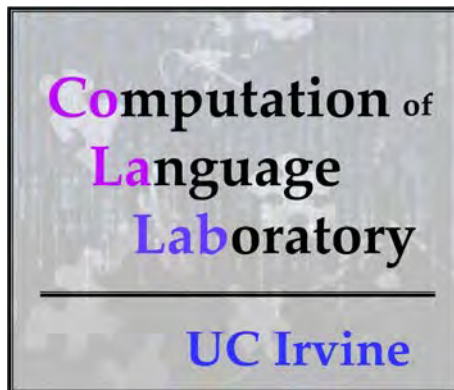# Frequency of what:
# How simple is the story of syntax acquisition?

Amanda Nili

In collaboration with Lisa Pearl

A Preview

- How do we know right from wrong, grammatically speaking?
- A simple theory of language acquisition: how often we hear something (a syntactic structure) determines how correct that structure is
- Is the simplest version of this story true?
- Looking at child-directed speech (because this is a study of acquisition)
- End result: syntax acquisition is not so simple (problems with the kind of frequency)

A Theory of Language Acquisition

- Relationship between frequency & grammaticality: If our brains are great computers (i.e., tracking statistics), they will be able to take in all of the data, analyze the frequency of every structure's use, then translate that to how acceptable the structure is, right?
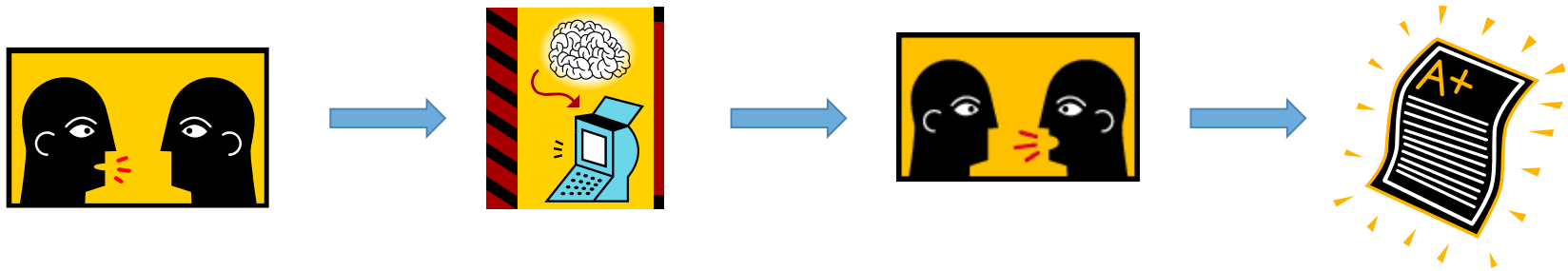
Frequency of what?

- Frequency of structure use

Frequency of what?

- Still, there are different ways of analyzing frequency:
    - (1) direct reflection (simple): we analyze the most surface forms and make minimal abstraction about the structure we hear
    - (2) something more sophisticated (less simple)

Frequency of what?

Subject (determiner + singular noun) + present tense intransitive verb.

This appears **20** times.

Ex: The pig grunts.

Frequency of what?

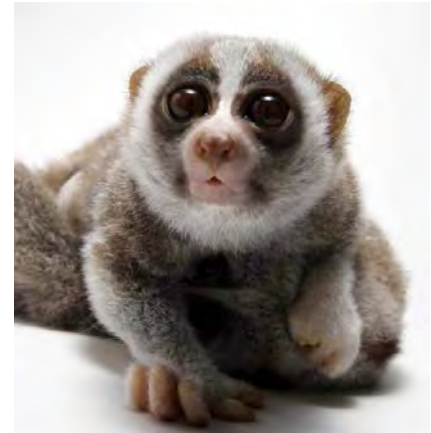*"Who" + auxiliary verb + subject (noun phrase) + transitive verb?*
This appears **43** times.
Ex: Who did Marvin poison?

Measuring grammaticality how?

- Grammaticality of that structure (how grammatical native speakers think it is) – often assessed by an acceptability score (an average of scores from multiple instances of each structure to control for semantic influence, which we hope reflects grammaticality in a well-controlled experiment)

Measuring grammaticality how?

- These acceptability scores are actually Z-scores (ranging from -1.19 to +1.13): they test standard deviation from the mean, and the direction of that deviation

Measuring grammaticality how?

Subject (determiner + singular noun) + present tense intransitive verb.

Acceptability score: **0.65**.

Ex: The pig grunts.

Reminder: Structure appears **20** times.

Measuring grammaticality how?

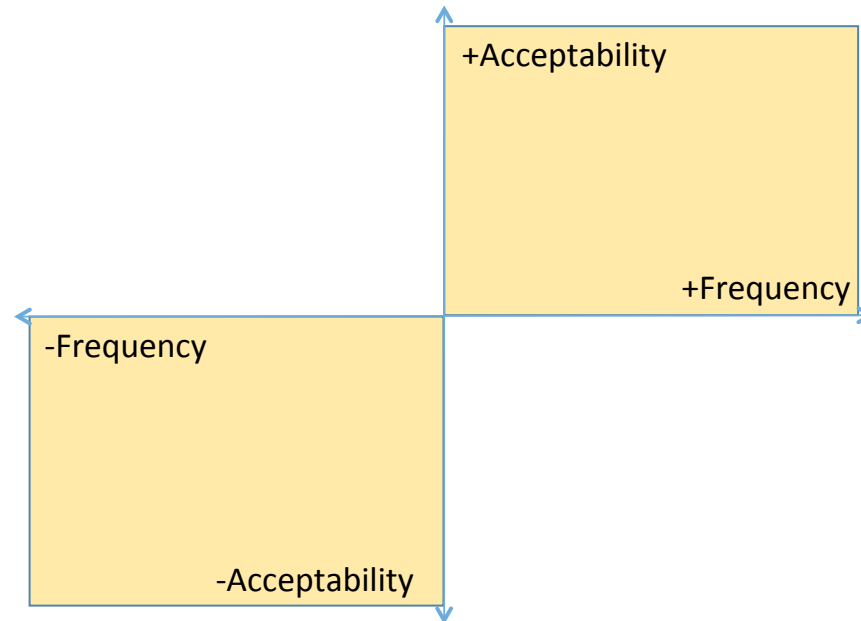*"Who" + auxiliary verb + subject (noun phrase) + transitive verb?*
Acceptability score: **1.12**.
Ex: Who did Marvin poison?

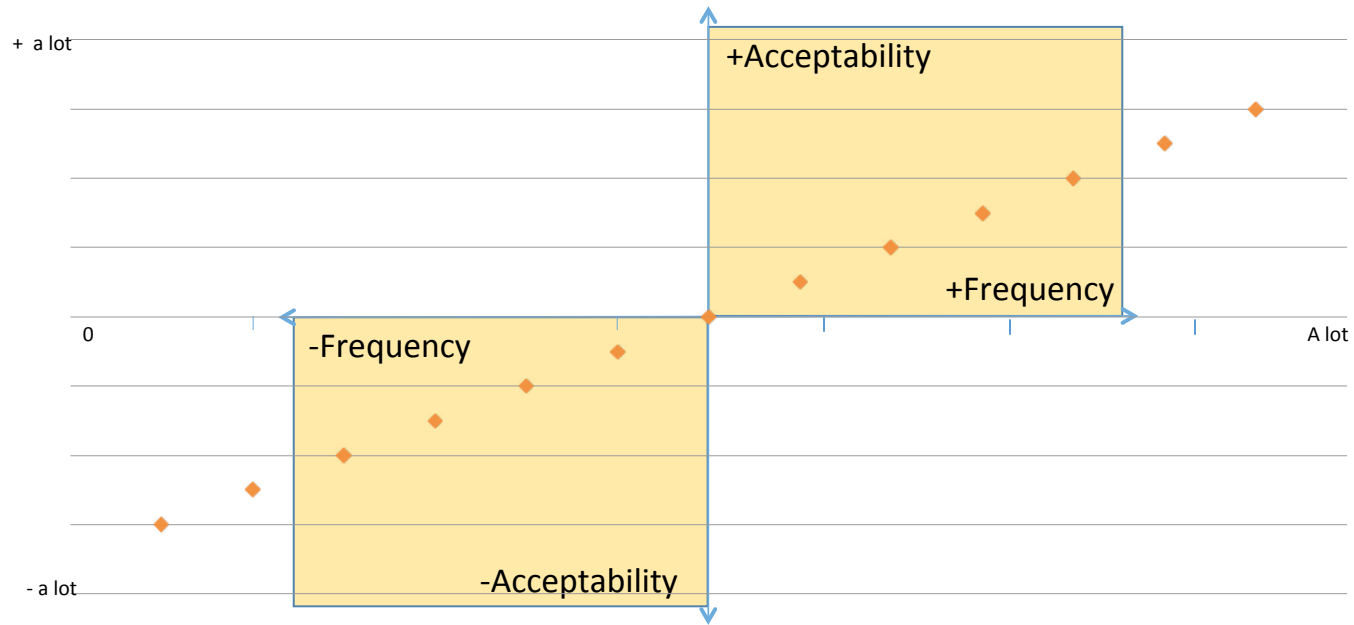Reminder: Structure appears **43** times.

The simple story would look something like this:

+Acceptability

+Frequency

-Frequency

-Acceptability
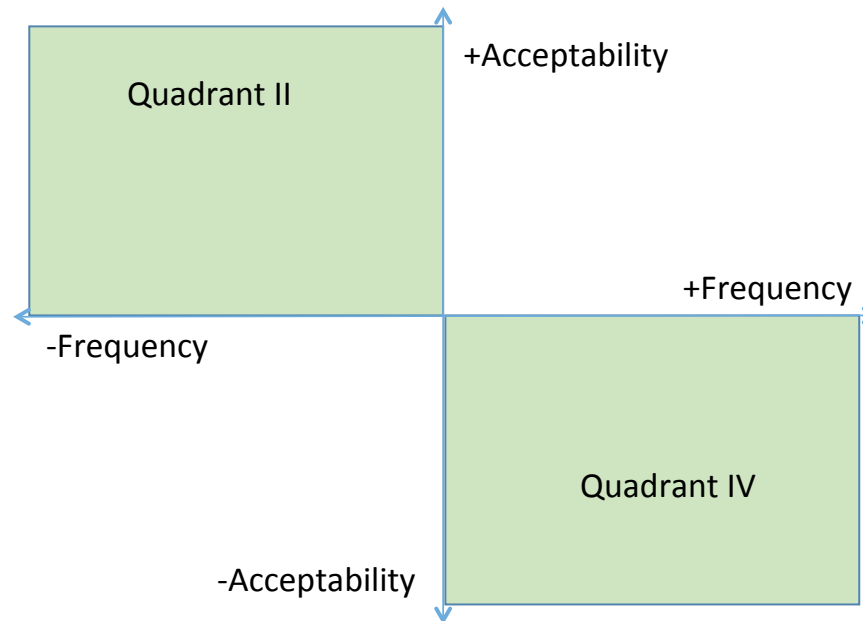
The simple story would look something like this:

Is the simple story true?

- Previous studies have found a gap: frequency values not matching perfectly (or even *well*) with acceptability scores
- Different explanations as to why that gap exists

The "gap" and other problems for the simple story:

Who believes the simple story?

- Kempen and Harbusch (2005): "[Object to findings of a frequency-grammaticality gap on the basis that] … quite a few orderings that are rated at least average to grammatical quality also have zero corpus frequencies."
  - Bad Method: The simple correlation is there, it's just that your naïve native speakers don't know what they find acceptable and what they don't, so they're rating utterances as more acceptable than they actually are!

Actually, the method is fine…

- Sprouse & Almeida (2012): different methods of collecting acceptability data accused of unreliability; a comparison of data from all methods shows that's not the case (all methods are pretty convincingly reliable)

Maybe it's not so simple…

- Jurafsky (2002): "… the mismatch between corpus frequencies and psychological norming studies is to be expected. These are essentially two different kinds of production studies, with different constraints on the production process."
  - When linguists try to compare these two data sets, they are doing something fundamentally incorrect: we don't find an r of 1 because, perhaps, acceptability data (from psychological norming studies) are expressive of something more abstract, whereas the frequency data are more concrete

Not so simple: Linear Optimality Theory

- Keller (2000): the acceptability-frequency relationship isn't so straight forward. It's about how significant a linguistic constraint (and its violation) is.

Not so simple: Linear Optimality Theory

- Hard constraint: Subject-Verb Agreement
  - Trish has painted a picture of Arthur.
  - *Trish have painted a picture of Arthur.

- Soft constraint: Definite Article Use
  - Which friend has Trish painted a picture of?
  - *Which friend has Trish painted the picture of?

Really not so simple

- Pearl & Sprouse (2013): these authors make the comparison between frequency and acceptability data.
  - Frequency of what? Really abstract representations (WH-questions).
  - Find no obvious correlation at that level of abstraction.

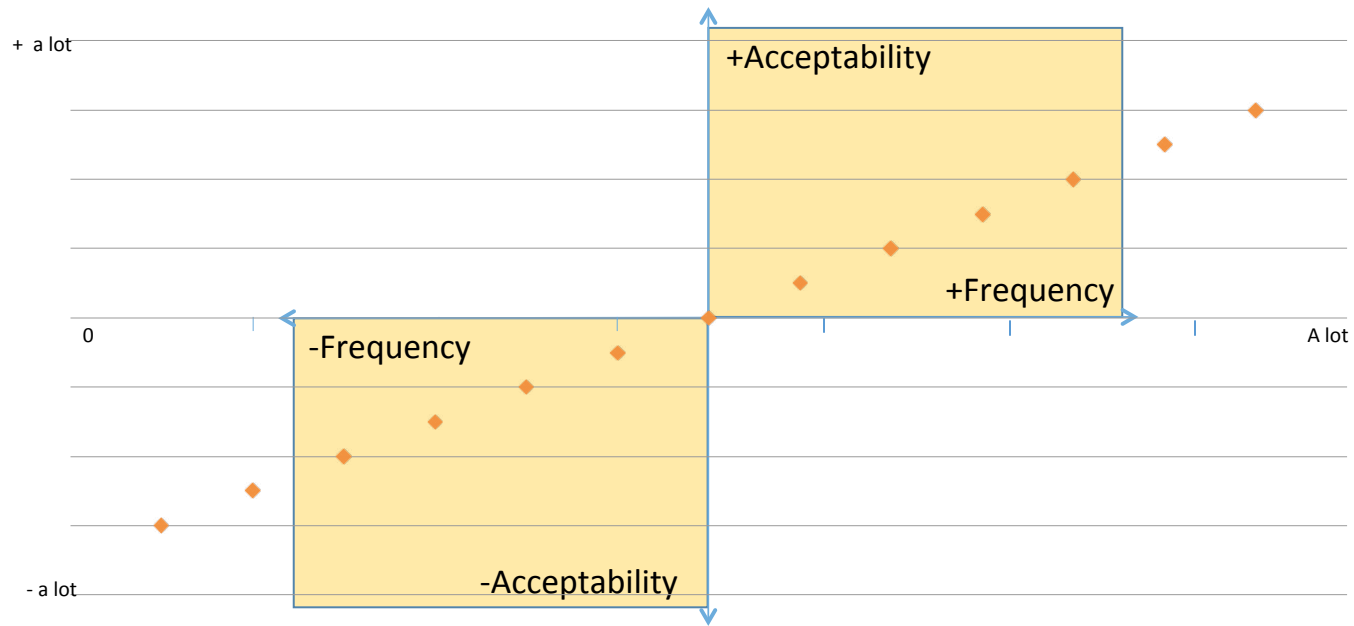[$_{CP}$ Who did [$_{IP}$ she [$_{VP}$ like _]]]?

Really not so simple

Note: Pearl & Sprouse's (in prep.) study of WH-questions has found a less-than-great correlation between adult-directed speech frequencies and grammaticality (as measured by acceptability).

'Frequency vs Acceptability'
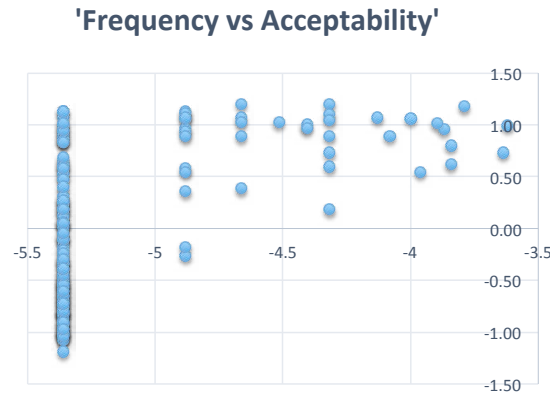
Not found (by Pearl & Sprouse):

What about child-directed speech?

(1) Important for understanding how we learn to have these grammaticality intuitions.

(2) Known differences at various levels between child-directed and adult-directed speech (e.g. motherese). Maybe there's a significant difference at the structural level, when looking at less abstract things. (Often there is a difference for sound distributions, words, and simple structures. Although Pearl & Sprouse find it does not apply for wh-dependencies).
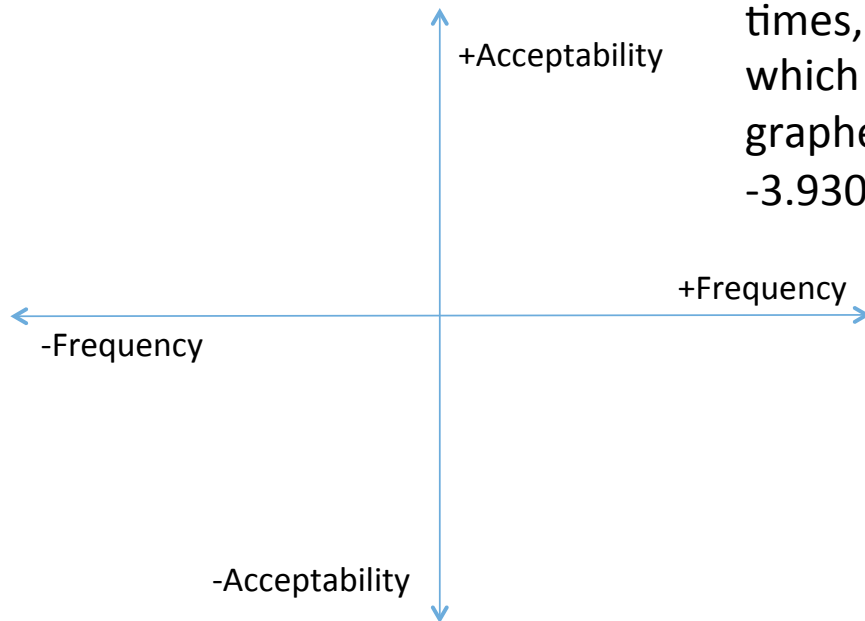
Acceptability Data

- Collected by Sprouse & Almeida (2012), using utterances in the Adger's *Core Syntax* textbook for linguistics students
  - Scores assigned by naïve native speakers using multiple data-collection methods
  - Each structure presented multiple times, using different words:
    - Collected averages of the multiple iterations => one averaged score per structure

Frequency Data

- Child-directed utterances from CHILDES data base, the following corpora:
  - Brown-adam (26,280 utterances): ages 2;3-4;10
  - Brown-eve (14,245): 1;6-2;3
  - Brown-sarah (46,948): 2;3-5;1
  - Soderstrom (21,334): 0;6-1; 0
  - Suppes (35,906): 1;11-3;11
  - Valian (25,550): 1;9.20-2;8.24

   Total => 170,263

Frequency values are negative because the normalized values are very small numbers, so each is the $\log_{10}$ of the calculated frequency.

For example: "The pig grunts." Appears 20 times, has a frequency score of 0.000117465, which we take $\log_{10}$ of, to get the more easily graphed (negative) value of -3.930090286.
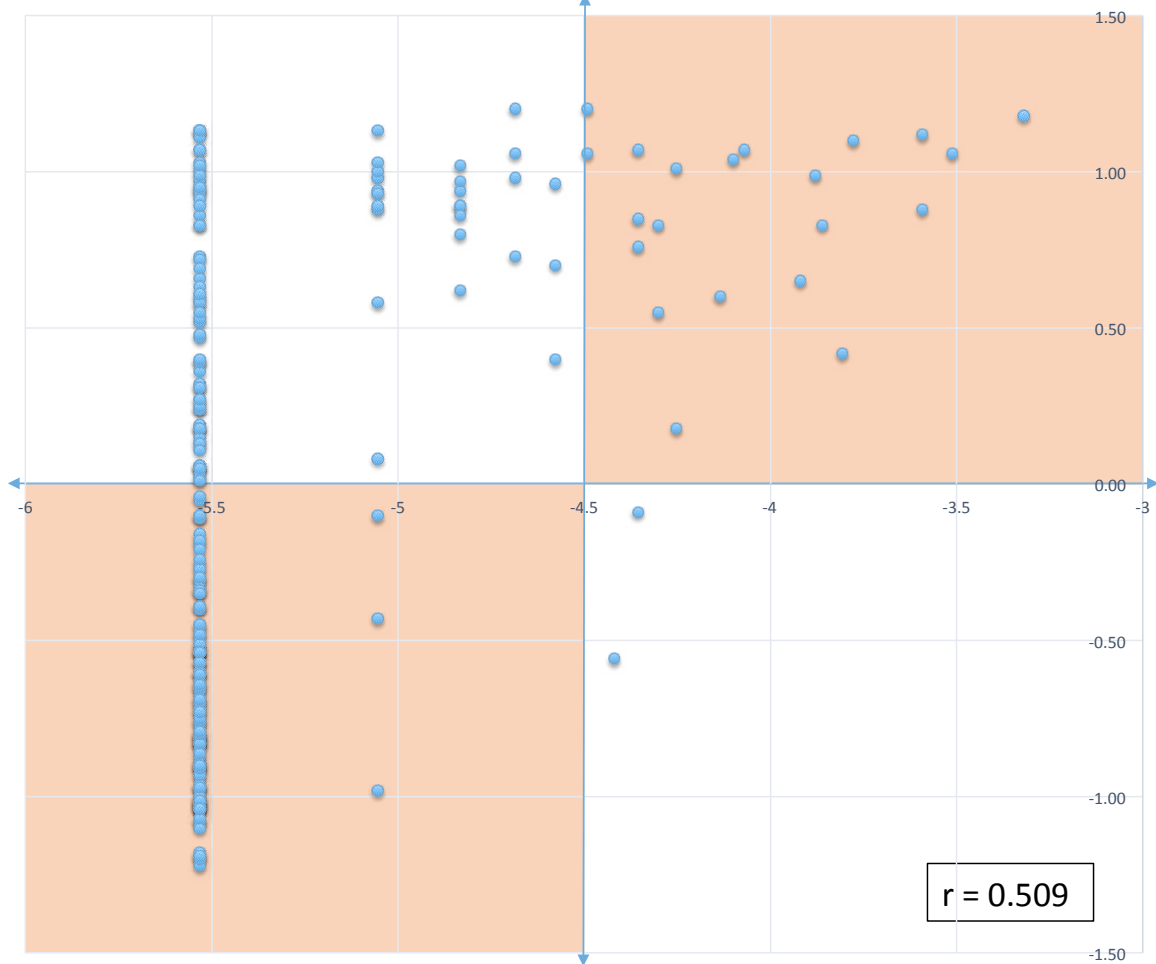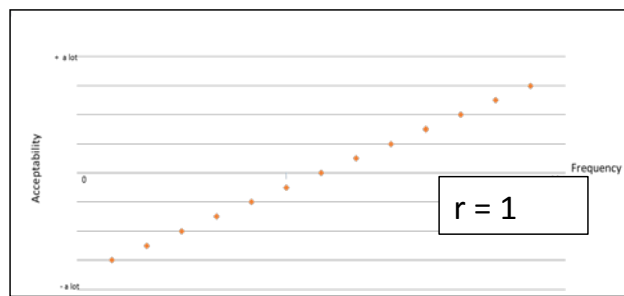
+Acceptability

+Frequency

-Frequency

-Acceptability

The expectation:
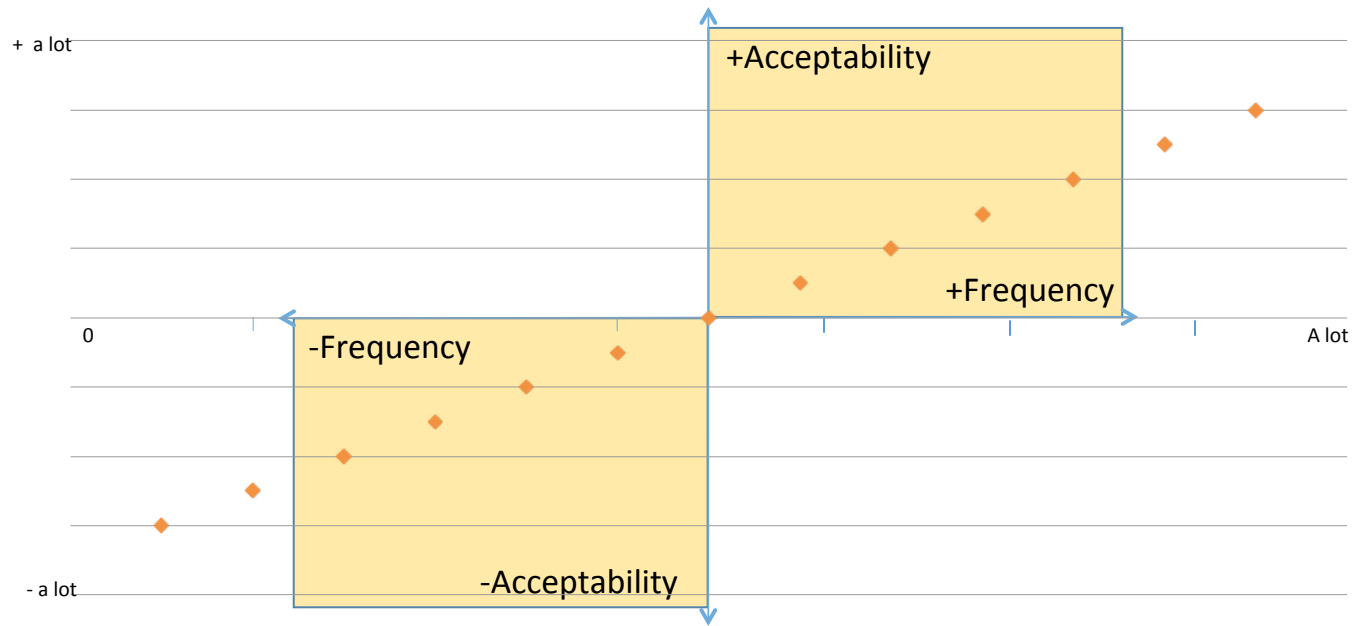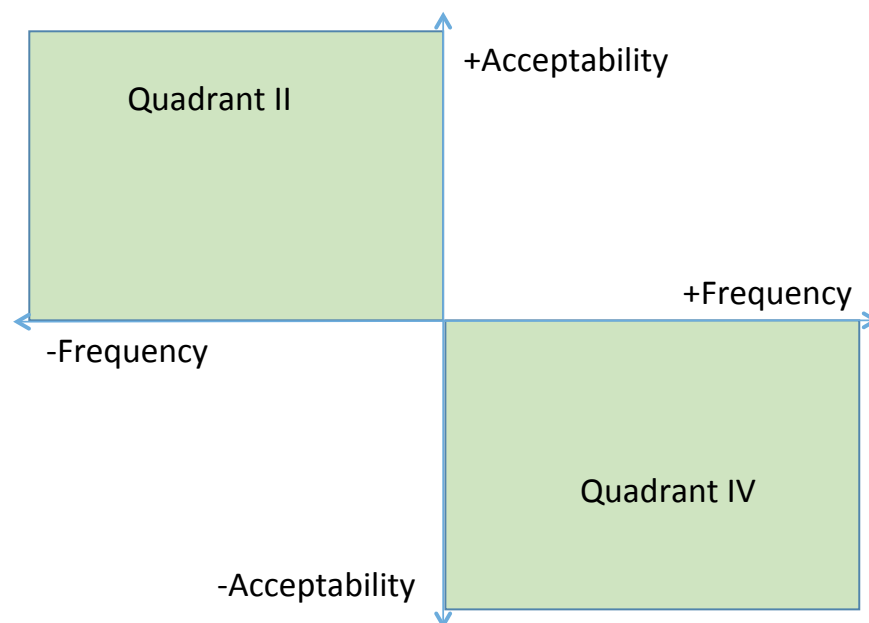
Frequency v Acceptability

Not Pictured:

r = 1

Frequency values are negative because they are $\log_{10}$

r = 0.509

- Comparing Sprouse & Almeida's 2012 study of acceptability scores [of 219 structures] rated by naïve native speakers with the actual frequency of utterance appearance in the CHILDES corpora (using child-directed speech only) results in some questionable data:

High acceptability and low frequency

- High acceptability (0.80) and low frequency (occurs 2 times): *Subject (nominative pronoun) + "have"-auxiliary +transitive verb past participle + object (accusative pronoun)*

Example: She has kissed her.
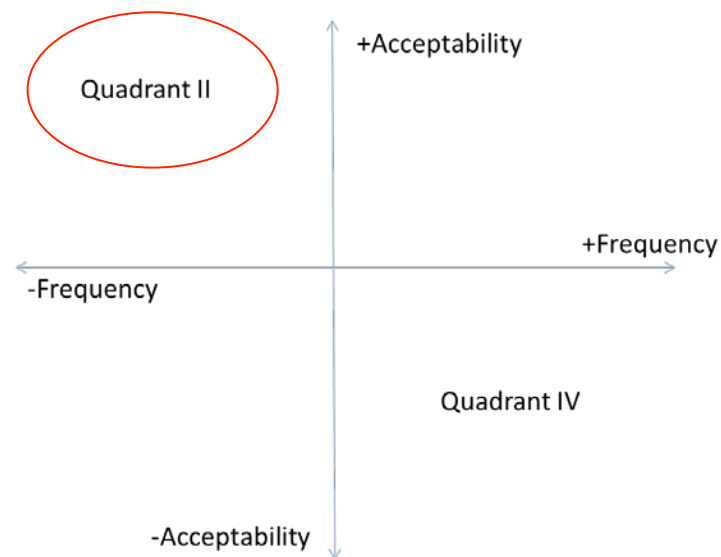
Low acceptability and existent frequency

- Low acceptability (-0.09) and existent frequency (occurs 7 times): *NP = [singular count noun with no determiner] + verb + PP, nothing after*

Example: *Letter is on the table.

- Consider that a high acceptability utterance ("Joss's idea is brilliant") has the *same* raw frequency score, but an acceptability score of 1.07.

+Acceptability

Quadrant II

+Frequency

-Frequency

Quadrant IV
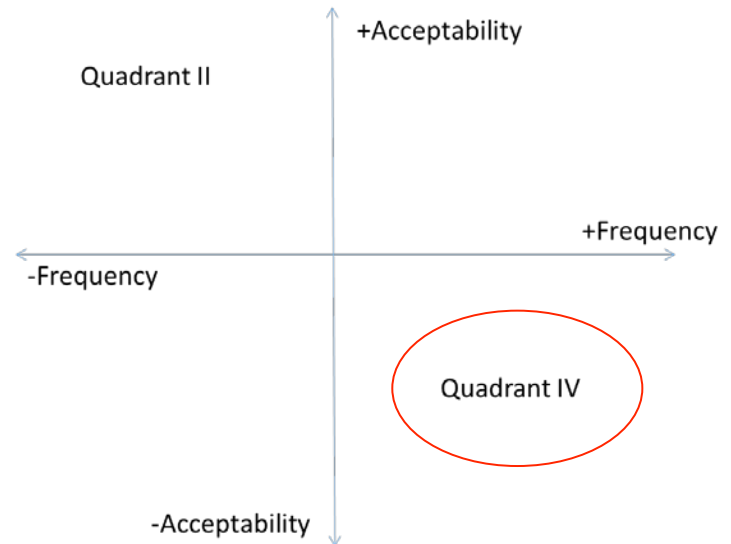
-Acceptability

Low frequency (occur 0 times) and varying degrees of low acceptability:
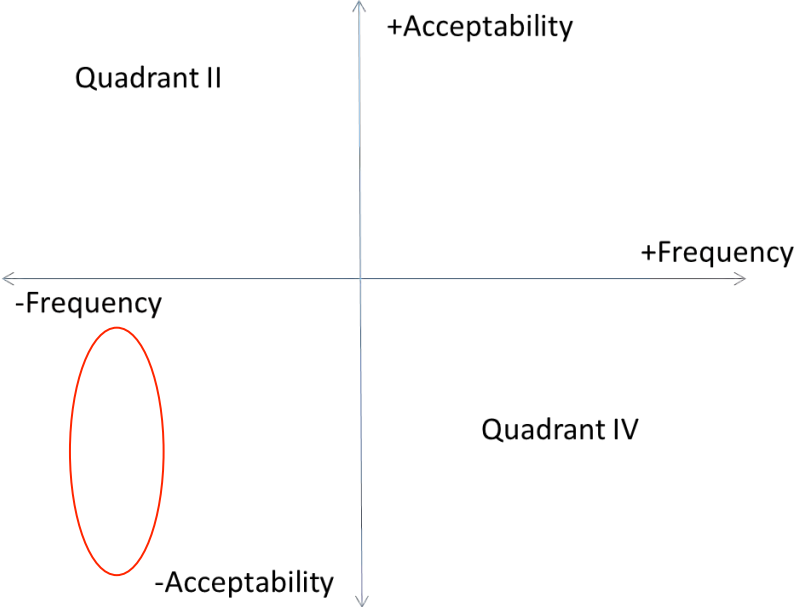
Low frequency (occur 0 times) and varying degrees of low acceptability:

- *Subject + tensed verb "wonder" + wh-object fronted + subject + auxiliary + transitive verb?*

  Example: I wondered who did Marvin poison?

  - Acceptability score: **-0.18**

- *Subject (name) + be + object (plural noun).*

  Example: Peter is pigs.

  - Acceptability score: **-1.20**

The data suggest:

- Frequency of simple structures is NOT the only factor to determine how acceptable structures are to native speakers
- (Not the simple version)

Future directions for research:

- Child-judgment data
- The different levels at which we study frequency

Future directions for research: Child judgment data

- Children learning a first language *very likely* don't perceive that language the same way adults do, but we compare frequency of linguistic input against adult judgment data – this may not be telling us enough about how children learn what's acceptable and what's not

Future directions for research: Different kinds of frequency

- Frequency of what?
  - Different levels of abstraction may be necessary for the different kinds of unacceptable utterances
  - Example: "The book ran." v "The thief ran."
    - We accounted for semantic category violations such as "The book ran." However…
- Consider: some of the utterances may show correlation between frequency and acceptability, because less abstract level of comparison is appropriate—other utterances may require something more abstract in order to correlate their acceptability scores with their frequencies.

About that question…

- Can we account for syntax acquisition with a simple story, or do we need something more sophisticated (a different understanding of frequency) to make sense of the data?

Frequency of what?:

- We know that base frequencies of structure usage don't correlate well with acceptability judgment data
- This could be because base frequencies are only part of the story
- Future research should focus on the frequency of what: determining how abstracted our information about syntax is

"The penguins."
 DT   NNS
     NP
DT-bird
NP-animate

A special thanks to:

- Lisa Pearl
- The Computation of Language Laboratory
- UROP