a permissible hypothesis, he [the child] can use inductive evidence for corrective action, confirming or disconfirming his choice. Once the hypothesis is sufficiently well confirmed, the child knows the language defined by this hypothesis; consequently, his knowledge extends enormously beyond his experience.”

From Evaluation to Discovery

- Evaluation: UG proposes (from biology), data disposes
- Discovery: Child proposes (from data), data disposes
- “Having selected a permissible hypothesis, he [the child] can use inductive evidence for corrective action, confirming or disconfirming his choice. Once the hypothesis is sufficiently well confirmed, the child knows the language defined by this hypothesis; consequently, his knowledge extends enormously beyond his experience.”

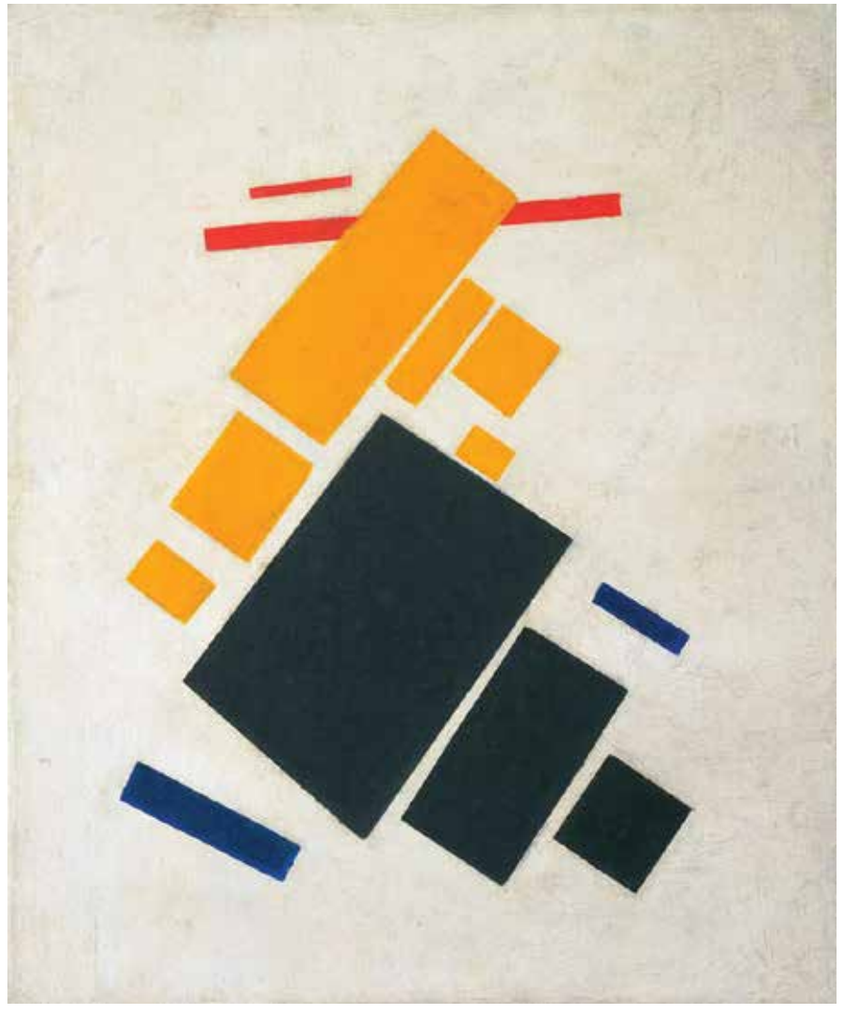


HOW CHILDREN LEARN TO BREAK THE RULES OF LANGUAGE

THE PRICE OF LINGUISTIC PRODUCTIVITY

CHARLES YANG

- Give a set of items:
  - If *many* do X, then all do X
  - if *few* do X, then remember the few that do

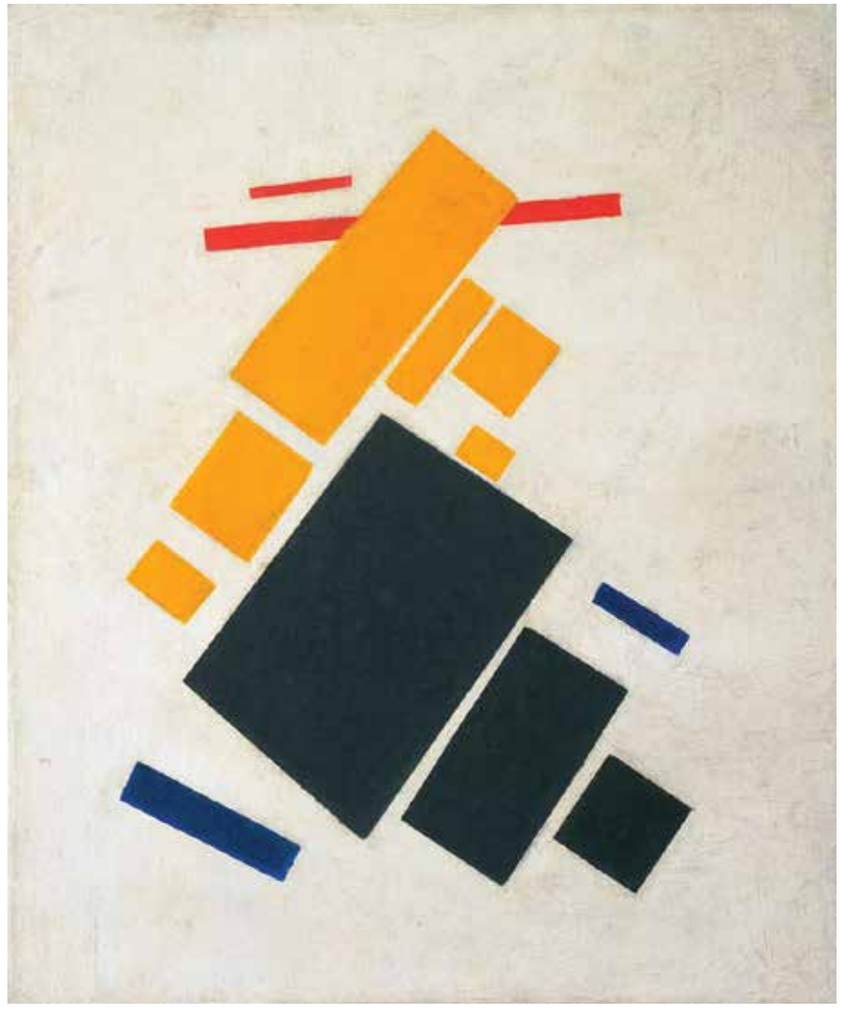


HOW CHILDREN LEARN TO BREAK THE RULES OF LANGUAGE

THE PRICE OF LINGUISTIC PRODUCTIVITY

CHARLES YANG

- Give a set of items:
  - If *many* do X, then all do X
  - if *few* do X, then remember the few that do
- How many is *many* or *few*?

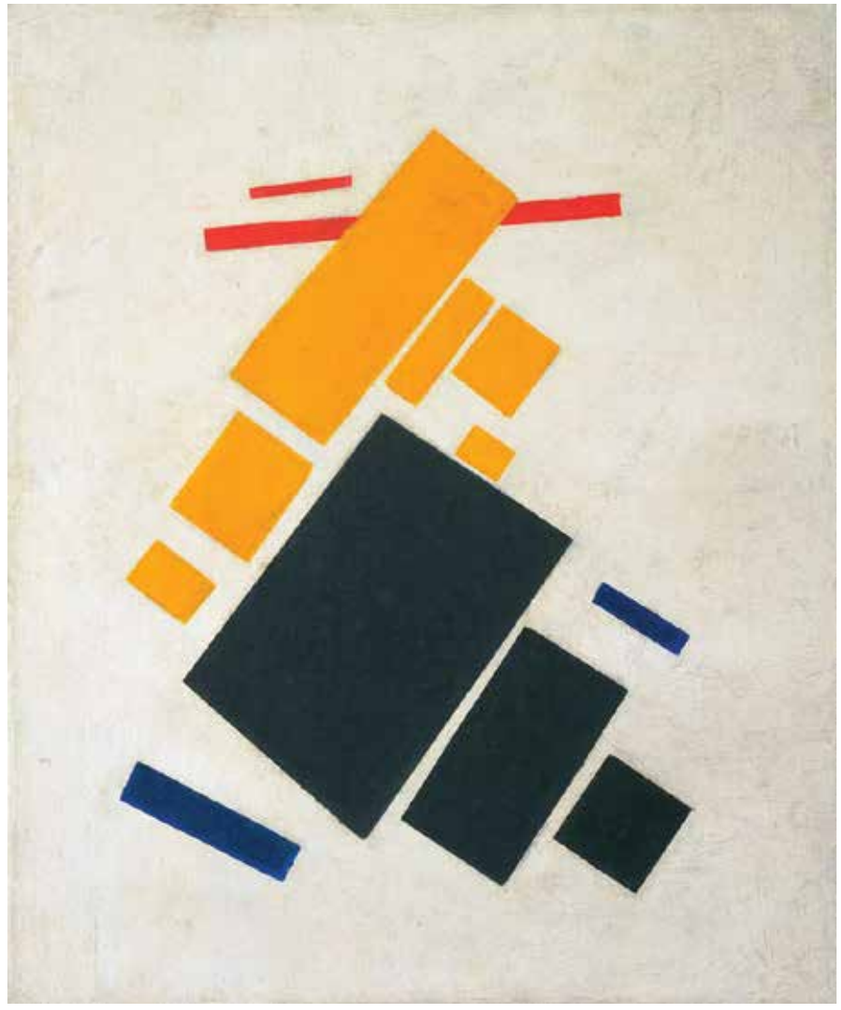


HOW CHILDREN LEARN TO BREAK THE RULES OF LANGUAGE

THE PRICE OF LINGUISTIC PRODUCTIVITY

CHARLES YANG

- Give a set of items:
  - If *many* do X, then all do X
  - if *few* do X, then remember the few that do
- How many is *many* or *few*?



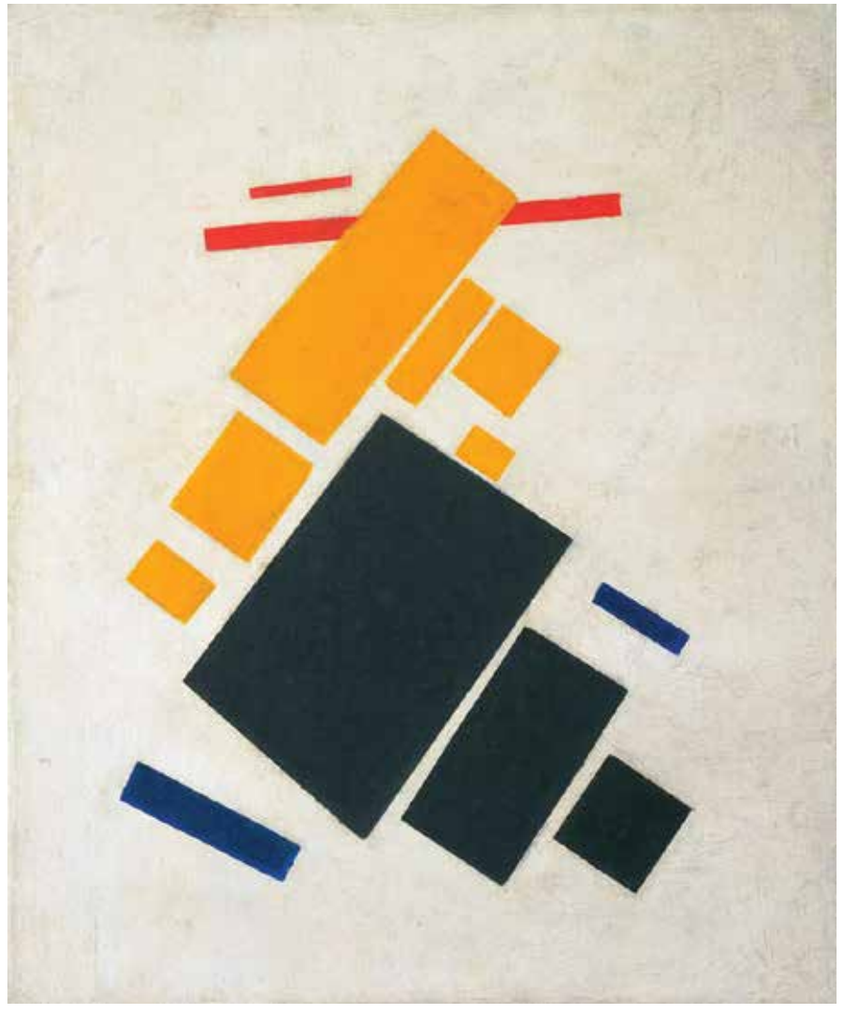
HOW CHILDREN LEARN TO BREAK THE RULES OF LANGUAGE

THE PRICE OF LINGUISTIC PRODUCTIVITY

CHARLES YANG

- Give a set of items:
  - If *many* do X, then all do X
  - if *few* do X, then remember the few that do
- How many is *many* or *few*?

- If a function is defined over *N* items, it generalizes iff it holds for at least *N-N/lnN* items.

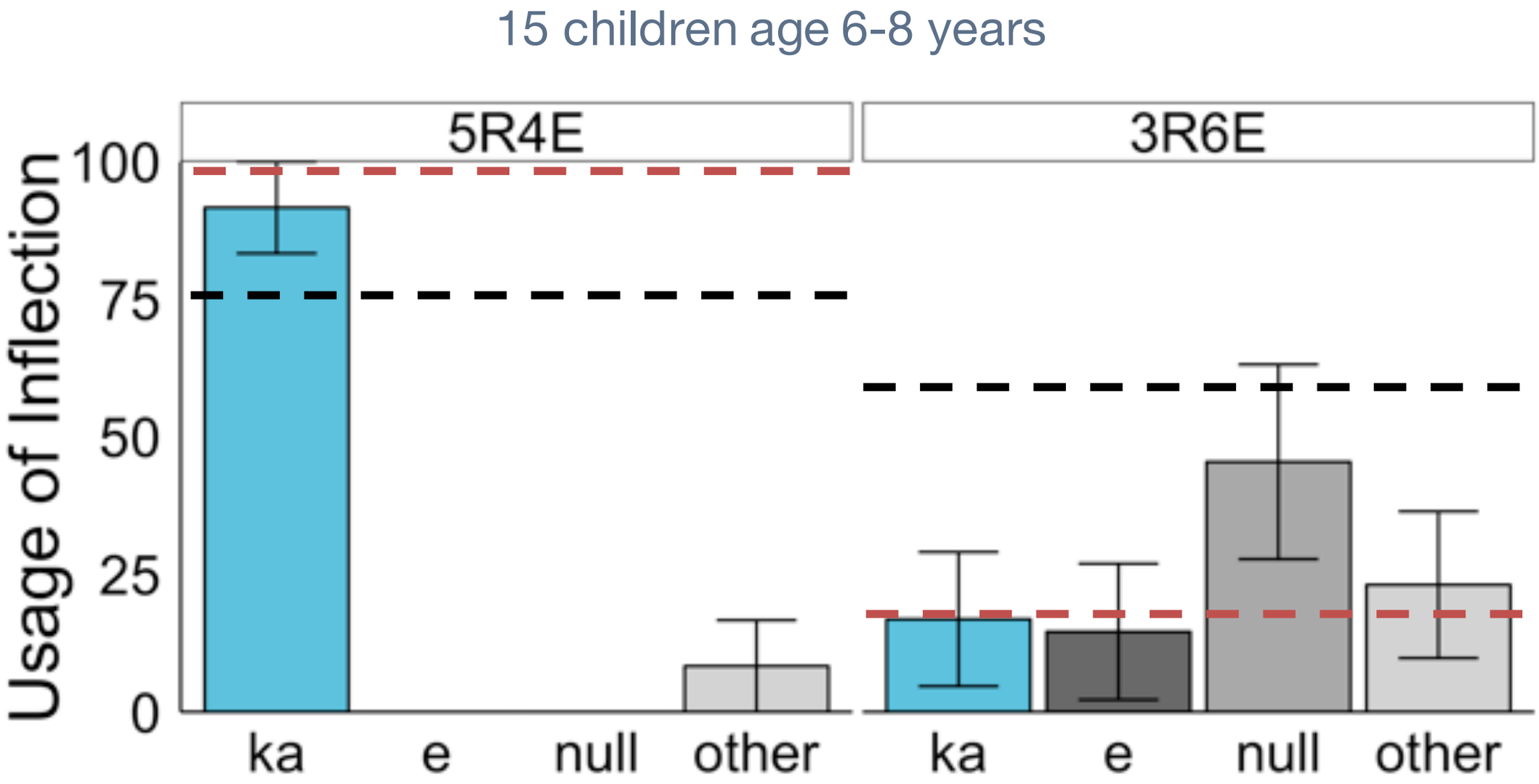


HOW CHILDREN LEARN TO BREAK THE RULES OF LANGUAGE

THE PRICE OF LINGUISTIC PRODUCTIVITY

CHARLES YANG

N=9, N/lnN=4.2



Schuler, Yang & Newport (2016, Submitted)

Emond & Shi (2021)

Emond & Shi (2021)

14-month-old non-Russian learning infants

“Movement” R1: ABC→BAC; R2: ABC→ACB

Machty gnutsja lukom → Gnutsja machty lukom

Machty gnutsja lukom → Machty lukom gnutsja

Emond & Shi (2021)

14-month-old non-Russian learning infants

“Movement” R1: ABC→BAC; R2: ABC→ACB

Machty gnutsja lukom → Gnutsja machty lukom

Machty gnutsja lukom → Machty lukom gnutsja

Emond & Shi (2021)

14-month-old non-Russian learning infants

“Movement” R1: ABC→BAC; R2: ABC→ACB

Machty gnutsja lukom → Gnutsja machty lukom

Machty gnutsja lukom → Machty lukom gnutsja

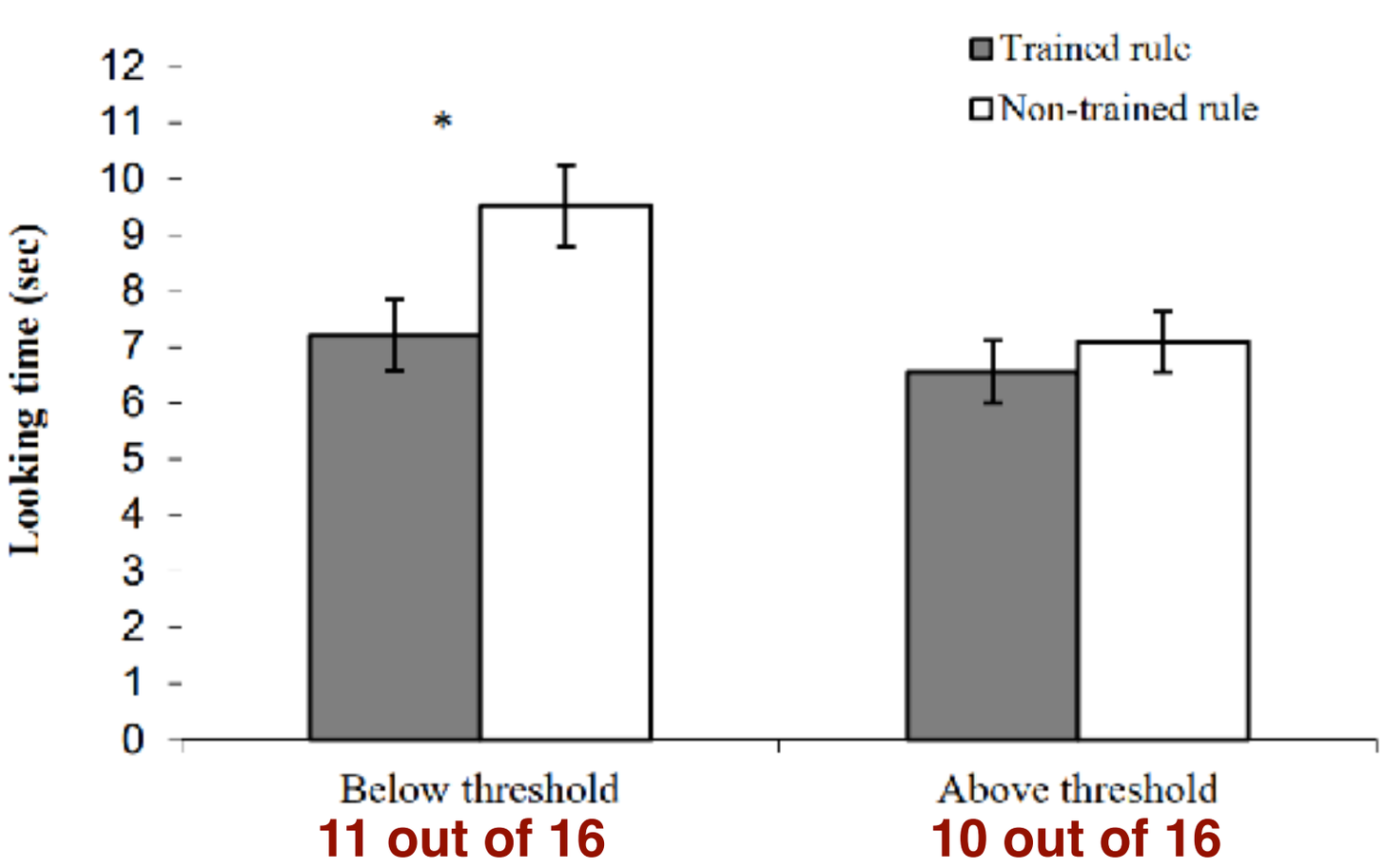
Emond & Shi (2021)

14-month-old non-Russian learning infants

“Movement” R1: ABC→BAC; R2: ABC→ACB

Machty gnutsja lukom → Gnutsja machty lukom

Machty gnutsja lukom → Machty lukom gnutsja



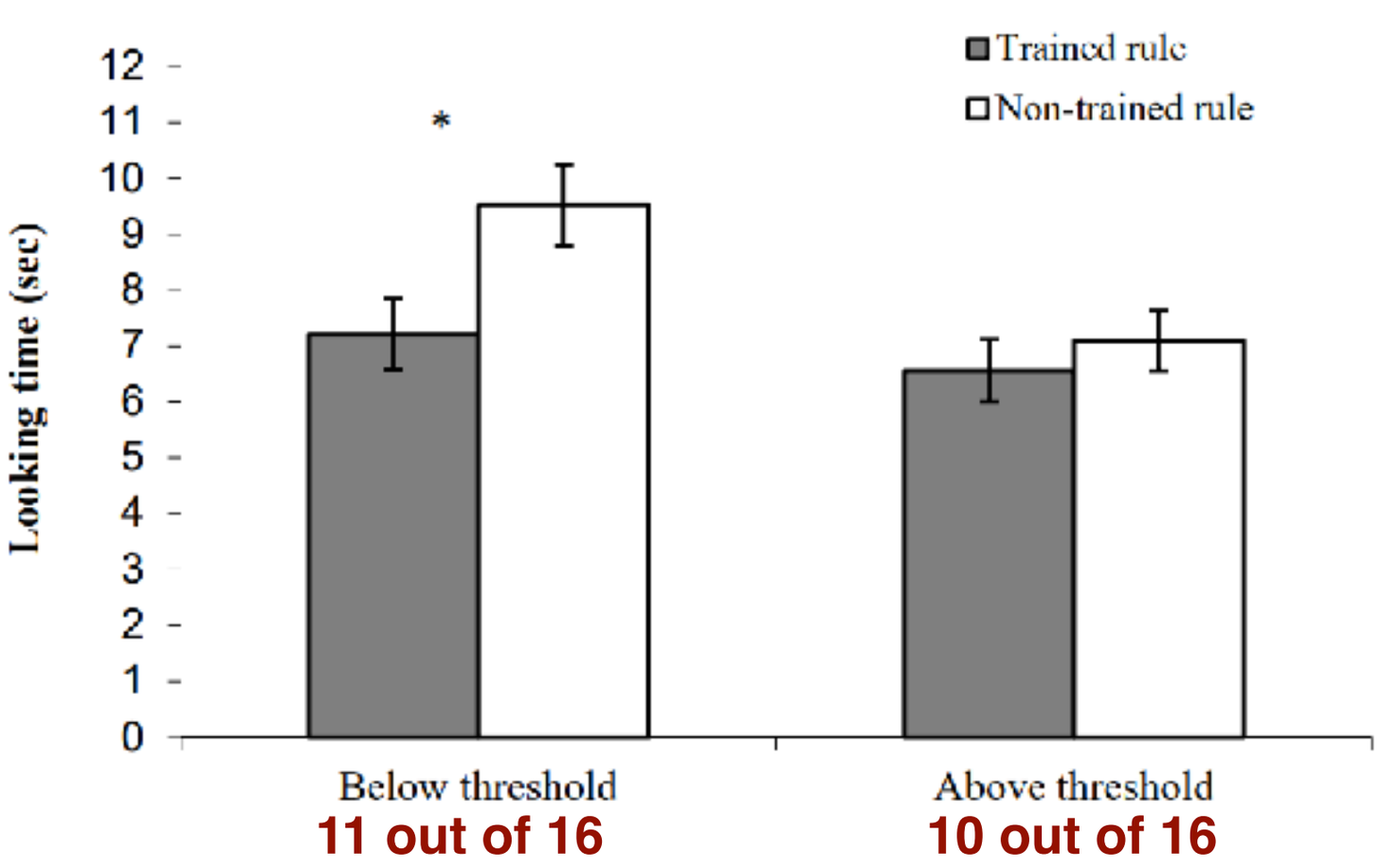
Emond & Shi (2021)

14-month-old non-Russian learning infants

“Movement” R1: ABC→BAC; R2: ABC→ACB

Machty gnutsja lukom → Gnutsja machty lukom

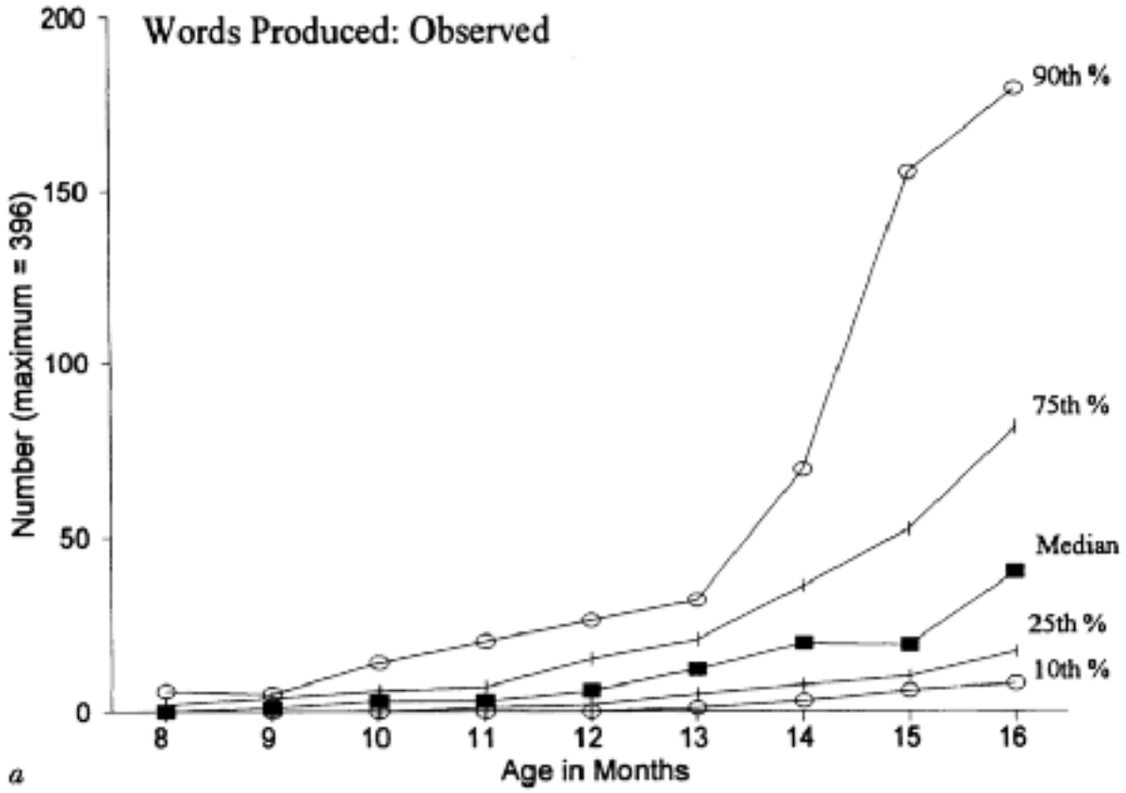
Machty gnutsja lukom → Machty lukom gnutsja



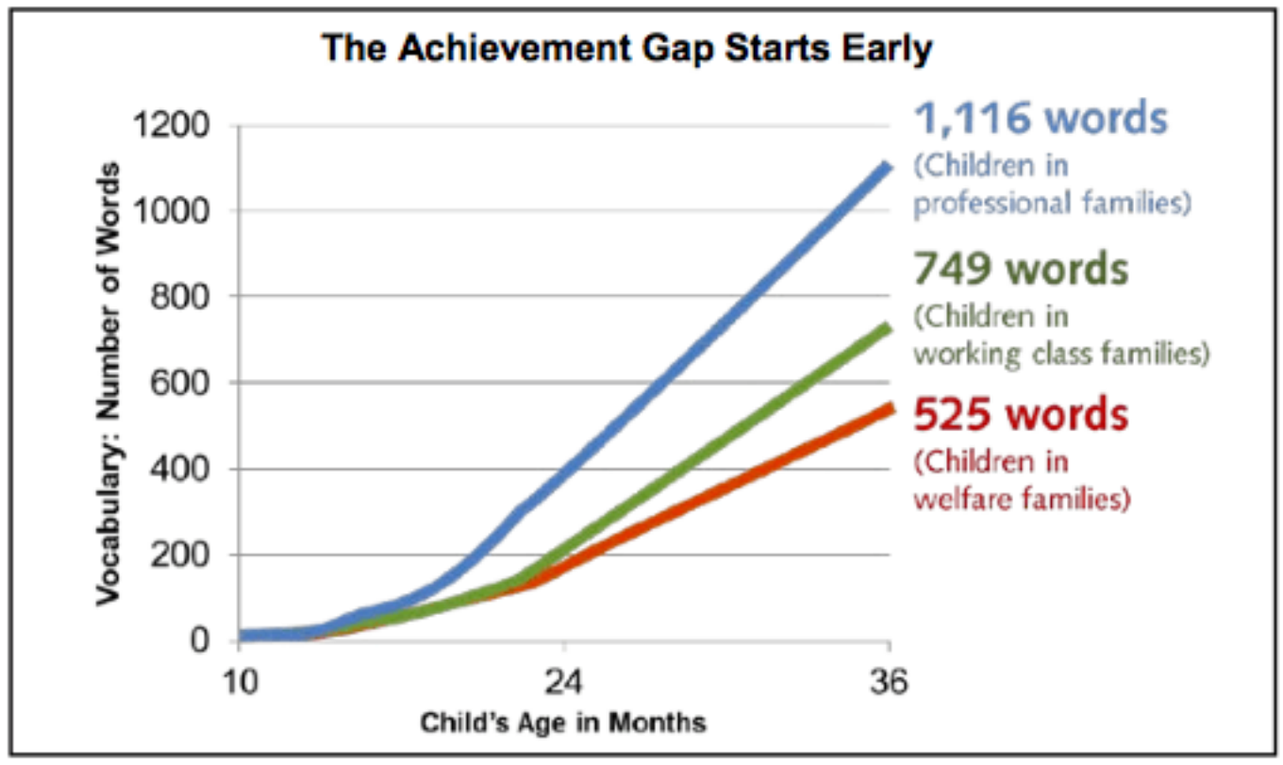
N=16

N/lnN=5.77

From Words to Rules



Fenson et al (1994)



Hart & Risley (1995)

For word learning, see Stevens, Trueswell, Gleitman & Yang (2017), Soh & Yang (2021)

Discovery Procedure as Abduction

1

2

3

4

5

6

7

8

9

10

11

Discovery Procedure as Abduction

1
2
3
4
5
6
7
8
9
10
11

- Find the most frequent color (6 vs. 5: ■)

Discovery Procedure as Abduction

1

2

3

4

5

6

7

8

9

10

11

- Find the most frequent color (6 vs. 5: ■)
- Hypothesize a rule that maps to ■

Discovery Procedure as Abduction

- Find the most frequent color (6 vs. 5: ■)
- Hypothesize a rule that maps to ■

Discovery Procedure as Abduction

1

3

- Find the most frequent color (6 vs. 5: ■)
- Hypothesize a rule that maps to ■

5

6

7

9

Discovery Procedure as Abduction

1

3

- Find the most frequent color (6 vs. 5: ■)
- Hypothesize a rule that maps to ■
  - Odd → ■

5

6

7

9

Discovery Procedure as Abduction

1

3

- Find the most frequent color (6 vs. 5: ■)
- Hypothesize a rule that maps to ■
  - Odd→ ■
- TP check (6~1): 1, 3, 5, **6**, 7, 9

5

6

7

9

Discovery Procedure as Abduction

- Find the most frequent color (6 vs. 5: ■)
- Hypothesize a rule that maps to ■
  - Odd→ ■
  - TP check (6~1): 1, 3, 5, **6**, 7, 9

Discovery Procedure as Abduction

- Find the most frequent color (6 vs. 5: ■)
- Hypothesize a rule that maps to ■
  - Odd→ ■
    - TP check (6~1): 1, 3, 5, **6**, 7, 9
- Test the rule Odd→ ■

Discovery Procedure as Abduction

1

3

- Find the most frequent color (6 vs. 5: ■)
- Hypothesize a rule that maps to ■
  - Odd→ ■
  - TP check (6~1): 1, 3, 5, **6**, 7, 9
- Test the rule Odd→ ■

5

7

9

11

Discovery Procedure as Abduction

1

3

- Find the most frequent color (6 vs. 5: ■)
- Hypothesize a rule that maps to ■
  - Odd→ ■
  - TP check (6~1): 1, 3, 5, **6**, 7, 9
- Test the rule Odd→ ■
  - TP check (6~1): 1, 3, 5, 7, 9, 11

5

7

9

11

Discovery Procedure as Abduction

1

3

- Find the most frequent color (6 vs. 5: ■)
- Hypothesize a rule that maps to ■
  - Odd→ ■
  - TP check (6~1): 1, 3, 5, **6**, 7, 9

5

- Test the rule Odd→ ■
  - TP check (6~1): 1, 3, 5, 7, 9, 11

7

- R1 productive: Odd→ ■ + 11

9

11

Discovery Procedure as Abduction

- Find the most frequent color (6 vs. 5: ■)
- Hypothesize a rule that maps to ■
  - Odd→ ■
    - TP check (6~1): 1, 3, 5, **6**, 7, 9
- Test the rule Odd→ ■
    - TP check (6~1): 1, 3, 5, 7, 9, 11
- R1 productive: Odd→ ■ + 11

Discovery Procedure as Abduction

1
2
3
4
5
6
7
8
9
10
11

- Find the most frequent color (6 vs. 5: ■)
- Hypothesize a rule that maps to ■
  - Odd→ ■
  - TP check (6~1): 1, 3, 5, **6**, 7, 9
- Test the rule Odd→ ■
  - TP check (6~1): 1, 3, 5, 7, 9, 11
- R1 productive: Odd→ ■ + 11

Discovery Procedure as Abduction

1
2
3
4
5
6
7
8
9
10
11

- Find the most frequent color (6 vs. 5: ■)
- Hypothesize a rule that maps to ■
  - Odd→ ■
  - TP check (6~1): 1, 3, 5, **6**, 7, 9
- Test the rule Odd→ ■
  - TP check (6~1): 1, 3, 5, 7, 9, 11
- R1 productive: Odd→ ■ + 11
- Remove words under R1 and repeat until {}

Discovery Procedure as Abduction

- Find the most frequent color (6 vs. 5: ■)
- Hypothesize a rule that maps to ■
  - Odd→ ■
    - TP check (6~1): 1, 3, 5, **6**, 7, 9
- Test the rule Odd→ ■
    - TP check (6~1): 1, 3, 5, 7, 9, 11
- R1 productive: Odd→ ■ + 11
- Remove words under R1 and repeat until {}

Discovery Procedure as Abduction

2

- Find the most frequent color (6 vs. 5: ■)
- Hypothesize a rule that maps to ■

4

  - Odd→ ■
    - TP check (6~1): 1, 3, 5, **6**, 7, 9

6

- Test the rule Odd→ ■
  - TP check (6~1): 1, 3, 5, 7, 9, 11

8

- R1 productive: Odd→ ■ + 11
- Remove words under R1 and repeat until {}

10

Discovery Procedure as Abduction

2

4

6

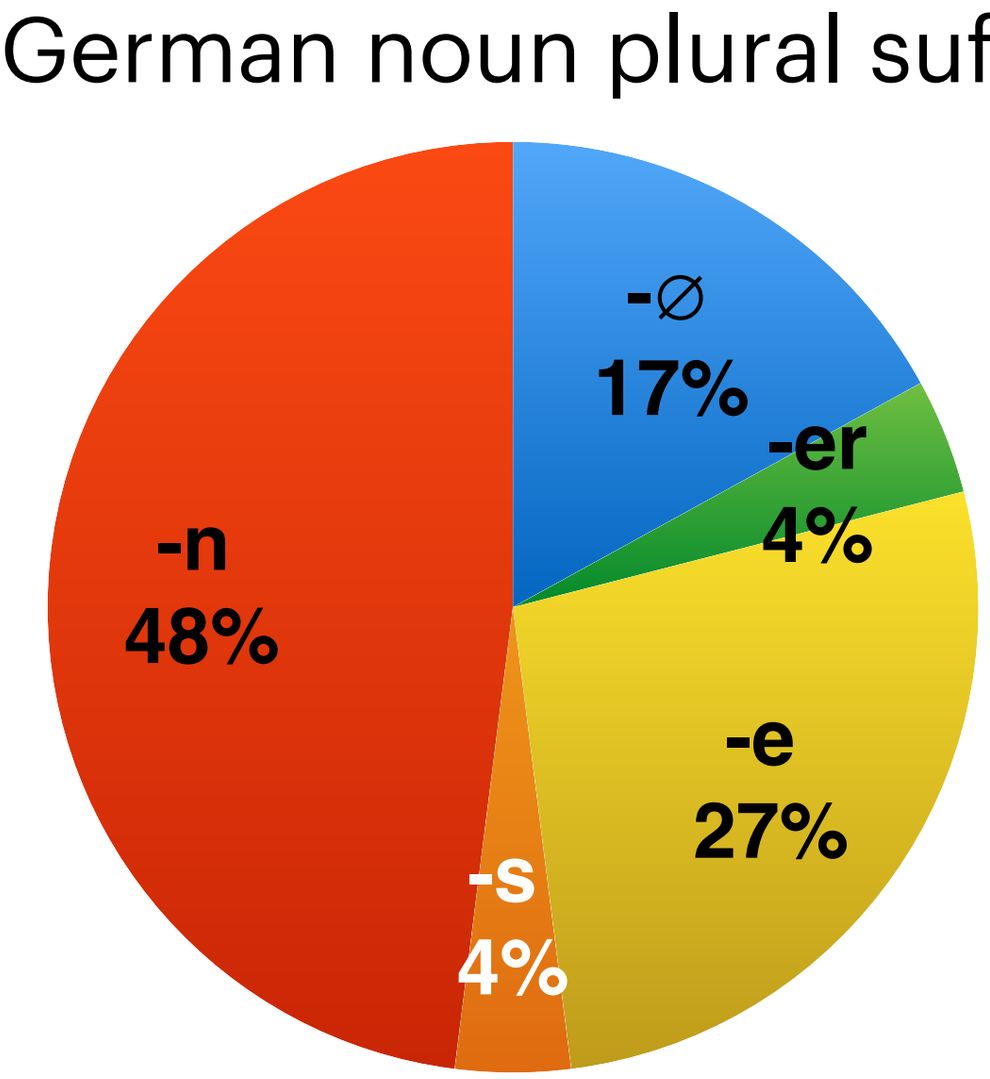
8

10

- Find the most frequent color (6 vs. 5: ■)
- Hypothesize a rule that maps to ■
  - Odd→ ■
  - TP check (6~1): 1, 3, 5, **6**, 7, 9
- Test the rule Odd→ ■
  - TP check (6~1): 1, 3, 5, 7, 9, 11
- R1 productive: Odd→ ■ + 11
- Remove words under R1 and repeat until {}
- R2 productive: Even → ■ + 6

The Awful German Language

The Awful German Language



Autos, Parks, Pizzas, ..., iPhones

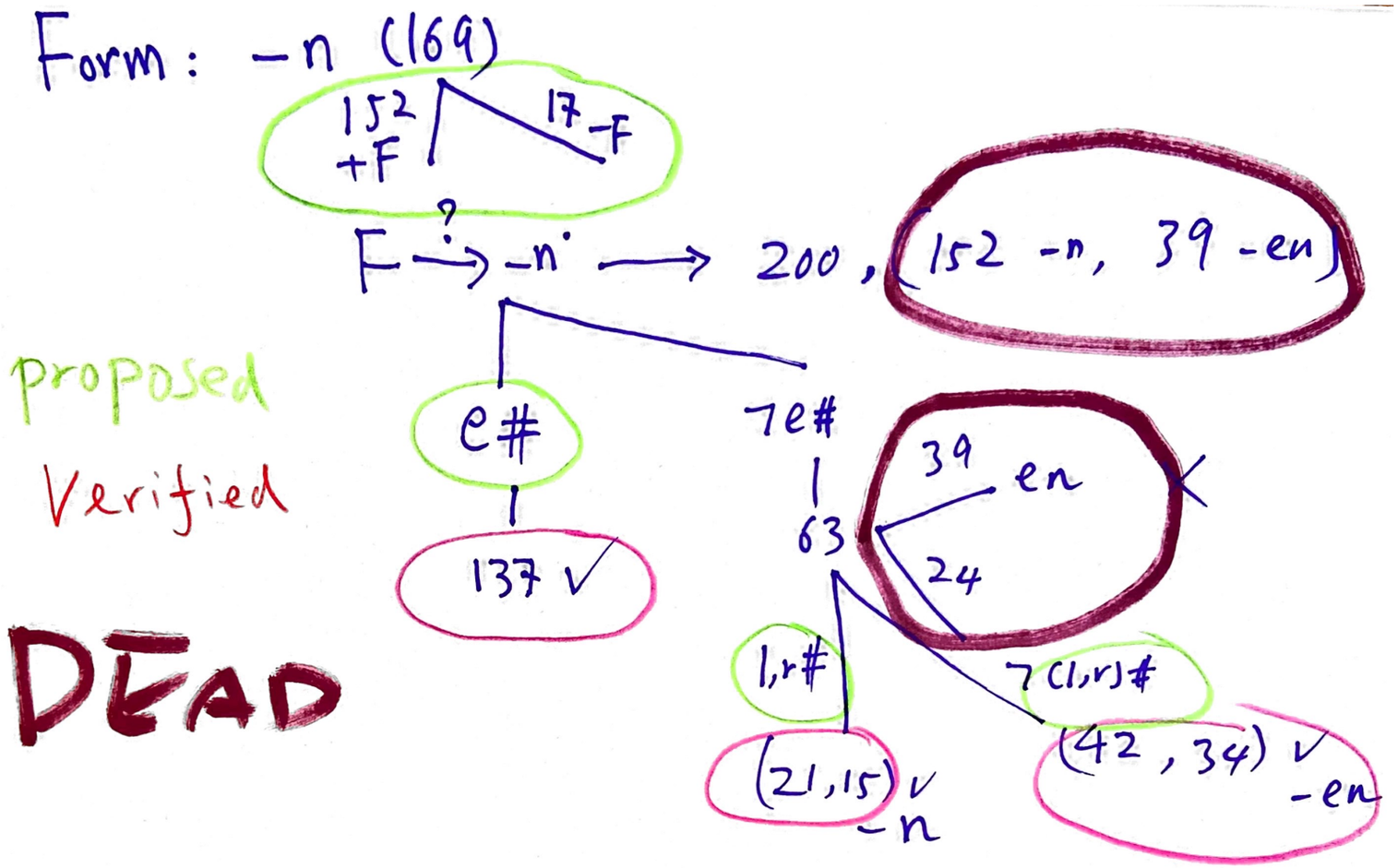
Simulating Leo

- Leo corpus: nom.sg nouns and their nom.pl forms
  - 200 F, 164 M, and 78 N

Suffix	Type	%
-n	169	38.2%
-null	121	27.4%
-e	80	18.1%
-en	47	10.6%
-er	15	3.4%
-s	8	1.8%
Other	2	0.5%
Total	442	100%

Suffix	Type	%
-n	**169**	**38.2%**
-null	121	27.4%
-e	80	18.1%
-en	47	10.6%
-er	15	3.4%
-s	8	1.8%
Other	2	0.5%
Total	442	100%

Suffix	Type	%
-n	**169**	**38.2%**
-null	121	27.4%
-e	80	18.1%
-en	47	10.6%
-er	15	3.4%
-s	8	1.8%
Other	2	0.5%
Total	442	100%



(+F, e#) → -n (137)

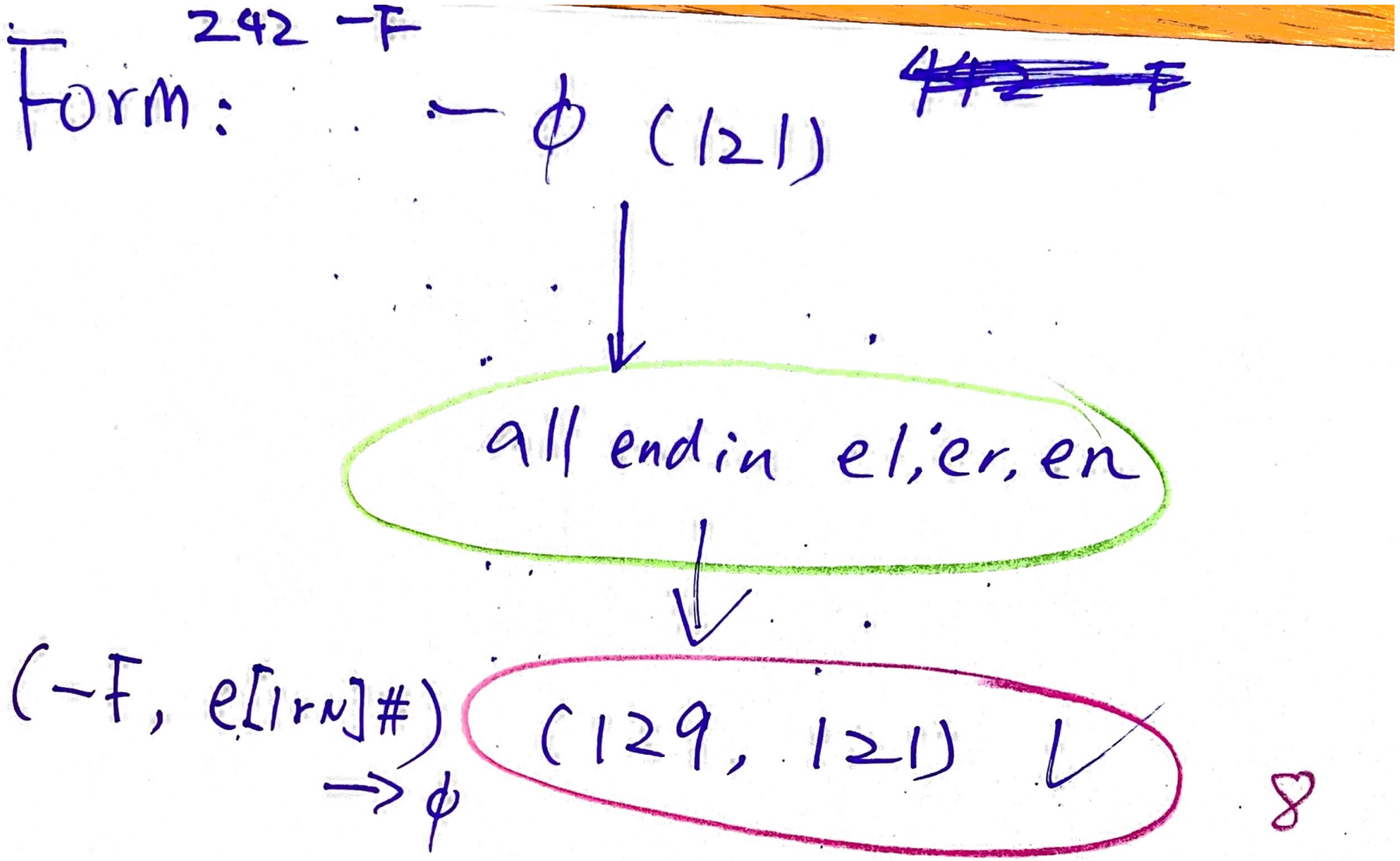
(+F, (l,r)#) → -n (21,15) 6

(+F, 7(l,r)#) → -en (42,34) 8

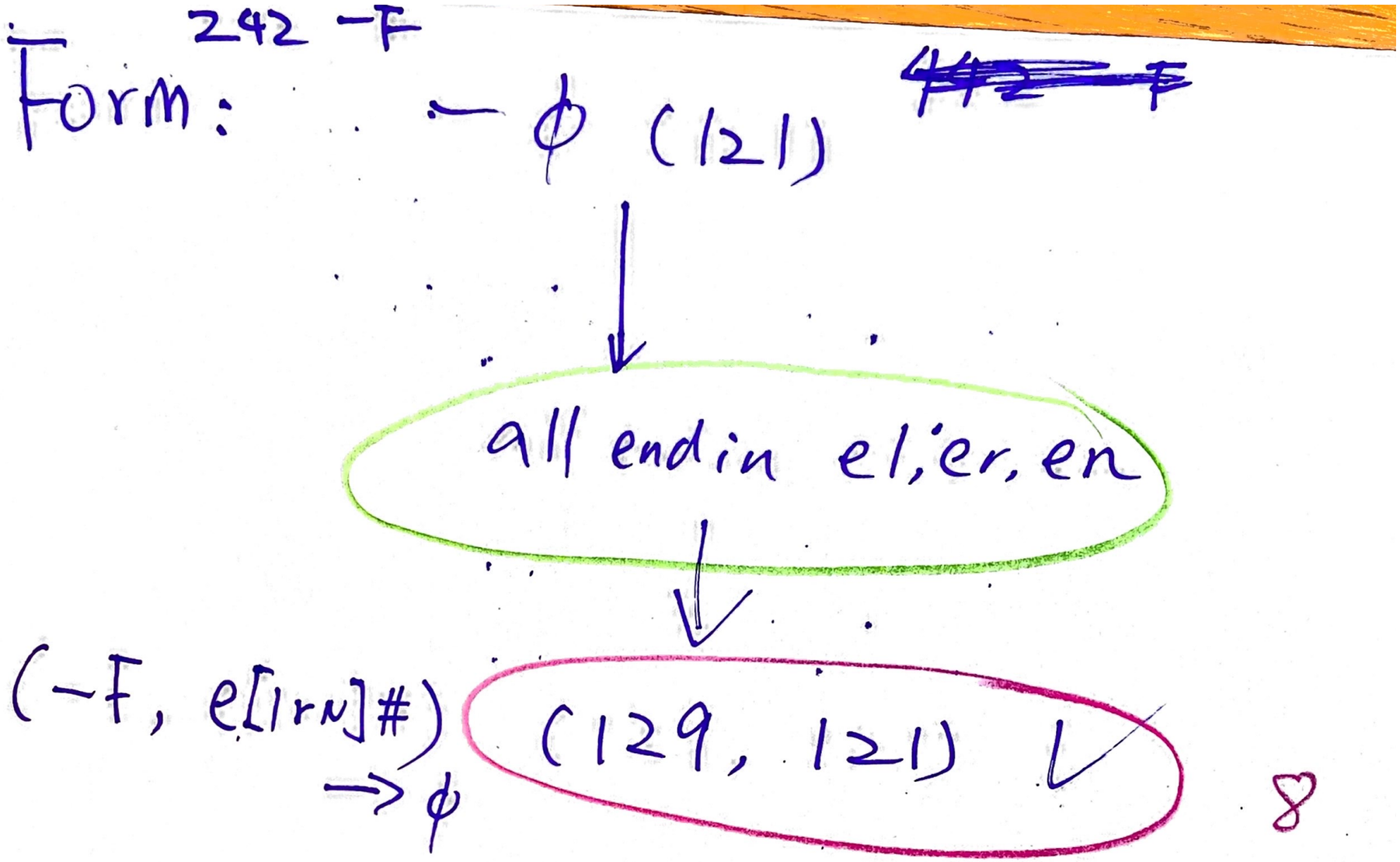
200 Fem Nouns accounted For with 14 exceptions

Suffix	Type	%
-n	169	38.2%
-null	**121**	**27.4%**
-e	80	18.1%
-en	47	10.6%
-er	15	3.4%
-s	8	1.8%
Other	2	0.5%
Total	442	100%

Suffix	Type	%
-n	169	38.2%
-null	121	27.4%
-e	80	18.1%
-en	47	10.6%
-er	15	3.4%
-s	8	1.8%
Other	2	0.5%
Total	442	100%



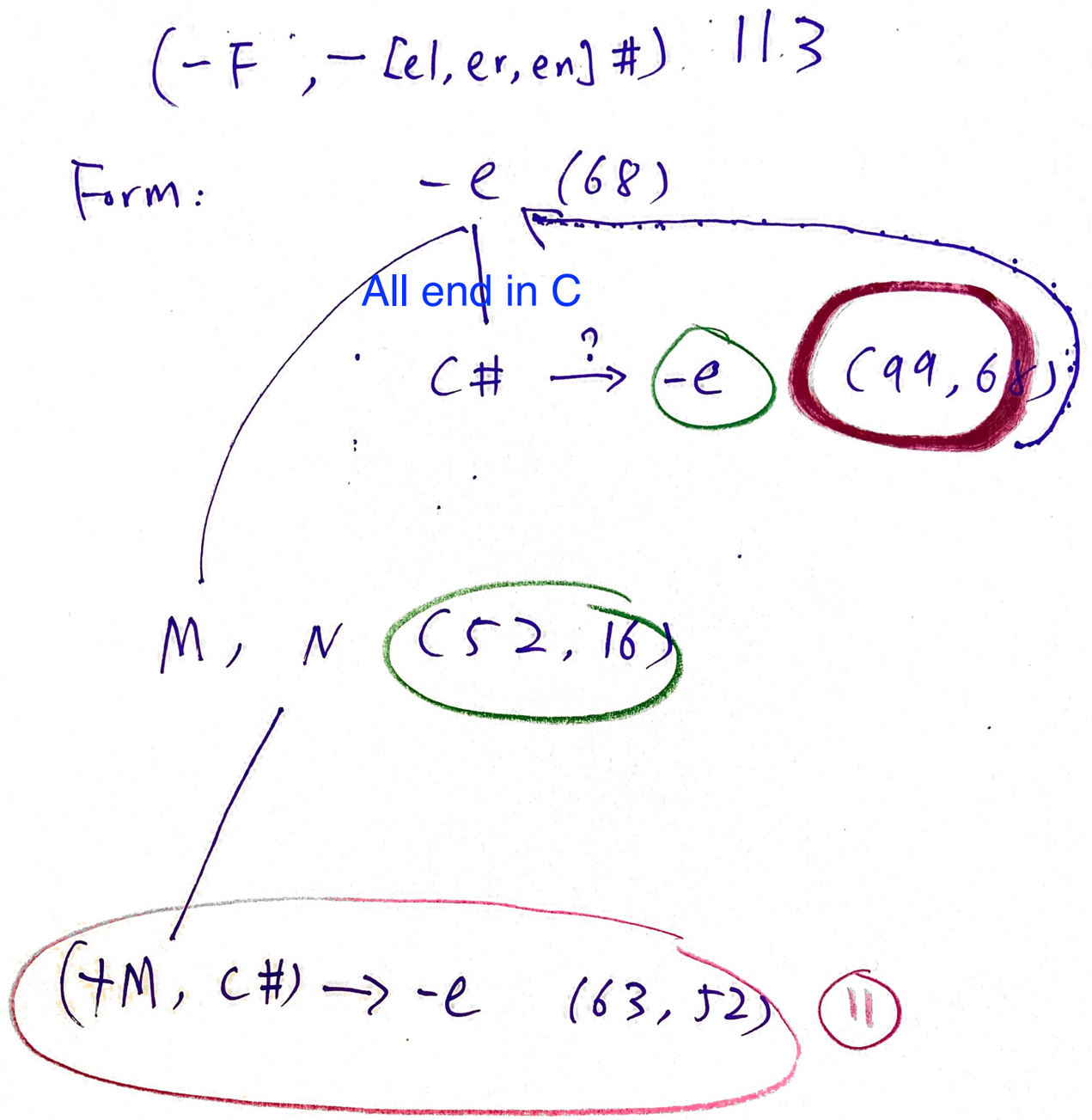
Suffix	Type	%
-n	169	38.2%
-null	121	27.4%
-e	80	18.1%
-en	47	10.6%
-er	15	3.4%
-s	8	1.8%
Other	2	0.5%
Total	442	100%



-F, e[lrn]#: add null (see Wiese 1996)

Suffix	Type	%
-n	169	38.2%
-null	121	27.4%
-e	**80**	**18.1%**
-en	47	10.6%
-er	15	3.4%
-s	8	1.8%
Other	2	0.5%
Total	442	100%

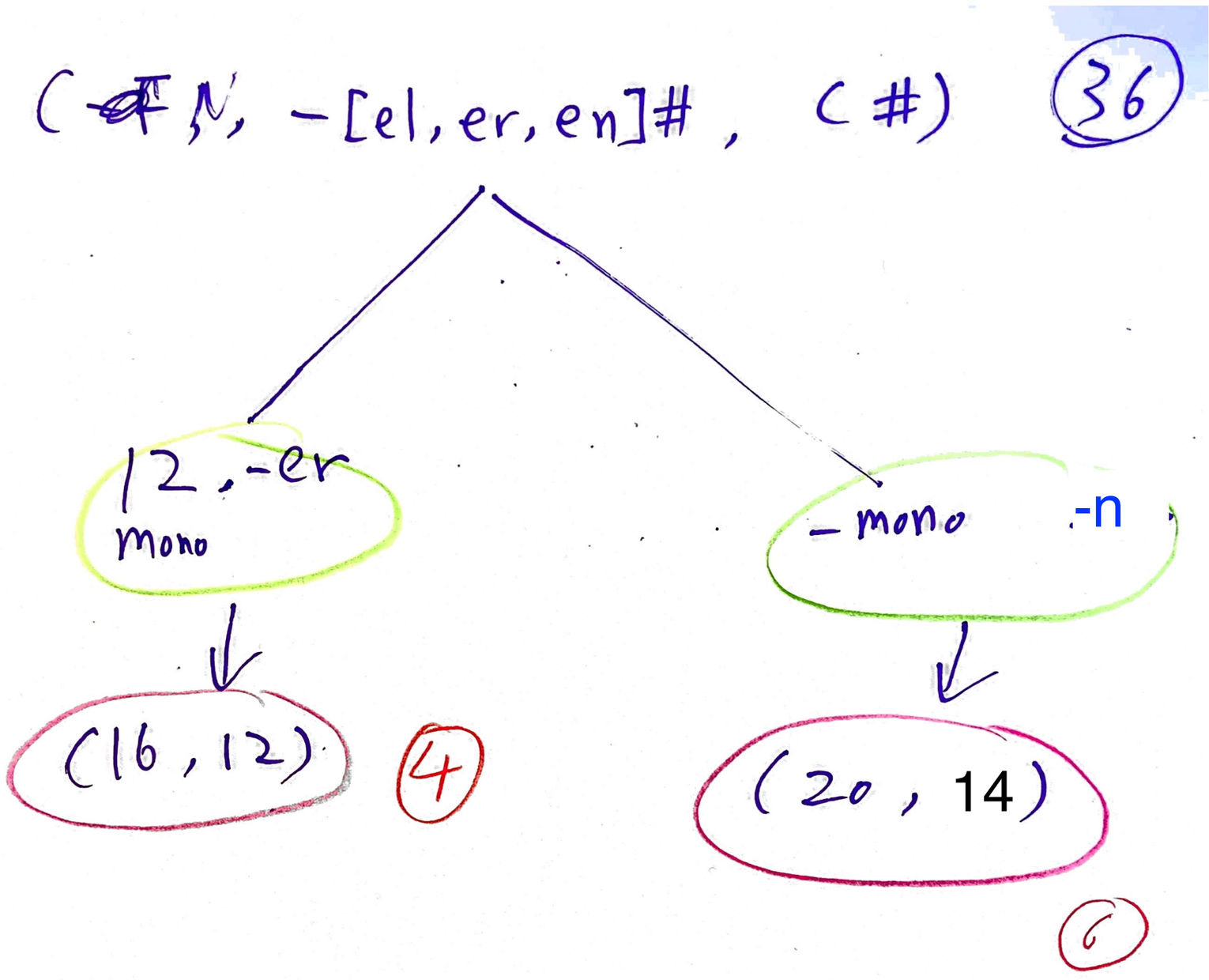
Suffix	Type	%
-n	169	38.2%
-null	121	27.4%
-e	80	18.1%
-en	47	10.6%
-er	15	3.4%
-s	8	1.8%
Other	2	0.5%
Total	442	100%



now ~~Are~~ (+N, C#) words

will recurse back to
(-F, -[el, er, -en] #,
e#) later

36 out of 50 (N, C#)



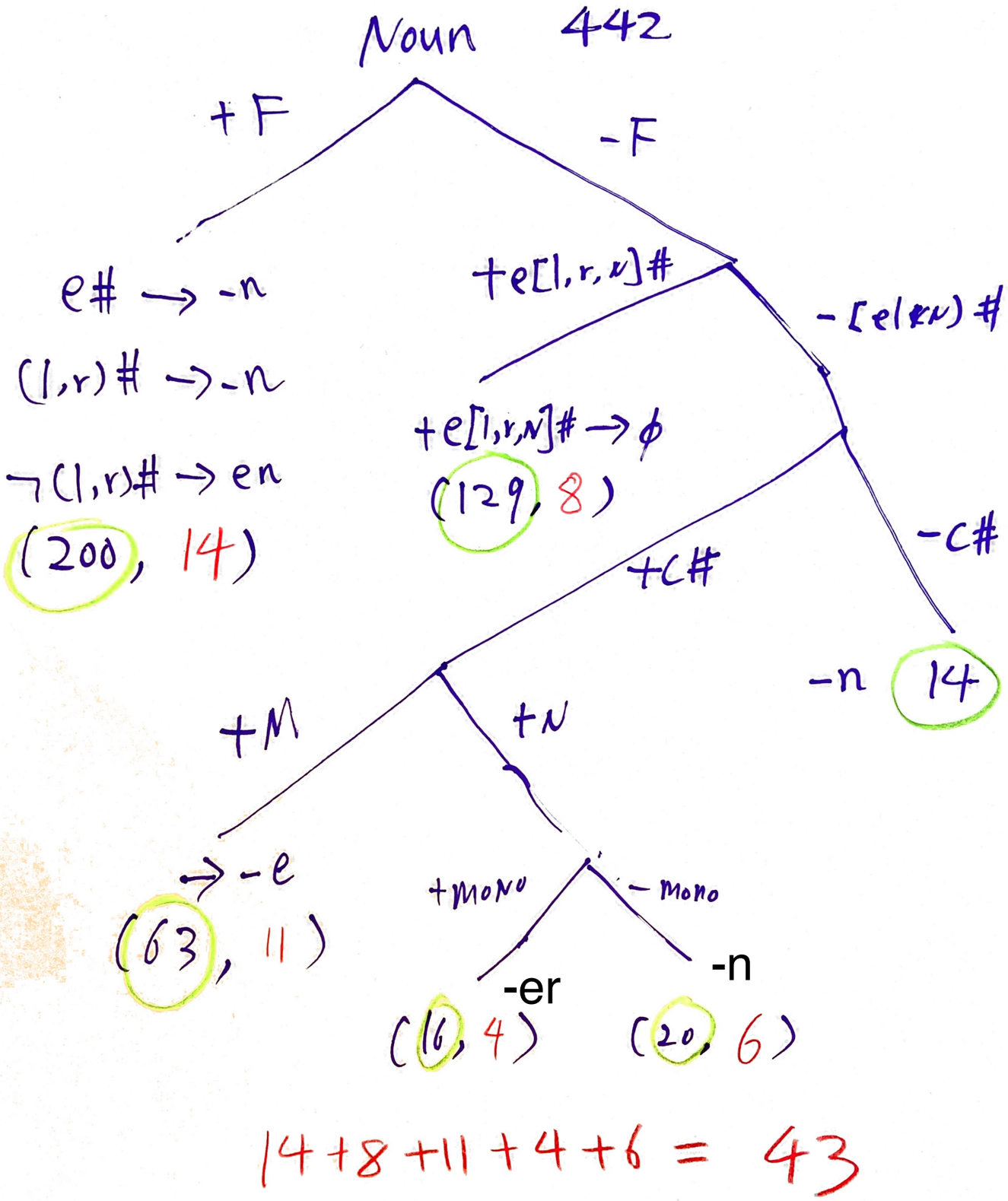
14 out of 14: (-F, e#)

Finally ..

(-F, -[e1, er en]#, e#)

14: add -n

The horrors of German (about 10%!)



Noun	Gender
Oma	F
Mail	F
Tunnel	M
Baby	N
Kanguru	N
Detail	N
Tief	N
Emu	N

Status of -s in the sample

All are exceptions to productive rules

Note productive rules require gender

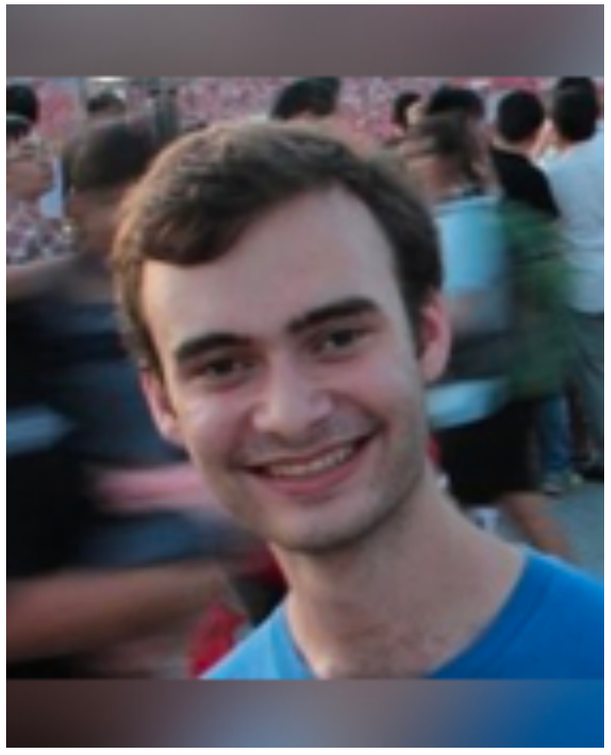
If a noun has unknown gender, as in foreign load words, they could fall under -s

As such, they pose no threat to the regularity of the plural system

Abduction of Tolerable Productivity (ATP)

Caleb Belth
(Michigan CS)

Sarah Payne
(Penn Linguistics/CS)

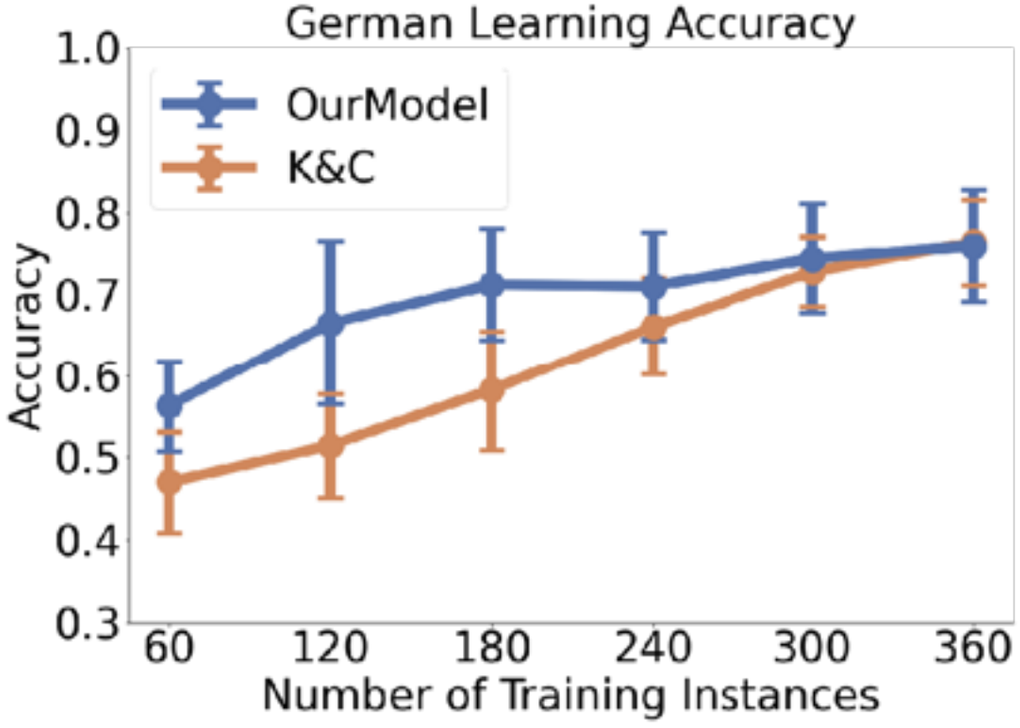


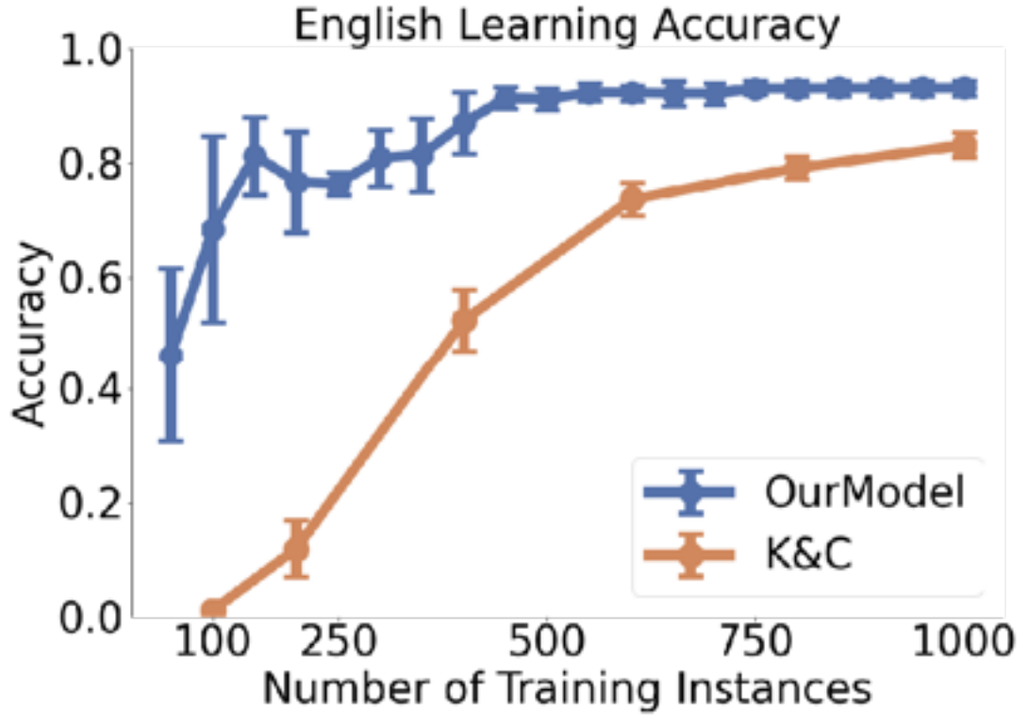
Jordan Kodner
(Stony Brook Linguistics)

CogSci 2021

Children vs. Machines

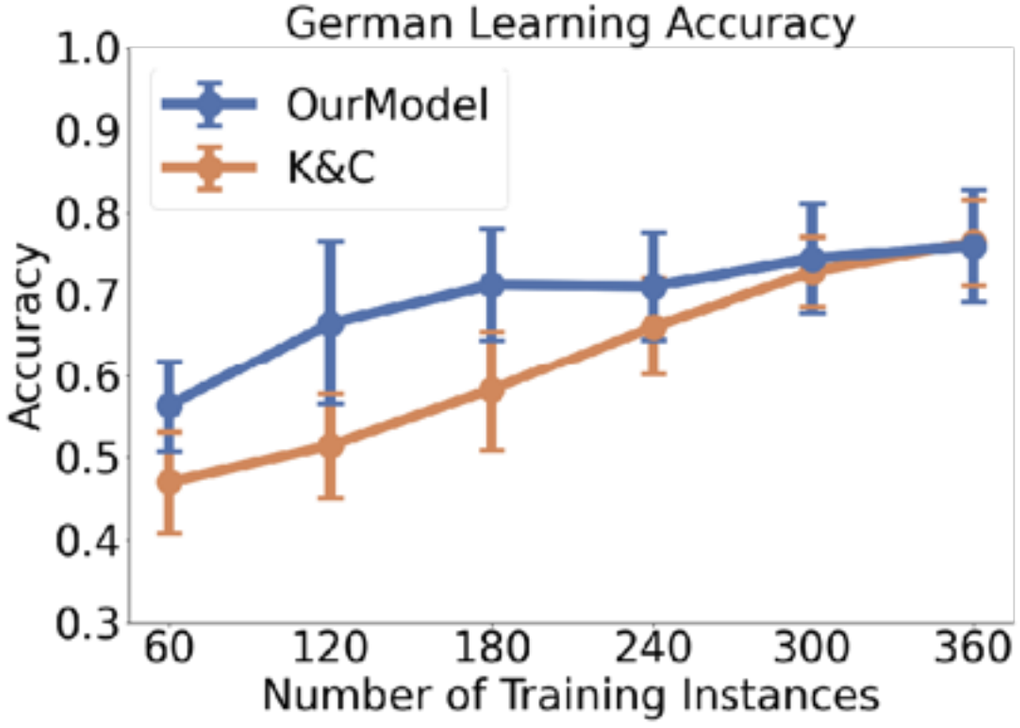
- Tested on a random sample of German nouns and English past tense in CELEX
- Compared against a state-of-the-art neural network (Kirov & Cotterell 2018)

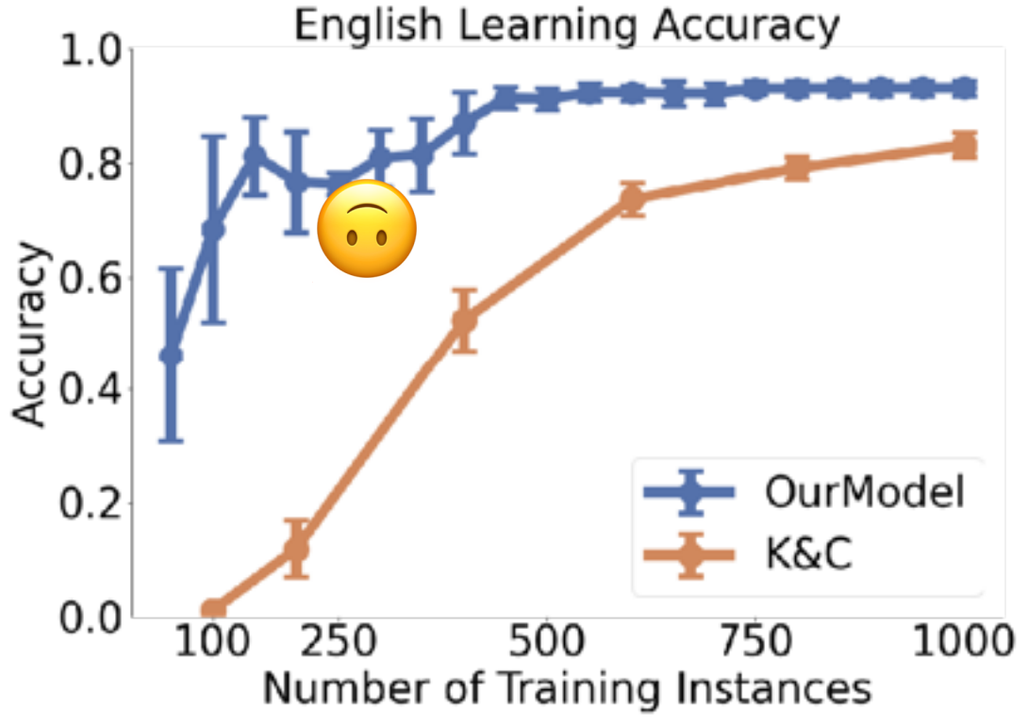




Children vs. Machines

- Tested on a random sample of German nouns and English past tense in CELEX
- Compared against a state-of-the-art neural network (Kirov & Cotterell 2018)





Morphology and Syntax

- Morphology is clearly language specific and must be learned
  - Various efforts to make morphology like syntax
- An adequate theory of learning morphology learning may be sufficient for learning syntax
  - Perhaps syntax is like morphology

REMARKS ON NOMINALIZATION*

(2) a. John is eager to please.
b. John has refused the offer.
c. John criticized the book.

(3) a. John's being eager to please
b. John's refusing the offer
c. John's criticizing the book

(4) a. John's eagerness to please
b. John's refusal of the offer
c. John's criticism of the book

Many differences have been noted between these two types of nominalization. The most striking differences have to do with the productivity of the process in question, the generality of the relation between the nominal and the associated proposition, and the internal structure of the nominal phrase.

REMARKS ON NOMINALIZATION*

(2) a. John is eager to please.
b. John has refused the offer.
c. John criticized the book.

(3) a. John's being eager to please
b. John's refusing the offer
c. John's criticizing the book

(4) a. John's eagerness to please
b. John's refusal of the offer
c. John's criticism of the book

Many differences have been noted between these two types of nominalization. The most striking differences have to do with the productivity of the process in question, the generality of the relation between the nominal and the associated proposition, and the internal structure of the nominal phrase.

Prolegomena to a Theory of Word Formation*

[1] The proposal just sketched might be modified somewhat as regards the treatment of words formed by rules that traditionally have been called "nonproductive". One might propose that all words formed by nonproductive rules are marked by these rules as [−Lexical Insertion]. The smaller subset of actually occurring words formed by such rules would then be listed in the filter with the feature [+Lexical Insertion]. That is, the nouns formed with the suffix *-al* would all be generated with the feature [−Lexical Insertion]; the relatively small number of actually occurring nouns of this type, like those listed in (3a), will appear in the filter marked [+Lexical Insertion]. In other words, it is assumed that words generated by a productive process are all actually occurring and that only exceptionally may a word of this type be ruled out of the language. On the other hand, words generated by a nonproductive rule are assumed not to be occurring except under special circumstances. In this fashion we might capture the difference between productive and nonproductive formations.

English Passives

- *The read book
- *The told story
- The advanced technology
- The missed opportunity
- The mixed ingredients
- The cooked meat
- *The mentioned bisque
- *The kicked ball
- The unread book
- The untold story
- *The unadvanced technology
- *The unmissed opportunity
- The unmixed ingredients
- The uncooked meat
- *The unmentioned bisque
- *The unkicked ball

Adjectival Passives

- A long history in syntax (Anderson, Wasow, Bresnan, Williams)
- Levin & Rappaport (1986)'s Sole Complement Generalization

- Arguments that can serve as the sole NP complement to a verb lead to an adjectival passive
- I offered a contract vs. *I offered a job candidate
- The offered contract vs. *The offered job candidate

But

- They read a book ⇏ *the read book.
- They told a story ⇏ *the told story
- The mentioned an example ⇏ *the mentioned example
- Not even the dative verbs: e.g., showed (them) a movie ⇏ *a shown movie, shot (him) an email ⇏ *a shot email

- I googled the topic ⇏ *the googled topic
- I friended my neighbor ⇏ *the friended neighbor
- Bieber dropped an album ⇏ *the dropped album

Verbal vs. Adjectival Passive

Verbal vs. Adjectival Passive

- Top **100** most frequent transitive verbs

Verbal vs. Adjectival Passive

- Top **100** most frequent transitive verbs
- **85** have verbal passive counterpart (e.g., The pizza was *eaten*) in CHILDES input: productive

Verbal vs. Adjectival Passive

- Top **100** most frequent transitive verbs
- **85** have verbal passive counterpart (e.g., The pizza was *eaten*) in CHILDES input: productive
- Only **9** have unambiguous adjectival counterpart (attributive usage in NPs):
  - Nowhere near the level for productive generalization
  - *baked, chopped, fried, squashed*: possible productive subclasses
  - In larger corpora: *fried, grilled, sautéd, boiled, baked, …* ⇒ sous-vided

Verbal vs. Adjectival Passive

- Top **100** most frequent transitive verbs
- **85** have verbal passive counterpart (e.g., The pizza was *eaten*) in CHILDES input: productive
- Only **9** have unambiguous adjectival counterpart (attributive usage in NPs):
  - Nowhere near the level for productive generalization
  - *baked, chopped, fried, squashed*: possible productive subclasses
  - In larger corpora: *fried, grilled, sautéd, boiled, baked, …* ⇒ sous-vided
- Verbal passive is productive, adjectival passive is unproductive (and you need to hear them, and individual variation is expected)

*read book vs. unread book

*read book vs. unread book

- If *read is not an adjective, how come *unread* is an adjective?
  - Only answer: **un-** is also unproductive!

*read book vs. unread book

- If *read is not an adjective, how come *unread* is an adjective?
  - Only answer: **un-** is also unproductive!
- 64 **un-** prefixed adjectives in 5 million words of CHILDES
  - Only 10 are morphologically simplex: happy, usual, even, fair, true, real, pleasant, dead, stable, safe, able
- Thus unred, unquick, unnice …

*read book vs. unread book

- If *read is not an adjective, how come *unread* is an adjective?
  - Only answer: **un-** is also unproductive!
- 64 **un-** prefixed adjectives in 5 million words of CHILDES
  - Only 10 are morphologically simplex: happy, usual, even, fair, true, real, pleasant, dead, stable, safe, able
  - Thus unred, unquick, unnice …
- Top 50 adjectival passives, only 16 have an **un-** counterpart
  - **Un-** is not productive: **unread book** is accept because we hear it!
  - advanced/unadvanced technology, missed/unmissed opportunity, recommend/unrecommeded dish, noted/unnoted scholar, …

English Passives

- *The read book
- *The told story
- The advanced technology
- The missed opportunity
- The mixed ingredients
- The cooked meat
- *The mentioned bisque
- *The kicked ball
- The unread book
- The untold story
- *The unadvanced technology
- *The unmissed opportunity
- The unmixed ingredients
- The uncooked meat
- *The unmentioned bisque
- *The unkicked ball

Child as Little Linguist?

Child as Little Linguist?

Linguist as Little Child!

Child as Little Linguist?

Linguist as Little Child!

START COUNTING!