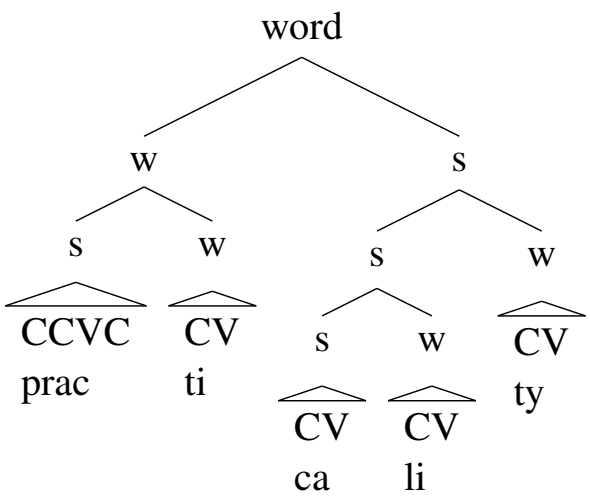
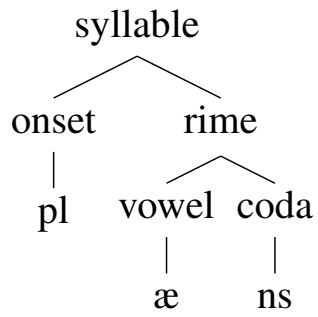
6.S077 Recitation 6
Animal “language”: Child vs. Chimp

1. Postscript to segmentation
2. Can animals learn human language? The case of Nim

How do children figure out stress for the word segmentation problem?
Consonant-Vowel pattern in babbling: universal

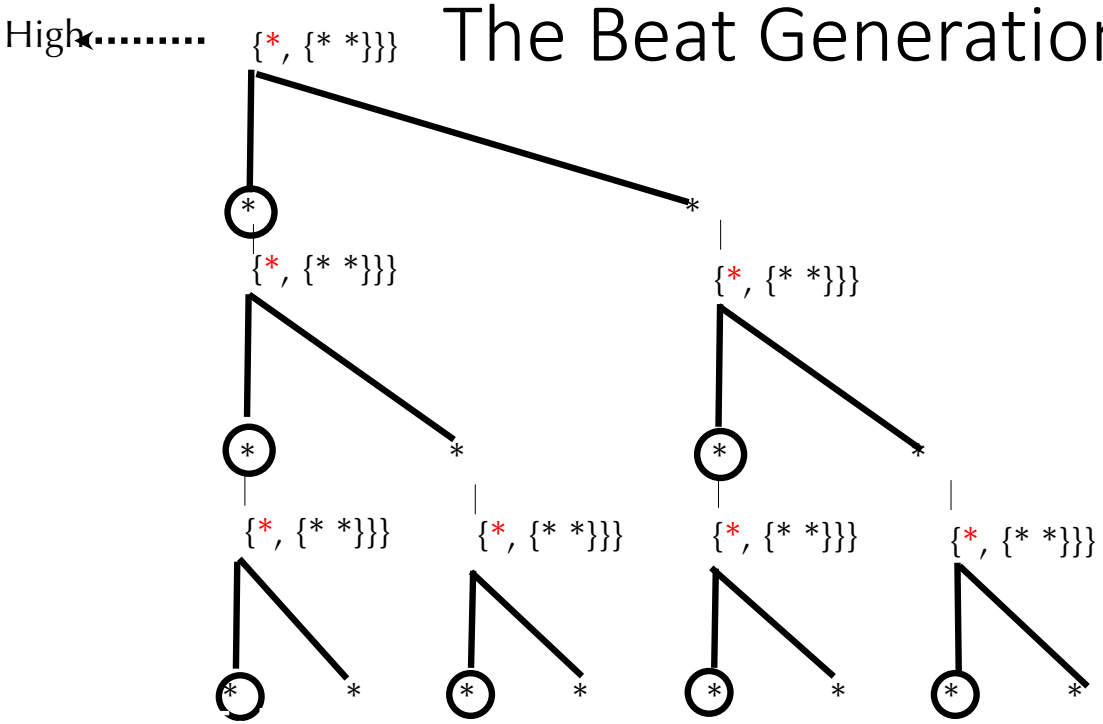




(a) Syllable structure of the word *plans.*

The right representation is <u>combinatorial</u>

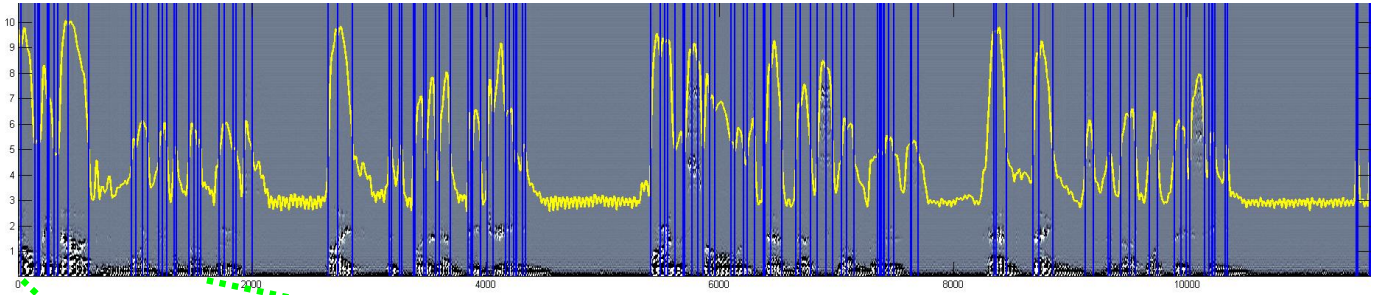
The Beat Generation



Iambic: mark <u>left</u> as “head” & project to next level

Suggests: there is an operation that takes <u>two</u> items & “merges” them

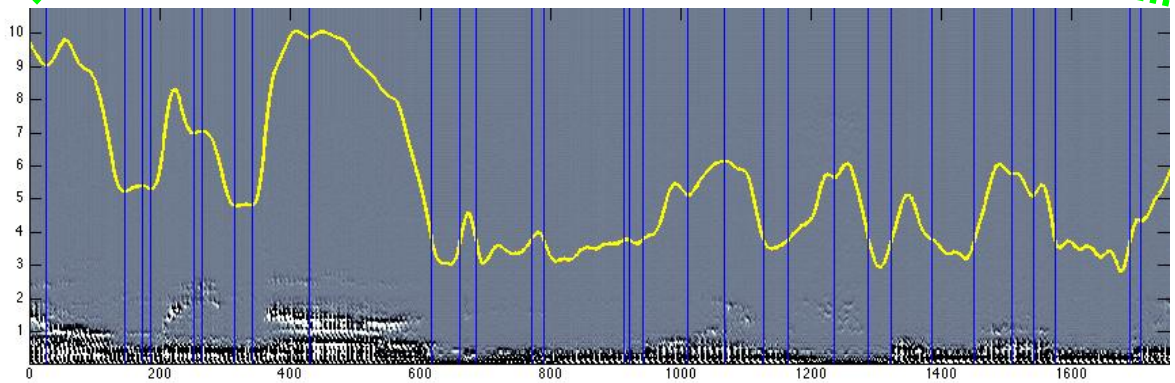
The Beat Generation



/ ∪ / ∪ / ∪ / ∪

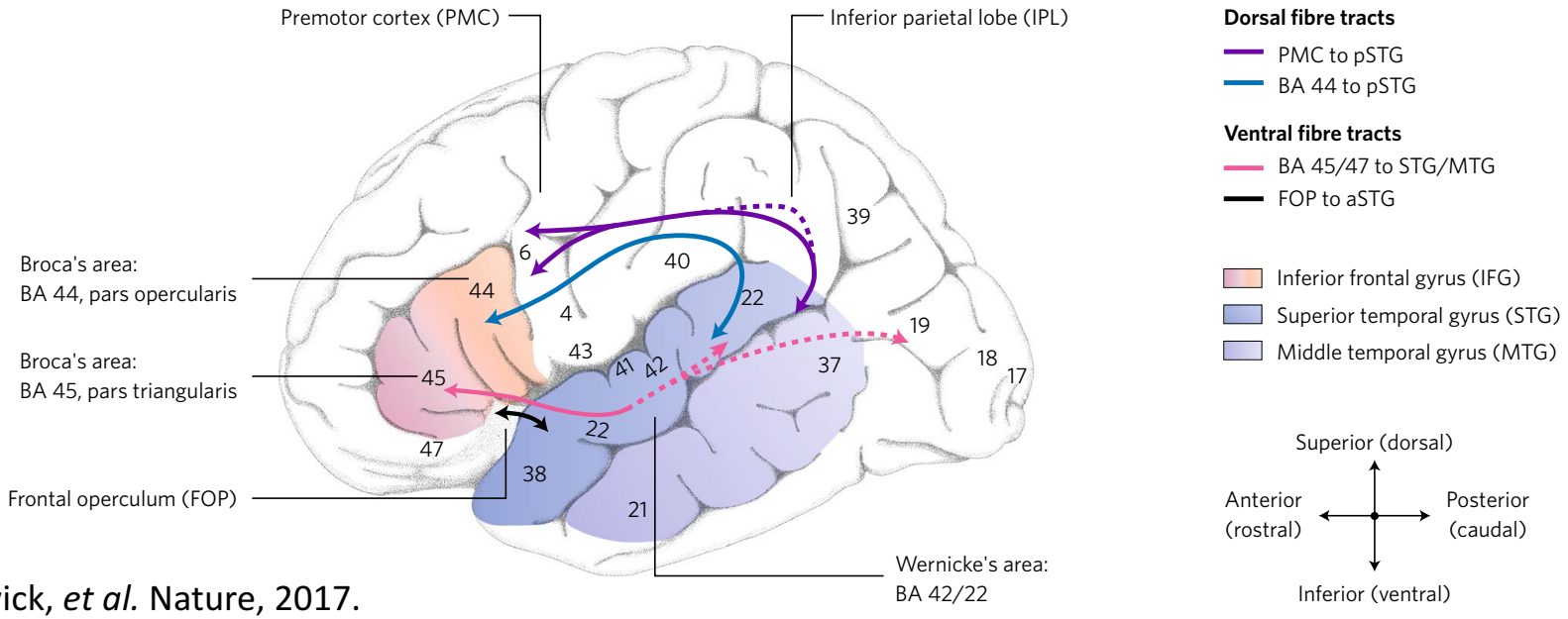
Tell me not in mournful numbers…

Yields beat pattern



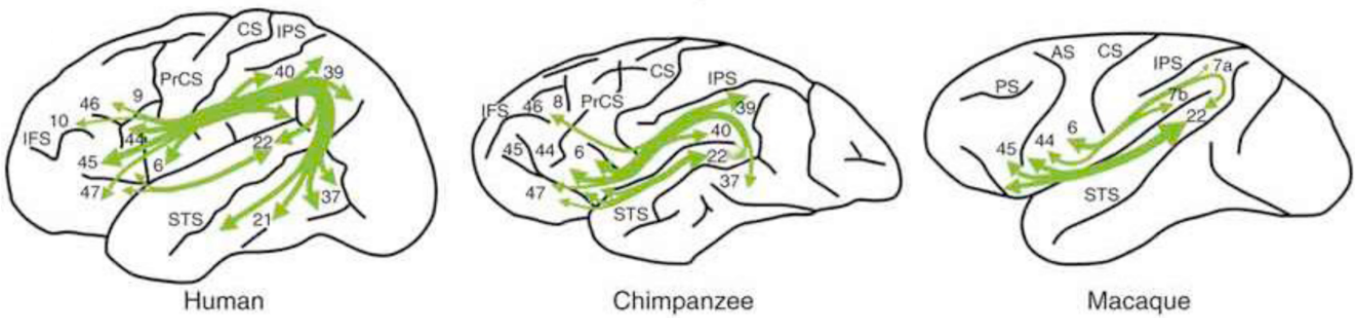
Human brain

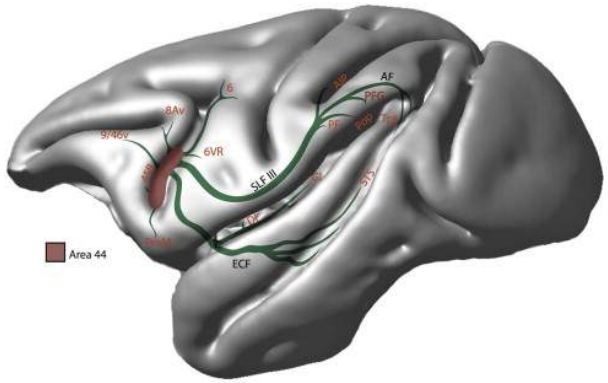
NATURE HUMAN BEHAVIOUR PERSPECTIVE



Federici, Berwick, *et al.* Nature, 2017.

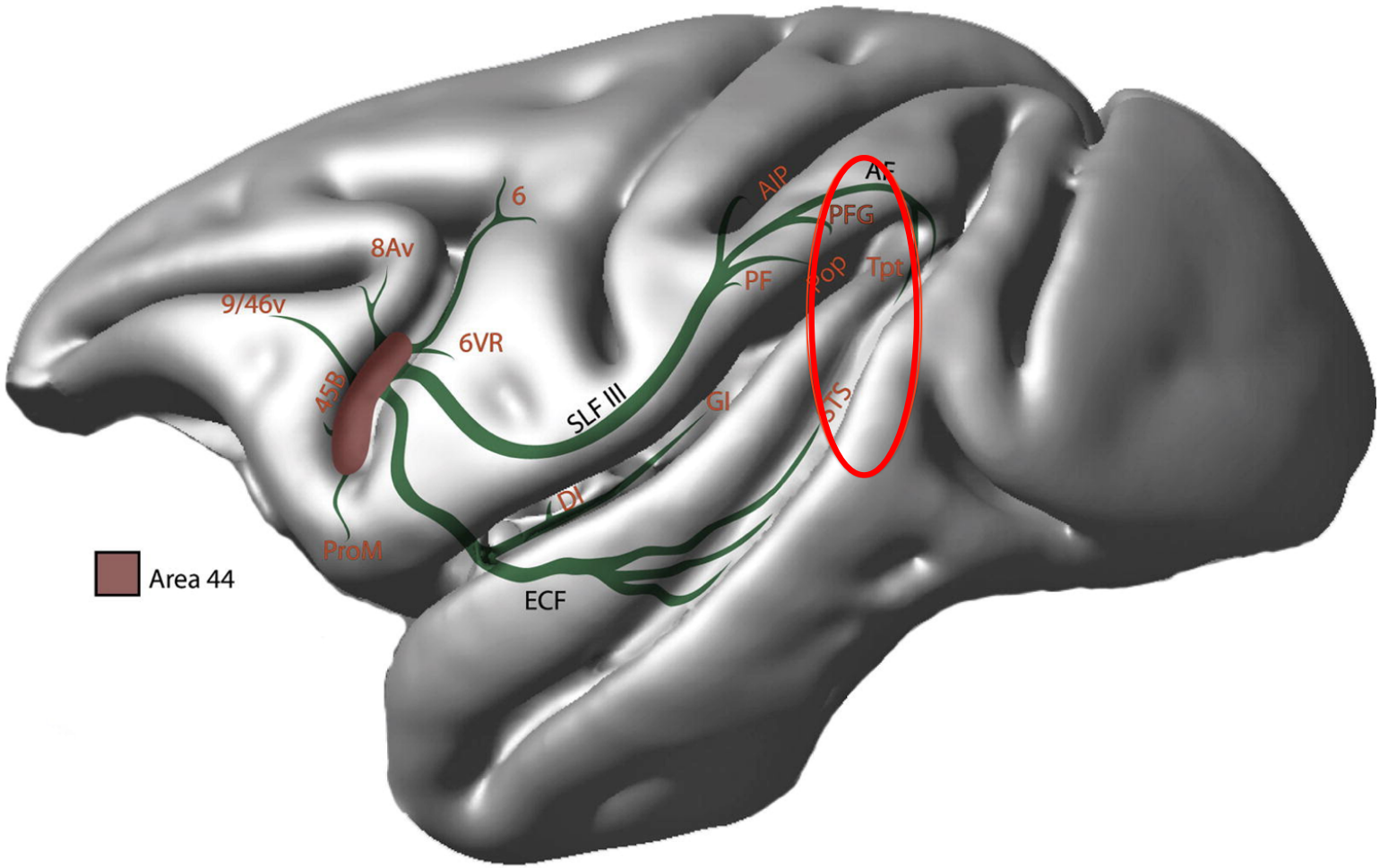
Fig. 1 | Structural connectivity between language regions. A schematic, condensed view of language-relevant brain regions and fibre tracts in the left hemisphere. The dorsal fibre tracts connecting the posterior temporal cortex (pSTG) with the frontal cortex involves the superior longitudinal fascicle and the arcuate fascicle. There are two streams with different termination points: one in the dorsal premotor cortex (PMC) (purple tract), and the other in BA 44 (blue tract). The ventral fibre tracts connecting the frontal cortex to the temporal cortex also consists of two streams: one going from BA 45 to the temporal, parietal and occipital cortex, involving the inferior fronto-occipital fascicle (pink tract); and the other going from the frontal operculum (FOP) to the anterior superior temporal gyrus (aSTG), involving the uncinate fascicle (dark grey tract). Adapted from ref. [24].





Macaque

Rilling *et al.,* Nature Neuroscience 2008*;* Frey *et al.,* Brain and Language, 2014


6
8Av
9/46v
6VR
45B
AIP
AF
PFG
PF
PGp
Tpt
SLF III
GI
STS
DI
ProM
ECF
Area 44

The rest of language – not just word segmentation
Insert punctuation to make readable:

The professors name is Bob The professor’s name is Bob

The is not always an article

I want to put hyphens between the words fish and and and and and chips in my fish and chips sign

I want to put hyphens between the words “fish” and “and” and “and” and “chips” in my fish and chips sign

Wouldn't the sentence I want to put a hyphen between the words fish and and and and and chips in my fish and chips sign have been clearer if quotation marks had been placed before fish and between fish and and and and and and and and and and and and and and and and and and and and and and and and and and chips as well as after chips

Wouldn't the sentence "I want to put a hyphen between the words Fish and And and And and Chips in my Fish-And-Chips sign" have been clearer if quotation marks had been placed before “Fish,” and between “Fish” and “and,” and “and” and “And,” and “And” and “and,” and “and” and “And,” and “And” and “and,” and “and” and “Chips,” as well as after “Chips”?

What about primates taught sign language?
Does Nim follow a grammar?
Why important?
“The continuity between primate language and early child language is believed to hold “the most promising guide to what happened in language evolution”
(Hurford 2011, p590)

Teaching Language to an Ape

Sarah, a young chimpanzee, has a reading and writing vocabulary of about 130 "words." Her understanding-goes beyond the meaning of words and includes the concepts of class and sentence structure

by Ann James Premack and David Premack

SCIENTIFIC AMERICAN
OCTOBER 1972
VOL. 227. NO.4 PP.92-99

Researcher Challenges Conclusion That Apes Can Learn Language

By DAVA SOBEL

Thirteen years of research involving gorillas and chimpanzees that were apparently taught human language have been thrown into question by new evidence that, according to a researcher, shows that the apes may be doing nothing more remarkable than a dog does in learning to sit or heel. Investigators are already at loggerheads in what promises to be a major dispute about whether the human capacity for language is unique.

Herbert S. Terrace of Columbia University, who spent four years teaching sign language to a chimpanzee named Nim Chimpsky, now asserts that the success of his own and related efforts can be explained as mere prompting on the part of the experimenters and mistakes in reporting the data. "Much of the apes' behavior is pure drill," he said. "Language still stands as an important definition of the human species."

But R. Allen Gardner of the University of Nevada, the first to train a chimp, Wa-

Nim's signs

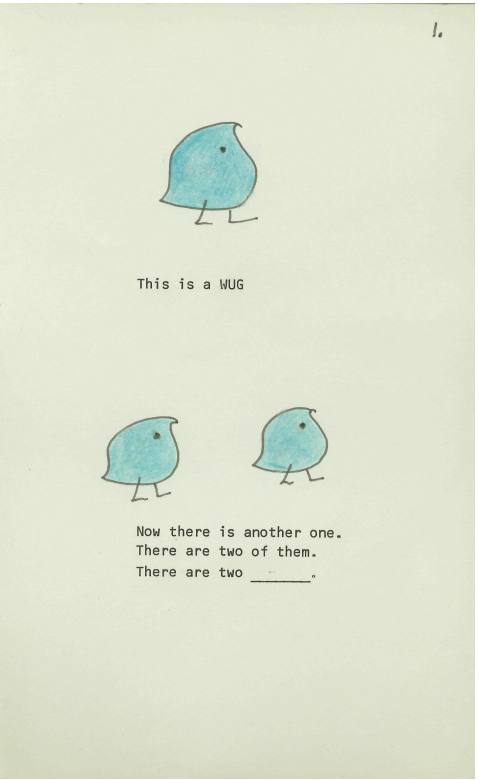
- Apple me eat
- Banana Nim eat
- Banana me eat
- Drink me Nim
- Eat Nim eat
- Eat Nim me
- Eat me Nim
- Eat me eat
- Finish hug Nim
- Give me eat
- Grape eat Nim
- Hug me Nim
- Banana Nim banana Nim
- Banana eat me Nim
- Banana me Nim me
- Banana me eat banana
- Drink Nim drink Nim
- Drink eat drink eat
- Drink eat me Nim
- Eat Nim eat Nim
- Eat drink eat drink
- Eat grape eat Nim
- Eat me Nim drink
- Grape eat Nim eat
- Grape eat me Nim
- Me Nim eat me
- Me eat drink more
- Me eat me eat
- Give orange me give eat orange me eat orange give me eat orange give me you

About 125 signs; thousands of multiple combinations, mostly 2-sign combinations
Like Det Noun rule, many look like "functor" e.g., "give", "more", combined with open class sign, like "apple", "Nim", or "eat"

Q: does Nim “know” human language?
Does Nim “have a grammar”?
Does Nim form rules?

Two views of language learning

Biology-driven: Open-ended, abstract syntactic <u>Rules</u>



Experience-driven, 'blank slate', usage-driven: Frozen "formulas", very specific, imitations

Evidence that kids form <u>rules</u> always, from the start
But how do we know?

Two views of language learning – with very different implications

- The constructionist or usage view: infants start with ”tabula rasa” and acquire fine little details, memorized word strings, repeating input from adults, and incrementally building up more and more general ‘templates’ or ‘formulaic patterns’ (“constructions”) – no pre-existing syntactic content attributed to child

e.g., Mother says: “more coffee”; Child: “more coffee”; M: “some coffee”; C: “some coffee”; but not “more juice” – no productive combinations until (perhaps much) later

- The biolinguistic view: infants start with knowledge of abstract syntactic categories like “Noun” and “Verb” and must learn details of how these categories behave, which words belong to which categories, etc.

Two views

- The two views lead to very <u>different</u> predictions about the trajectory of language learning, & what we expect to see kids do during language learning
- Constructionist: we ought to hear the kids repeat back exactly the words the adults say, very gradually forming matching what the parents say, certainly never something from another language
- If there is <u>no</u> input, then there ought to be <u>no</u> learning (blank slate + zero input → zero)
- Biolinguistic: we ought to hear very new things; child grammar and categories like adult grammar; we should hear things from other languages
- If there is <u>no</u> input, then we might see what the "initial state" looks like – how could we test this? (There is one way we'll see later.)

How could we tell?

- If language acquisition is more experience-driven, we'd expect children to "track the data"– follow the pattern of the examples they get, basically "input matching"
- If it's more internal, biologically-driven, we'd expect children to follows their "inner child" (which might mean some *other* adult language constraint, not attested by the input)
- Experience-driven: child English less articulated versions of "patterns" produced by adult English speakers
- Biologically-driven: child English differs from adult English in ways that adult languages differ from each other
- What does the evidence from child language acquisition tell us?

Rules vs. memorized “constructions” how can we tell them apart? Case 1: early child language (“two word” stage)

<u>Rule</u>

Give me milk

Gimme coffee (my favorite)

Gimme money (later child stage)

Gimme shelter (Winston’s favorite)

…

<u>Formula</u> (give + me) - tag team

Give me milk

Etc.

How to distinguish rules from frames/formulas? Errors kids make? No… (why?)

Litmus test to distinguish use of a grammar, i.e. "rule use" vs. "memorized sequences"

Apply to output from:

1. Children in 2-word stage (about 1 ½ – 2 yrs old)
2. Nim, the chimpanzee taught sign language from birth

We shall see that the test shows:

1. Children construct syntactic grammar rules, and have the abstract categories for them <u>from the start</u>
2. Nim does not use rules

But: Children don't just replicate the input data

English kids:

Want cookies
Where going?
How wash it?
Erica took
I put on

Are these "errors"? What's the pattern?

But we **never** see this kind of error: *Who see*?

There's another language that acts just like this; kids produce "want cookies"; "Erica took" etc.... but never "Who see". Which language? Ans:

Chinese: can drop subject/object (topic) if you can recover it from discourse

In this respect, English kids (all kids) start out "as if" they are speaking Chinese

What tells English kids <u>not</u> to do this?

- English kids eventually don't produces examples like Chinese (why not? What happens if they go to Dim Sum a lot?)
- What input tells them that every English sentence requires a Subject? Most of the time, there *are* Subjects that play a real role…but...
- "There are toys on the floor"
- (Spanish/Chinese: "are toys on the floor") Hay juguetes en el suelo / 有地板上的玩具 (Yǒu dìbǎn shàng de wánjù)
- 1% of English sentences: "There are ..."
- Once again: you need an initial, <u>universal</u> constraint <u>plus</u> some statistics...

Intuition behind detection of grammar/rule use
Let's focus on {a, the}

- Suppose Grammar rule: <u>productive</u>, freely used:

  a cookie, the cookie, a book → will also produce "the book"

  Free generation of the the two <u>parts</u> of the rule, Determiner + Noun
- Suppose "Template" or "usage": fixed and only parrots back what has been found in the input [a-cookie] or [a-thing-to-eat]

  a cookie, the cookie, a book → will <u>not</u> generalize to produce "the book" (because not heard yet)
- Expected Difference in observed output:

1. Rule Use: Will see <u>all</u> Nouns in child's vocabulary paired with both Determiners
2. Canned Template use: Will see only a <u>fraction</u> of the Nouns with both

A simple measure: overlap or “diversity”

- Let the child’s vocabulary be: {a, the} x {cookie, toy, bath}
- Suppose we observe kid say: a cookie, a toy, the toy, the bath (maybe multiple times)
- Then we say the <u>diversity</u> is 1/3 = 50% (because only “toy” is seen with **both** Determiners)
- If all the Nouns in the child’s vocabulary appear with both Determiners, then the diversity is 3/3 = 100% (fully productive rule use)
- If **no** Noun appears with **both** Determiners, the diversity is 0/3 = 0%
- Fully productive rule use = choose Determiner and Noun separately and combine them (but the observed frequency will depend on other factors)

Is diversity a good test to detect “rules” vs. “memorized frames”?

- Evidence for Frames or “Formula templates”? If frames, expect mismatch between children’s diversity and adults’ diversity (children’s diversity lower)
- Look at kids: 25% diversity, this seems low…
- So maybe they <u>do</u> just memorize...?
- Not so fast...

More careful study by Virginia Valian (2005) – child diversity is <u>not</u> low for *a, the*

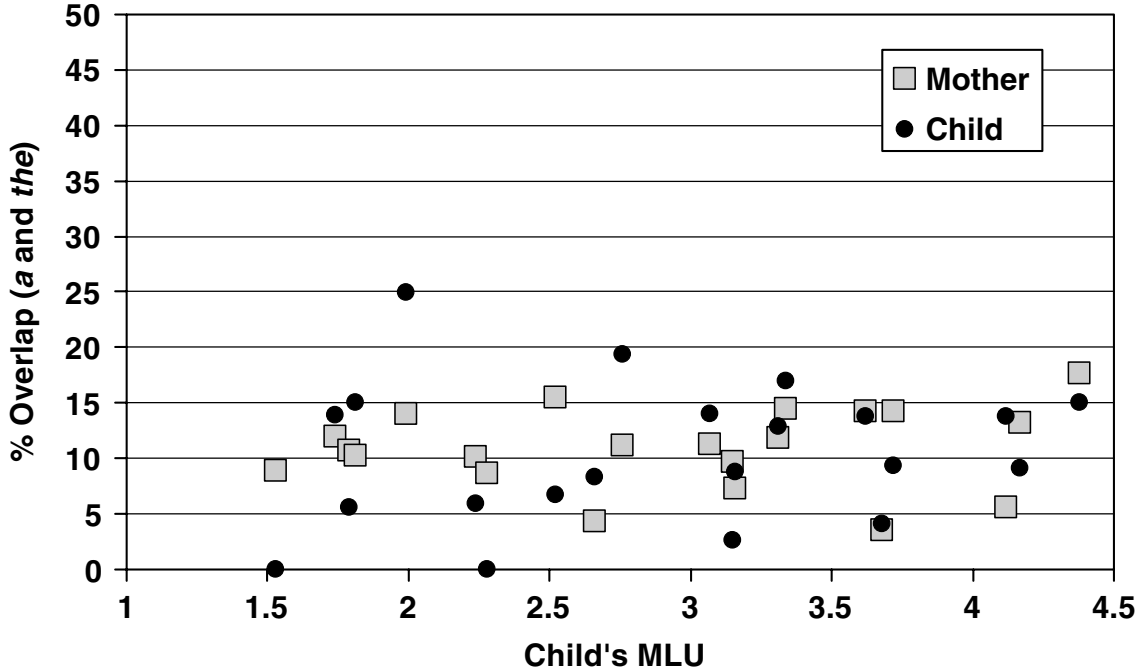


Fig. 1. Analysis 1: Overlap of *a* and *the* for children and mothers as a function of child MLU. (<u>m</u>ean <u>l</u>ength of <u>u</u>tterance)

(Valian et al. 2005)

All Determiners – mother, child overlap

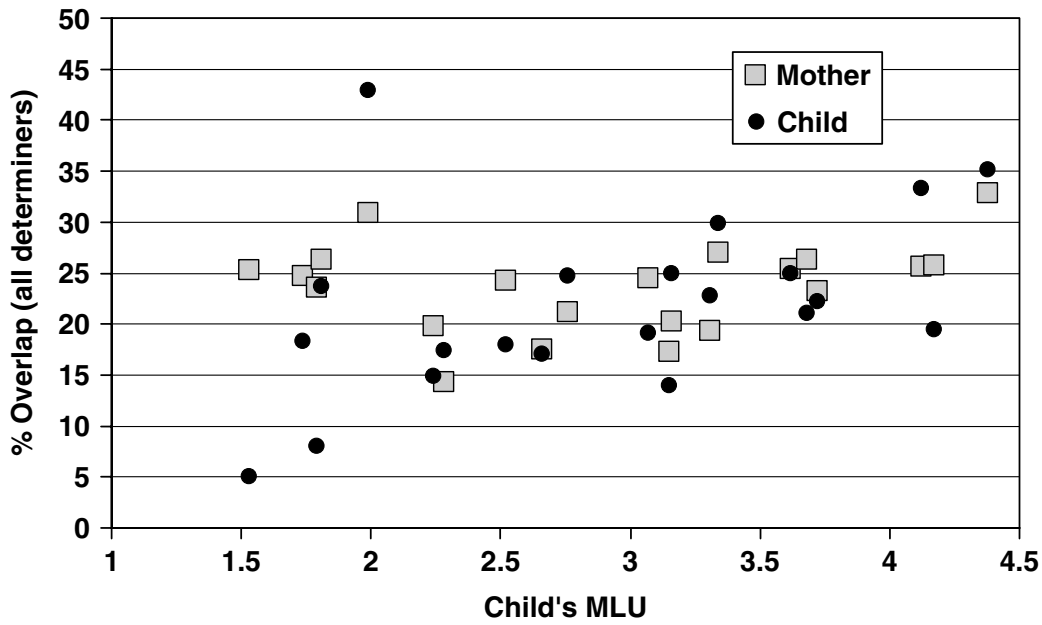


Fig. 2. Analysis 2: Overlap of all determiners for children and mothers as a function of child MLU.

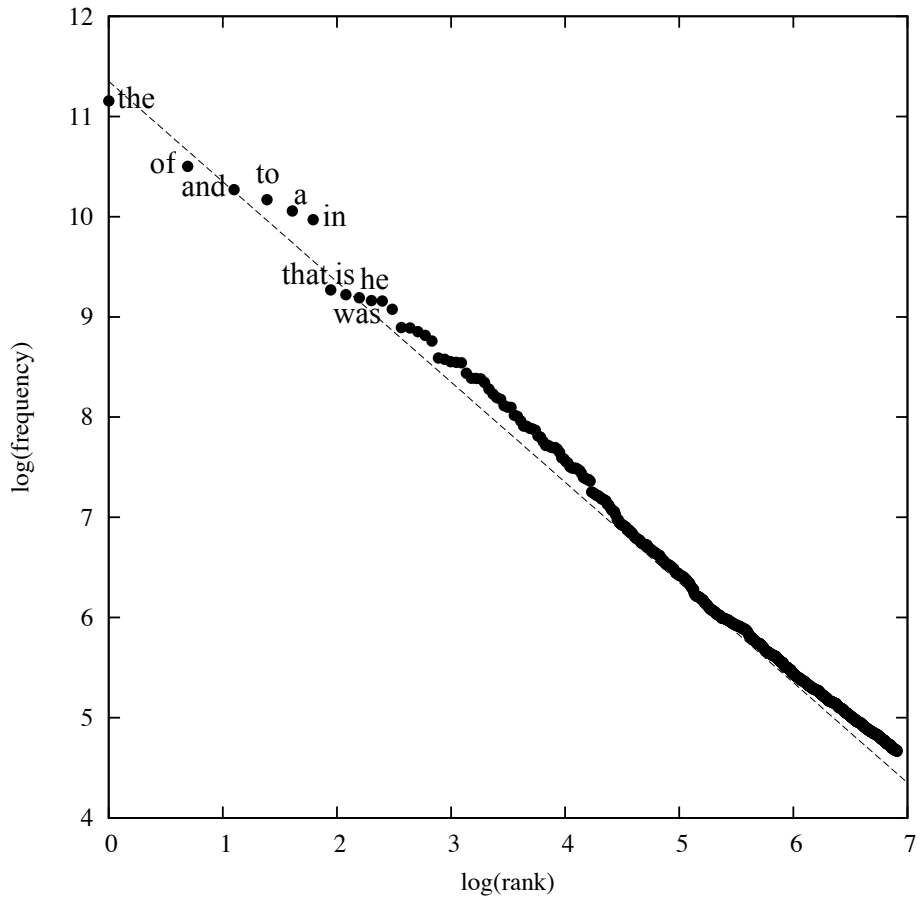
Further doubt

- Wait a sec...look at the Brown corpus of 1 Million English words; that is by professional writers
- Diversity: only 25% of singular nouns that combine with either “a” or “the” combine with <u>both</u>
- Why?

So: sampling effects can lead to <u>apparently</u> low diversity

- Asymmetry: Nouns favor certain determiners over others by a factor of 2.5
- E.g., “the bathroom” is much more frequent than “a bathroom”, but “a bath” is much more frequent than “the bath”
- This means the chance of pick both determiners can fluctuate, due to chance
- A noun has to appear at least twice for the diversity (or overlap) to be > 0
- We will adjust for this by using Zipf’s law to estimate the probabilities for a particular determiner or a noun

Zipf's law, word *frequency* x *rank* = *Constant*
rank is inversely proportional to word's frequency

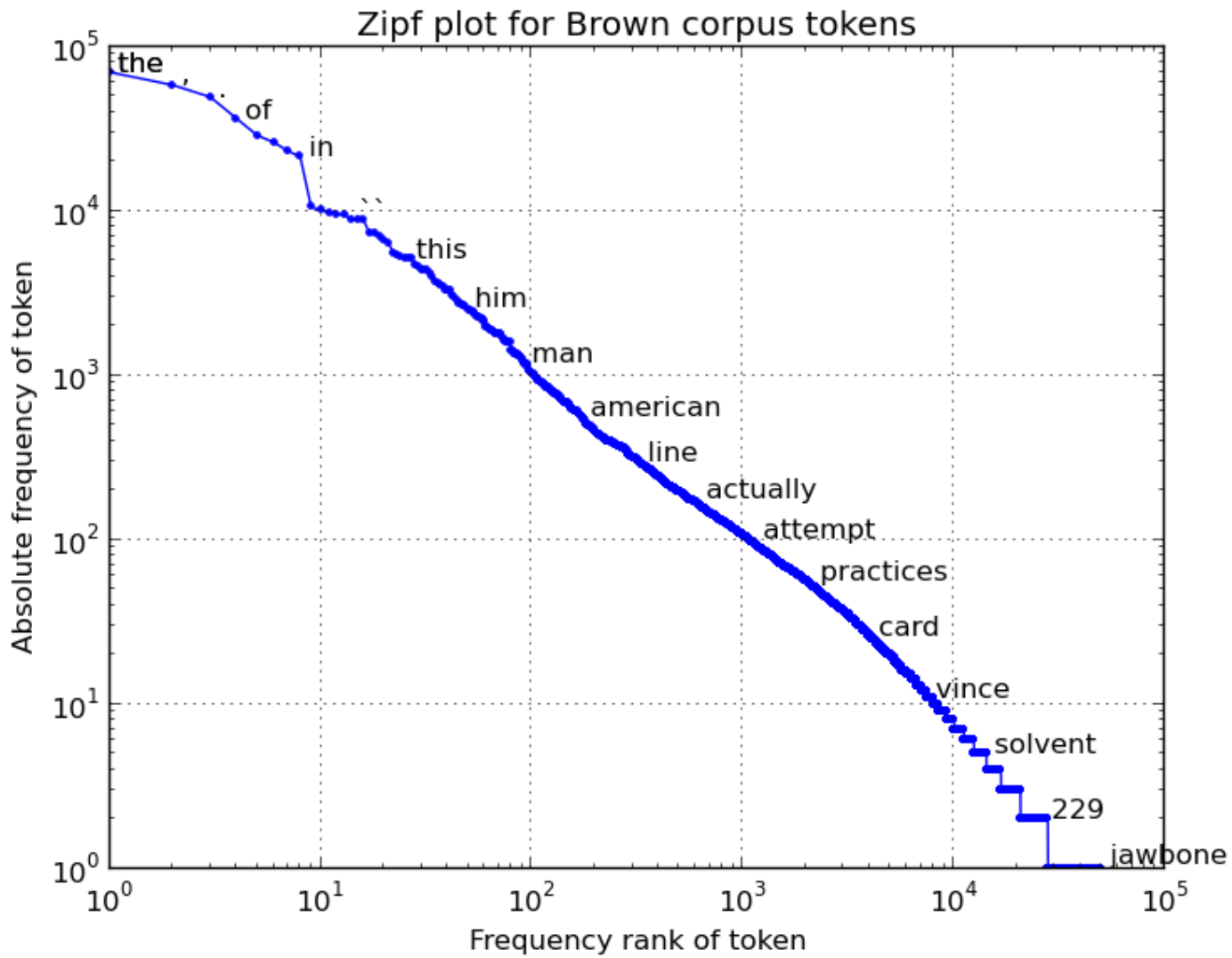


Measuring what you <u>say</u> vs. what you <u>know</u>

- If a kid knows 100 distinct Noun types, and pairs them with “a” or “the” 500 times, how many of the 100 will be paired with them both, assuming the rule Determiner Phrase → Determiner Noun?
- Many of the 100 nouns will fall on Zipf’s long and flat tail – appearing just once, so will never meet up with both Determiners
- Others might turn up 2 or 3 times, but still be monogamous (like getting 3 heads in coin tosses)

What you <u>say</u> vs. what you <u>know</u>

- The grammar isn't a fair coin anyway – sometimes a Noun has a favored determiner
- Imbalances: "the bathroom" more common than "a bathroom", but we say "a bath" more often than "the bath" (not a part of grammar, which probably doesn't encode the frequency of our biological needs) – but gives <u>impression</u> of grammatical absence
- Need statistical test of diversity profile based on the sample size and the vocabulary size (e.g., 500 and 100), <u>correcting</u> for Zipf-law sampling effects, to estimate probability of a particular Noun & probability of particular Determiner-Noun pairs


Zipf plot for Brown corpus tokens
Absolute frequency of token
Frequency rank of token
the
,
.
of
in
``
this
him
man
american
line
actually
attempt
practices
card
vince
solvent
229
jawbone
10^0
10^1
10^2
10^3
10^4
10^5

Using this to develop a statistical profile for use of a grammar rule

- What would it mean to use a rule <u>productively</u>?
- Combine Determiner + Noun freely – how to test this?
- This means picking items from each kind of category <u>independently</u>, as opposed to <u>dependently</u>
- For the case of 2 Determiners (*a, the*) and *N* Nouns, imagine three urns. One holds a blue marble (the Determiner, *a*); one holds a red marble (*the*) and the third holds *N* different green marbles, numbered 1, 2, …, *N* (the Nouns)
- You randomly pick one of the green marbles, and <u>independently</u> one of the blue or red marbles, each with some probability (all with replacement); this gives you one particular <u>independent</u> Determiner-Noun pairing; repeat *S* times