Why grammar and processing are sensitive to different event properties?

Eunkyung Yi, Doug Roland, Hongoak Yun, Hongmo Kang, Gail Mauner, and Jean-Pierre Koenig

Where we are

- Semantic similarity *of word senses* affects retrieval (shared features activation model)

- Semantic similarity *between words* affects choice of syntactic frame (shared features priming and verb-frame association model)

Where we are going

- Semantic similarity is quite distinct from the kind of semantic properties that we make reference to as linguists
  - Continuous measure (rather than categorical)
  - Gradient effects (rather than categorical)
- The difference is more general: semantic properties that matter to processing ≠ properties that matter to grammar, it seems

WHY (if it is not an illusion)?

Outline

1. What kinds of properties matter to the grammar of events
2. Semantic similarity of verbs affects production (direct manipulation)
3. Plausibility (another continuous conceptual variable) affects comprehension and differs from predictability
4. Semantic similarity of dependents affects comprehension and differs from predictability
5. Why are grammars and human sentence processors sensitive to different kinds of properties?

What the grammar of event is sensitive to

- How event participants are encoded is sensitive to rules

  1. #The wall hit the horse

- The “rules” tend to be rather stable cross-linguistically with some variations

- The rules are sensitive to (1) a limited set of semantic properties and (2) are based on “lexical entailments” and (3) (typically) do not gang up

Examples of event grammar rules

- If a verb (sense) describes a causal event, the cause is the subject and the affected entity the object
- If a verb (sense) described a mental representation the experiencer is the subject and the mental representation the object
- ….

Events properties that matter to Grammar (I)

(For subject/object selection, valence alternations, or case assignment.)

- Causality (including intermediary causes=tools);
- Volition;
- Mental representation;
- Change of state;
- Control (more or less in control);

Events properties that matter to Grammar (II)

- Affectedness;
- Modality (including negation);
- Boundedness;
- Contact;
- Motion;
- …

The two-level theory

- Why are *these* properties relevant? Why are *only* these properties relevant? Why are these *kinds* of properties relevant? Why is there no *ganging up*?

"Perhaps there is a set of semantic elements and relations that is much smaller than the set of cognitively available and culturally salient distinctions, and verb meanings are organized around them" (Pinker, 1989)

- – Grammar rules cannot "see" differences among verbs other the ones we listed;
- – Children's hypothesis space for learning valence alternations is constrained;

Be ware of limited samples

- The range of properties is limited by our current knowledge
  - There may be “exotic” properties that are not part of the current list of properties relevant to syntax;
- Example, absolute and relative age in “subject” and “object” assignment in Oneida (Iroquoian)

Generational subject/object selection

(4) **lo-**nulhá·
3ZOIC.SG>3MASC.SG-mother
‘his mother’

(5) **luwa**-yʌha
3FEM.SG>3MASC.SG-child
‘her son‘

- Rule I (refers to generation): The argument that corresponds to the older generation maps onto the “subject,” while the argument that corresponds to the younger generation maps onto the “object.”

Absolute age subject/object selection

(6) **lake**-ʔkʌha
3MASC.SG>1SG-sibling
'my older brother'

(7) **khe**-ʔkʌha
1SG>3FEM.SG-sibling
'my younger sister‘

- Rule II (refers to age, not generation): The argument that corresponds to the older person in a kin relation corresponds to the agent prefix, while the argument that corresponds to the younger person corresponds to the patient prefix.

Oneida’s event grammar is radically different

- Koenig and Michelson (2015a, 2014):
  - Oneida does not make use of syntactic features (particularly categorial features)
  - Every construction above the word level only has a semantic composition component
  - There is no notion of argument realization, but there is “agreement” inflection on most words

Basic distinctions and distribution of pronominal prefixes

- 58 portmanteaux-like pronominal prefixes
- Transitive prefixes mark 2 animate arguments of dyadic (or triadic) verbs

(1) **shukwa**-hlo·lí-heʔ

3Msg>1pl-tell-ASP

‘he tells us'

Basic distinctions and distribution of pronominal prefixes

- Intransitive prefixes mark single animate arguments of monadic verbs
- Two categories: A(gent) and P(atient)

(2) **lo**-nu·sé-heʔ
3MsgP-lazy-ASP
'he is lazy'

(3) **wa**-ha-ya·kʌ́-neʔ
FACT-3MsgA-go.out-ASP
'he went out'

Linking rules in Oneida are quite similar to that in other languages

- Generation and age-sensitivity is certainly different
- But, other “linking” rules to agreement structure are the same as those found in English for argument structures

What characterizes semantic information relevant to structure

- Limited number of properties matters
- Related almost always to “lexical entailments”
- So, notwithstanding Oneida and other less studied languages, the generalization is correct about restriction on the *number*, *identity*, *kind* of rules relevant to the grammar of events and no ganging up

What matters to on-line sentence processing…

- Reduced relatives garden-pathing is affected by how good/ bad a patient the subject is:

  The horse raced past the barn fell

(8) The shrewd heartless gambler (who was) manipulated by the dealer had bid a lot more money than he could afford to lose.

(9) The young naive gambler (who was) manipulated by the dealer had bid a lot more money than he could afford to lose.

Why such a difference?

- The *number*, *identity*, and *kind* of information relevant to processing and grammar are quite different
- Why is there such a difference?
  - Grammar = Quasi-innate partitioning of conceptual space
  - Grammar = Subset of processing routines that develop out of frequency
  - …

What’s a nice man doing in such a place?

- It could be that the linguists in us:

    1. Are treating as “lexical entailments” what is really gradient
    2. That the categorical nature of grammatical rules are an illusion
    3. That the small number and kind of properties are the result of ethnocentrism
    4. That grammars are built consciously and this is how our conscious mind works

- True, let’s build comparable systems and let’s compare (cf. discriminative learning models of inflectional morphology vs. Information-based Morphology)

Picking up where we left off

- Semantic similarity is not one of the properties that matter to the grammar of events
    - Semantic similarity between verb senses matters to lexical access
    - Semantic similarity between verbs affects *choice* of syntactic frame
- Today: it matters to sentence processing (both production and comprehension)

Can semantic similarity modulate syntactic priming?

- Research I discussed last time bears on the effect of *stored* associations on syntactic frame selection

- What happens when you hear a verb used in the ditransitive or prepositional frame? Does the semantic similarity between that verb and the verb you are to use affect your syntactic choice?

    - Can semantic similarity between verbs modulate syntactic priming?

Syntactic priming

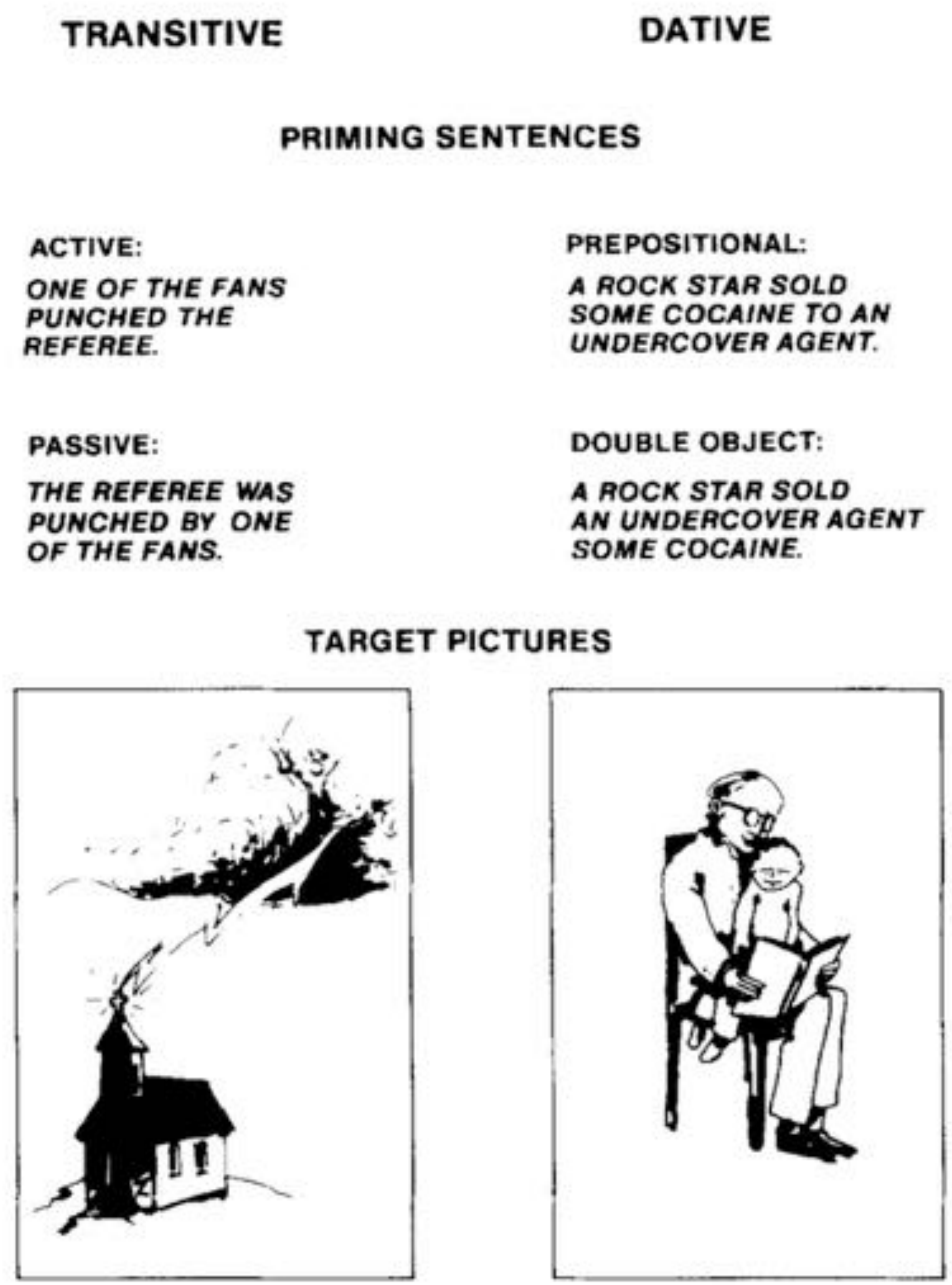


FIG. 1. Examples of transitive and dative priming sentences and target pictures used in Experiment 1. Only one of the two alternative priming sentence forms was presented on each priming trial, followed by a target picture. Note that the target pictures can be described with either of the two primed syntactic forms, as in *Lightning is striking the church* or *The church is being struck by lightning* for the transitive picture, and *The man is reading a story to the boy* or *The man is reading the boy a story* for the dative picture.

Semantic similarity and syntactic priming on A+N combinations

- Cleland and Pickering (2003) showed that semantic similarity affects choice of pronominal adjective or relative clauses
- Cards described by a confederate and a participant
- Red sheep/Red knife/Red goats (or their RC equivalents) as primes, “Red sheep” card as target
- 47% priming if same nouns for prime and target, 31% priming if semantically related nouns, 8% if semantically unrelated nouns

Stimulus set from Cleland and Pickering

For example, one item consisted of the pairing of (2a–f) with a picture of a red sheep:

2a. The red sheep (same noun, pre-nominal)
2b. The sheep that's red (same noun, relative clause)

2c. The red goat (semantically related noun, pre-nominal)
2d. The goat that's red (semantically related noun, relative clause)

2e. The red knife (unrelated noun, pre-nominal)
2f. The knife that's red (unrelated noun, relative clause)

Does semantic similarity of verbs affect priming of argument structure constructions?

Two pairs of syntactic priming experiments

- Manipulated *verb similarity* between prime and target
- Sentence reading & recall paradigm
- Two argument structure alternations

DATIVE ALTERNATION (Experiments 1-2)

John gave ***his son*** a toy. (Double Object: **DO**)

John gave a toy *to* ***his son***. (Prepositional Object: **PO**)

LOCATIVE ALTERNATION (Experiments 3-4)

John loaded ***the truck*** with hay. (Location Object: **LO**)

John loaded **hay** on *the truck*. (Material Object: **MO**)

Verb semantic similarity in dative alternation

EXPERIMENTS 1 & 2

TARGET SENTENCE

The producer **promised** a large part to the actress

PRIME SENTENCES

HIGH	The CEO **guaranteed** all employees a Christmas bonus
LOW	The ball boy **bounced** the player a new ball
CTRL	Organic food is increasing in popularity recently

HIGH similarity vs. control **>** **LOW** similarity vs. control

Verb semantic similarity in locative alternation

EXPERIMENTS 3 & 4

TARGET SENTENCE

The kid **smeared** mom's lipstick on her face

PRIME SENTENCES

HIGH	The New Yorker **spread** a toasted bagel with cream cheese
LOW	The freight driver loaded the huge truck with lots of boxes
CTRL	The congressman decided to run for the next election

HIGH similarity vs. control > LOW similarity vs. control

Verb similarity norming

STEP 1 LATENT SEMANTIC ANALYSIS

HIGH similarity pair > **LOW** *similarity* pair

STEP 2 JUDGMENT EXPERIMENT

On a Likert scale (1~7)

	Dative	Locative
HIGH similarity pairs	5,55	4,96
LOW similarity pairs	1,81	1,67

A trial: Reading & Recall

(Potter & Lombardi, 1998; Griffin & Weinstein-Tull, 2003)

READING

- The producer promised a large part to the actress — **PO**
- The CEO guaranteed all employees a Christmas bonus — **DO**

RECALL

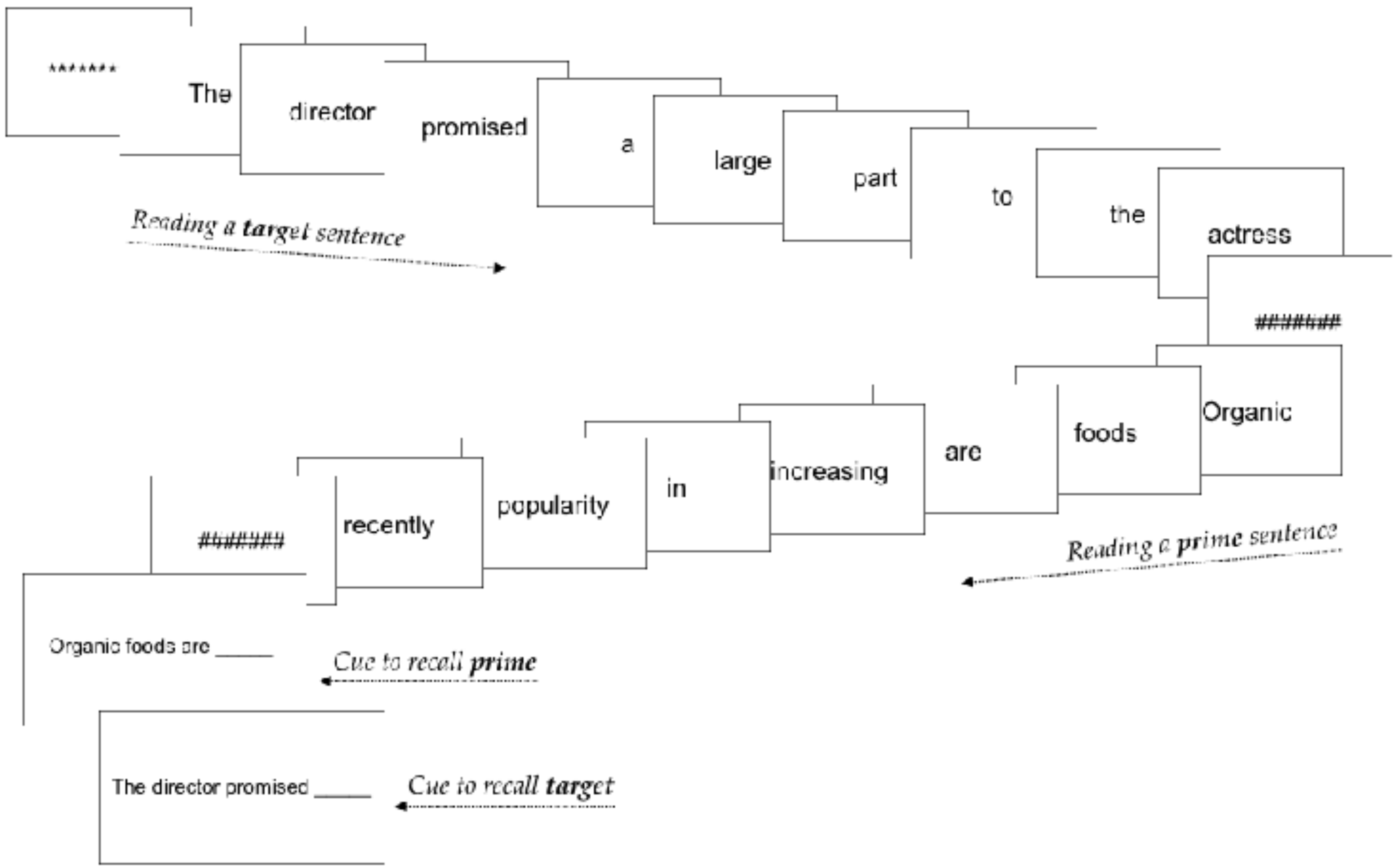
- The CEO guaranteed ___all employees a Christmas bonus___ — **DO PRIME**
- The producer promised ___a large part to the actress___ / ___the actress a large part___ — **TARGET** PO *or* DO

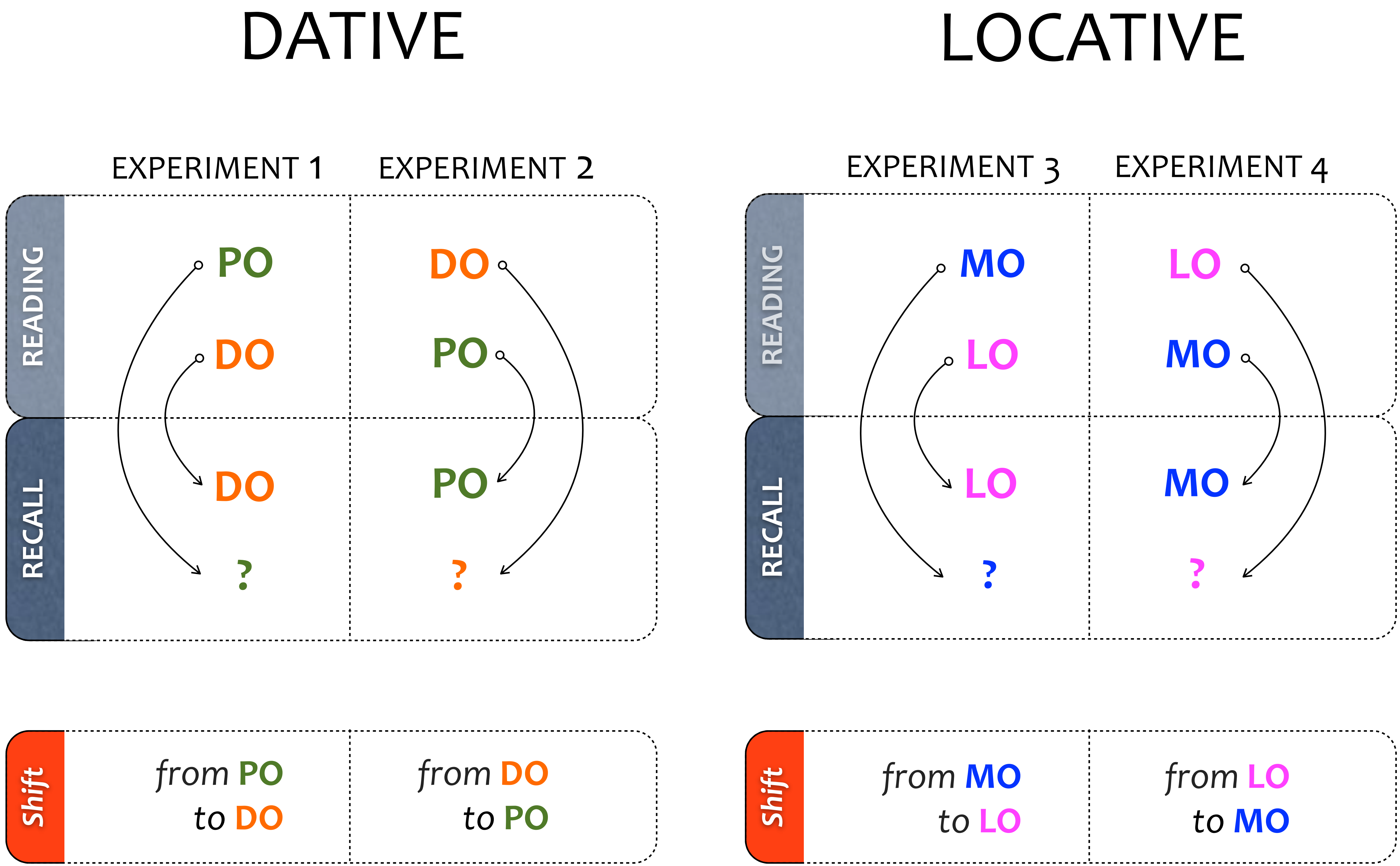
If the target is recalled in

- **PO** = recalled as was read
- **DO** = shifted to the prime construction

➡ Priming effect was measured by shifts from *the construction read* to *the construction primed.*

RSVP

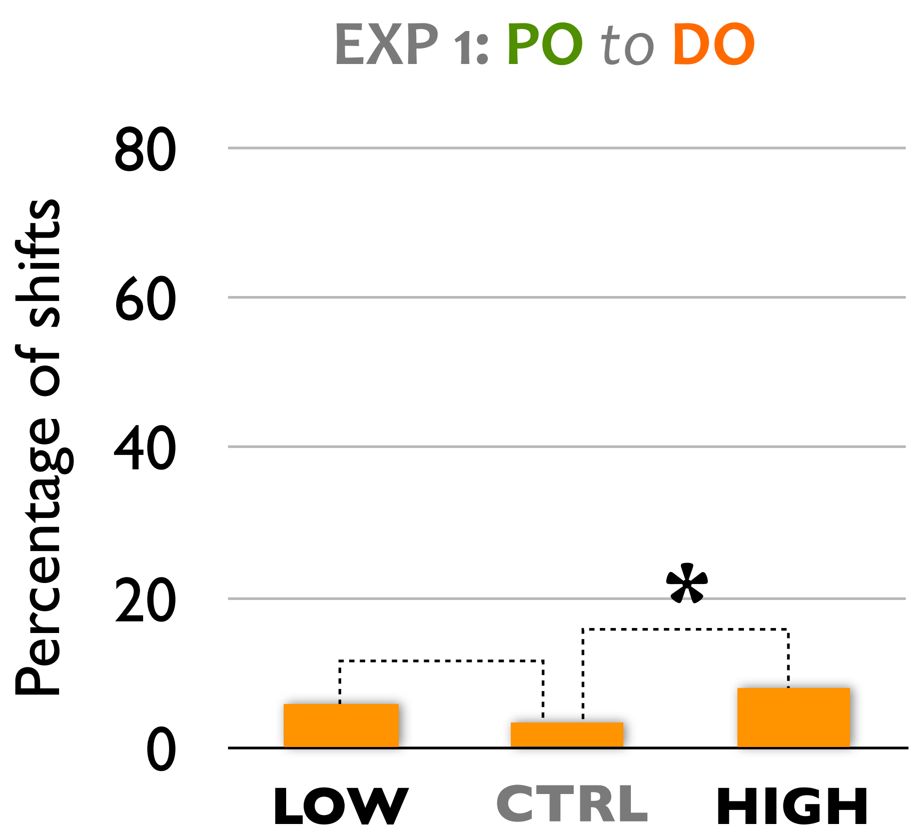

****
The
director
promised
a
large
part
to
the
actress
Reading a target sentence
######
Organic
foods
are
increasing
in
popularity
recently
Reading a prime sentence
######
Organic foods are _____
Cue to recall prime
The director promised _____
Cue to recall target

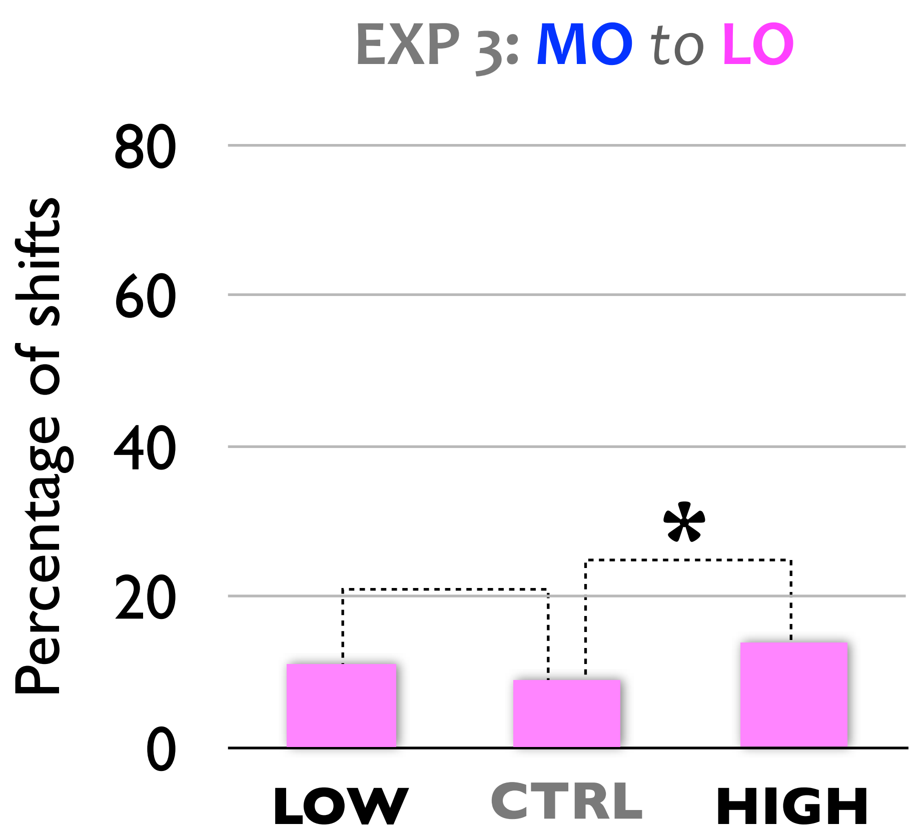

DATIVE
LOCATIVE
EXPERIMENT 1
EXPERIMENT 2
EXPERIMENT 3
EXPERIMENT 4
READING
RECALL
PO
DO
DO
?
DO
PO
PO
?
MO
LO
LO
?
LO
MO
MO
?
Shift
from PO
to DO
from DO
to PO
from MO
to LO
from LO
to MO

Frequency of prime sentence structures

Shift to	**DATIVE** **DO**:**PO**=16:84	**LOCATIVE** **LO**:**MO**=26:74
a **less frequent** construction	*from* **PO** *to* **DO** EXPERIMENT 1	*from* **MO** *to* **LO** EXPERIMENT 3
a **more frequent** construction	*from* **DO** *to* **PO** EXPERIMENT 2	*from* **LO** *to* **MO** EXPERIMENT 4

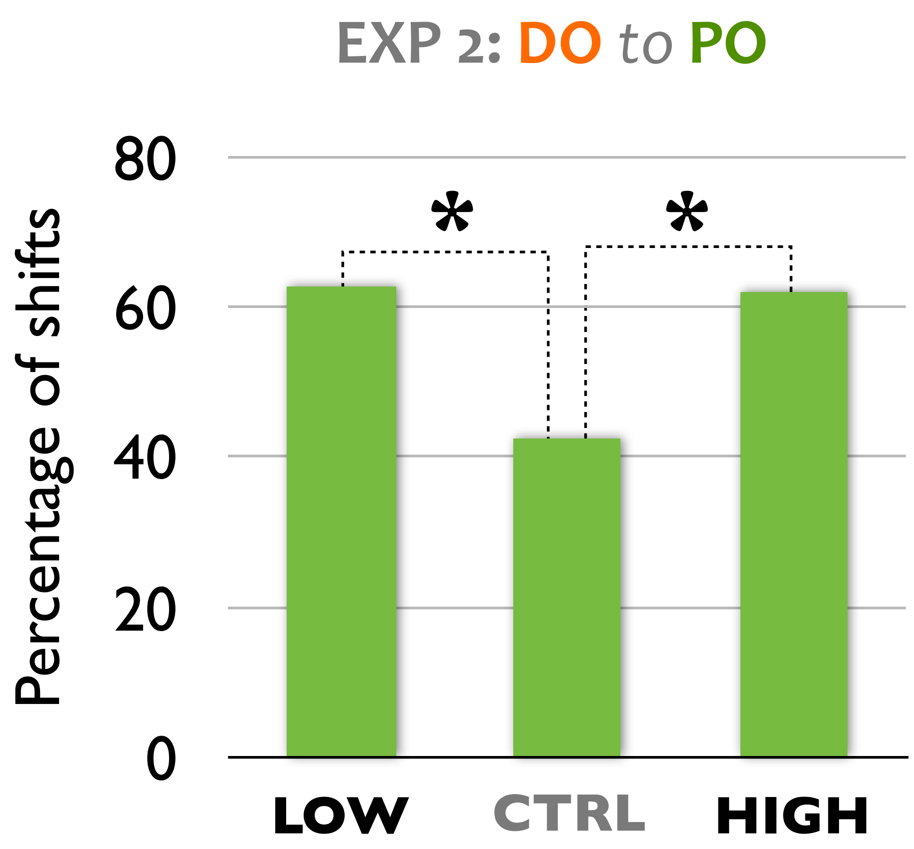
Shift to the **less frequent** construction

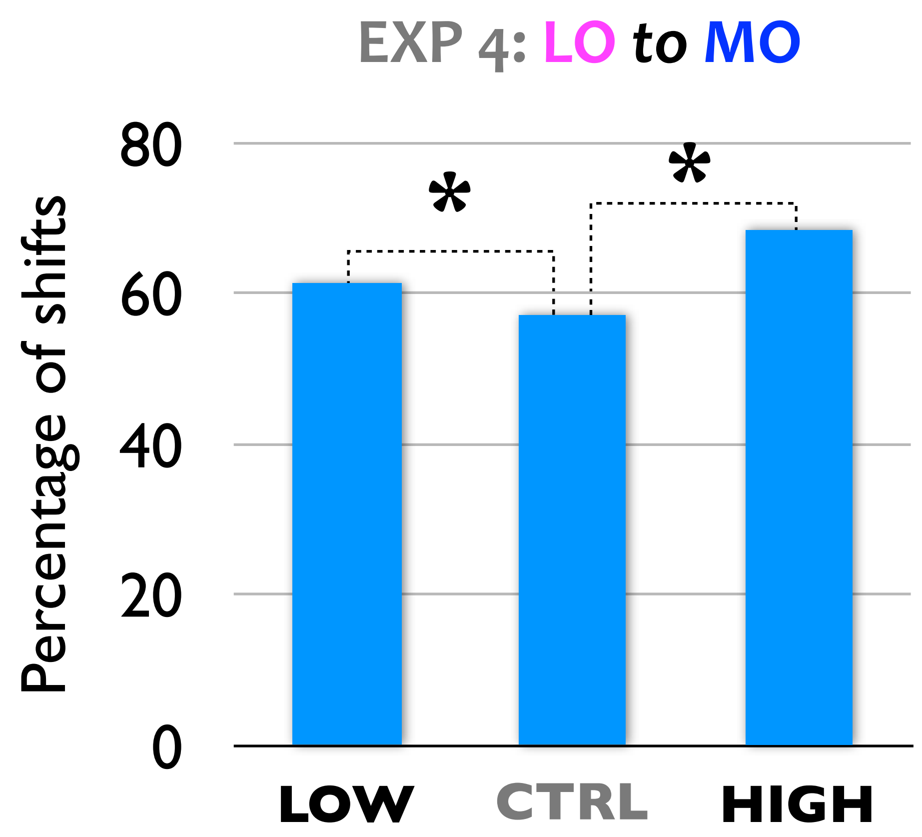




- Only prime verbs HIGHLY similar in meaning to target verbs lead to syntactic priming.

Shift to the **more frequent** construction





- Both HIGH and LOW semantic similarity primes lead to syntactic priming.

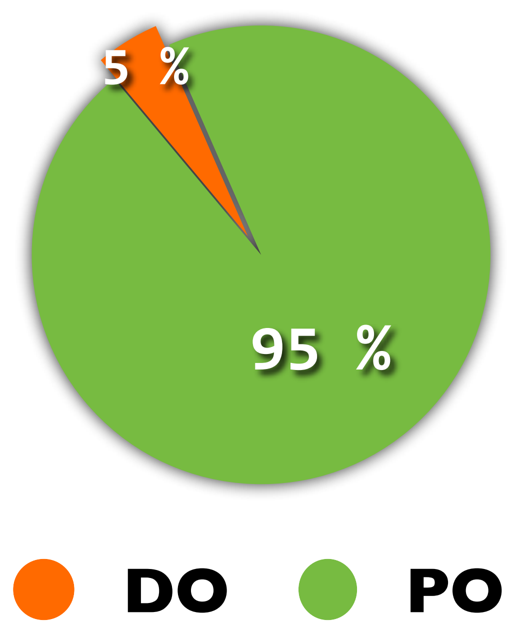
The interaction between syntactic priming and syntactic frame bias

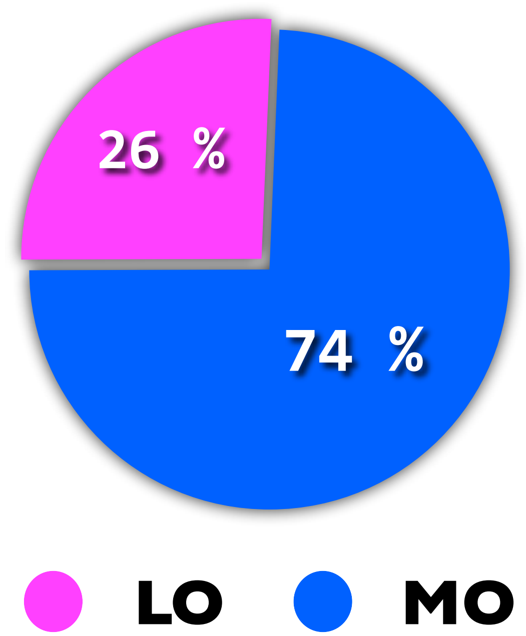
Experiment 1 *Shifts to* DO	Experiment 2 *Shifts to* GO	Prime Type	Experiment 3 *Shifts to* PO	Experiment 4 *Shifts to* FO
7,9 %	7,9 %	Control (Baseline)	42,3 %	54,3 %
11.0% (3.1%↑)	14.6% (6.7%↑)	Low-similarity	62.8% (20.5%↑)	63.2% (8.9%↑)
14.8% (6.9%↑)	18.9% (11.0%↑)	High-similarity	62.5% (20.2%↑)	71.3% (17.0%↑)

What is responsible for the results in the difference in shifts across expt's 1-2 and 3-4?

- One frame is strongly favored particularly when the 1st post-verbal NP is not a pronoun.

The 1st post-verbal NP is a non-pronoun





- Priming occurs regardless of meaning similarity when the target syntactic structure is overwhelmingly preferred.

The mechanism underlying our priming effect

- At least two views of priming
  - Residual activation (Branigan, Pickering)
  - Implicit learning (Chang, Dell, Bock)
- A cue-based memory retrieval model (Reitter et al)?

  “Any chunk that is present in a buffer may serve as a cue to other chunks held in memory if the model assumes an association S_{ji} between the two chunks j and i.”

PLAUSIBILITY IS NOT REDUCIBLE TO PREDICTABILITY

Hong Mo Kang, Jean-Pierre Koenig, Gail Mauner

(University at Buffalo)

University at Buffalo The State University of New York

Purpose of our studies

- Examine the relationship between predictability and plausibility
- Determine whether the effect of plausibility on processing can be separated from the effect of predictability

Background

- Effect of world knowledge on reading (plausibility):

  John used a knife to chop the large carrots for dinner.

  John used an axe to chop the large carrots for dinner.

  (Rayner et al., 2004)

- Effects of plausibility on processing tested through manipulation of *thematic fit*

Does predictability solely explain reading difficulty?

- Predictability as the sole source of processing difficulty
  - Plausibility=predictability (of role fillers) (Padó, Crocker, & Keller, 2009)
    - Plausibility of a role filler for a particular verb is defined as the joint probability of lexical, semantic and syntactic information
    - $\text{Plausibility}_{v,r,a} = P(v, s, gf, r, a)$
  - Effect of plausibility is mediated by effect of predictability of words (Levy,2008)
    - $\text{difficulty} \propto -\log P(w_i \mid w_{1\ldots i-1}, \text{CONTEXT})$
    - Any representation that affects reading difficulty does it through predictability of next word.

The right corpus size is not that “big”

- 130 (spoken) words per minute
- 8 hours of *continuous* speech per day
- 20 years median age for our participants
- Right corpus size: 455,520,000
- In corpus of this size (or larger), our target regions for sentences in our experiments have *P=0*

Effect of plausibility above and beyond predictability

- DeLong, Quante, and Kutas (2014)

  (1) It was difficult to understand the visiting professor. Like many foreigners he spoke with a/an [accent/lisp/apron] when conversing in English.

  - A posterior post-N400 ERP component associated with plausibility

condition	word	Mean predictability(SD) (0-1)	Plausibility (1-5)
High cloze/ High plausibility	accent	.88 (.13)	4.9 (.09)
Low cloze/ Somewhat plausible	**lisp**	**.03 (.08)**	**2.8 (.86)**
Low cloze/ low plausibility	**apron**	**<0.01(<0.01)**	**1.2 (.18)**

Methodological issues

1. Categorical predictors
   - Discrete rather than continuous predictors: Not enough data points to be able to tease apart the effect of predictability and plausibility
2. Effect of plausibility was not tested across all levels of predictability
   - High plausibility=high predictability, Low plausibility=low predictability (Abbott & Staub, 2015; Matsuki et al., 2011)
3. Limitation of the measure of predictability
   - "Low-predictability" items were often high-predictability items for some participants

Methodological focus of our study

- Expanded range of predictability and plausibility
  - The predictability of low predictability items is truly low
  - Maximize the range of plausibility within each predictability level
- Better measures of predictability and plausibility
  - Predictability and plausibility treated as continuous variables to maximize chances of teasing them apart

New method for measuring predictability

- 31 sentence frames with 3-argument verbs

(1) The girl sent her boyfriend ________.

- 10 continuations per sentence frame
- 103 undergraduate students of University at Buffalo
- Predictability calculated as proportion of weighted sum (Roland et al., 2012)
  - Weighted sum of continuations: 1st response: 10, 10th response: 1
  - predictability=weighted sum of continuation/sum of the weighted sum of all continuations

Is plausibility reducible to predictability?

- Objective
  - To investigate whether predictability explains plausibility
- Participants
  - 143 undergraduate UB students
- Task
  - Plausibility judgment task
- Two possibilities:
  - If plausibility is predictability, predictability will predict plausibility judgments
  - If plausibility is not predictability, predictability will not predict plausibility judgments

How to not artificially match levels of plausibility and predictability

- Grouped continuations into three predictability intervals for each sentence frame:
  - High predictability : pred.>$10*10^{-3}$
  - Medium predictability: $2*10^{-3}$<pred.<$5*10^{-3}$
  - Low predictability: pred.<$1*10^{-3}$ and answered only once (**never answered earlier than 5th**)
- Selected most/least plausible continuations within each predictability level for each sentence frame

How materials for plausibility judgment task were selected

- We randomly selected a subset of cloze responses (9 for high pred.; 9 for medium pred.; 15 for low pred.)
  - Selection needed because too many responses per sentence frame to have all continuations rated (or sorted) (820 on average per sentence frame)
- 98 undergraduate UB student chose the most/least plausible event participant among the randomly selected responses
  - 6 sentences (2 plausibility levels x 3 predictability levels) for each sentence frame

Plausibility judgment study

- Participants asked to judge the plausibility of 24 sets of sentences on a 7 point Likert scale.
- Analysis
  - *Ordinal* package of R (version 2015.6-28)
  - Predictors: log-transformed predictability, lemma frequency (COCA)
  - Dependent variable: plausibility judgment data
    (7: highly plausible, 1: not at all plausible)
  - Model fit index: Mcfadden's pseudo R-squared (McFadden,1973)
    - $R^2_{McFadden} = 1 - \log(L_c)/\log(L_{null})$
    - 0.2~0.4 as "excellent fit" (McFadden,1977)

Example stimulus set

Condition		Sentence
Pred.	Plau.	
High	High	The girl sent her boyfriend a present.
High	Low	The girl sent her boyfriend flowers.
Med.	High	The girl sent her boyfriend brownies.
Med.	Low	The girl sent her boyfriend drugs.
Low	High	The girl sent her boyfriend a selfie.
Low	Low	The girl sent her boyfriend breakfast.

Overall (McFadden's ps R-squared=0.03)

	Coef.	SE	*z*	*p*
Predictability	0.89	0.06	14.50	<0.001
Lemma frequency	-0.76	0.06	-13.63	<0.001

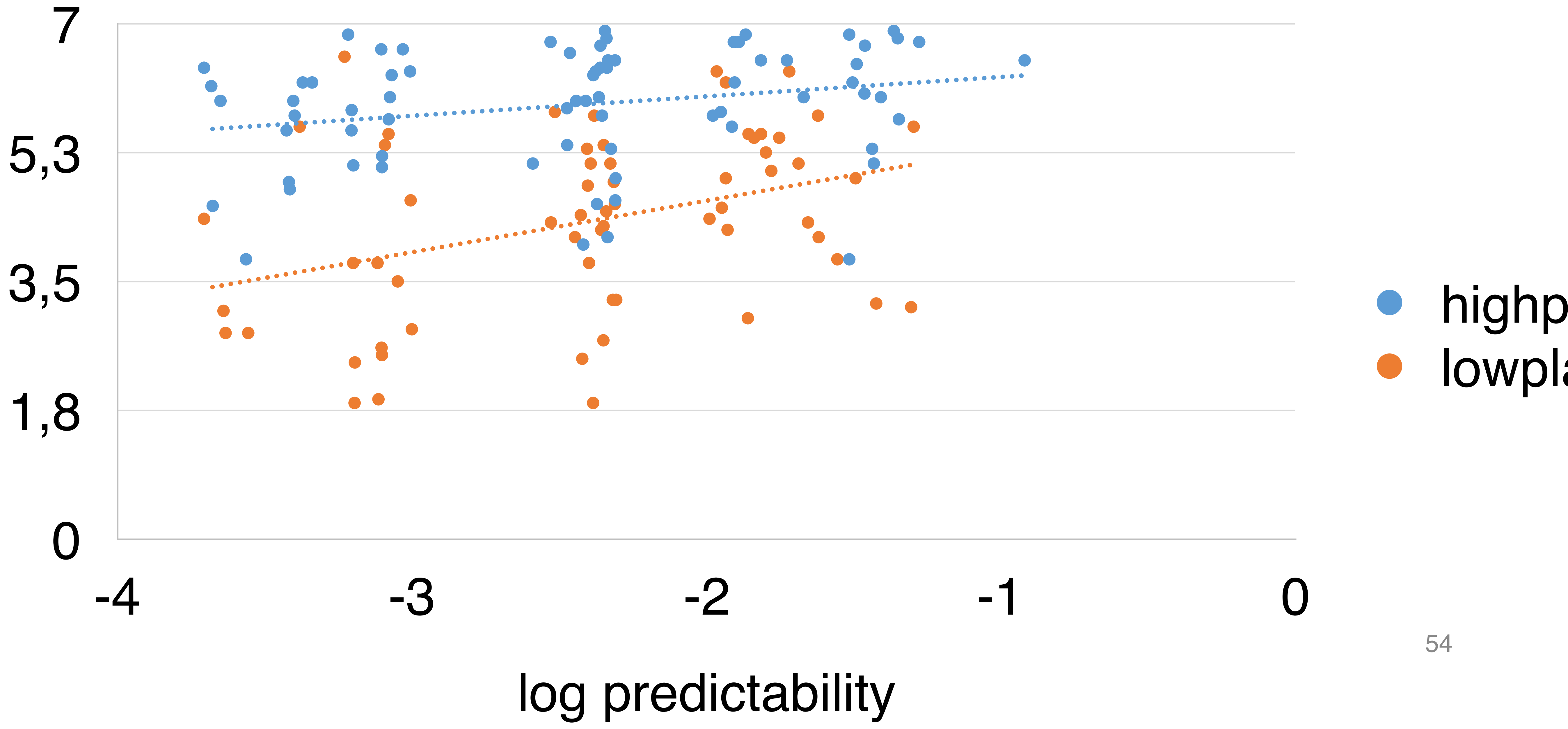
High plausibility (McFadden's ps R-squared=0.03)

	Coef.	SE	*z*	*p*
Predictability	**0.12**	**0.10**	**1.30**	**0.19**
Lemma frequency	0.14	0.09	1.60	0.11

Low plausibility(McFadden's ps R-squared=0.02)

	Coef	SE	*z*	*p*
Predictability	**1.06**	**0.09**	**11.14**	**<0.001**
Lemma frequency	-0.54	0.10	-5.19	<0.001

Predictability explains plausibility only in low-plausibility condition



Predictability does not explain much of plausibility judgement responses

- Low explanatory power of predictability
  - Very low pseudo-R-squared
- Predictability predicts plausibility when plausibility is low
- Predictability does *not* predict plausibility when plausibility is high

Can the effect of plausibility be separated from the effect of predictability in online processing?

- Task
  - Self-paced region-by-region reading
- Predictions
  - Effects of predictability and plausibility can be dissociated
  - The effect of plausibility will occur later than the effect of predictability
- Participants
  - 143 undergraduate students of University at Buffalo

Example stimulus set (2 spillover regions added to normed sentences)

Condition		Sentence
Pred.	**Plau.**	
High	High	The girl/ sent/ her boyfriend/ a present/ together with a card/ yesterday.
High	Low	The girl/ sent/ her boyfriend/ flowers/ together with a card/ yesterday.
Med.	High	The girl/ sent/ her boyfriend/ brownies/ together with a card/ yesterday.
Med.	Low	The girl/ sent/ her boyfriend/ drugs/ together with a card/ yesterday.
Low	High	The girl/ sent/ her boyfriend/ a selfie/ together with a card/ yesterday.
Low	Low	The girl/ sent/ her boyfriend/ breakfast/ together with a

Analysis

- R's *lme4* package (*t*-test with *LmerTest* package)
- Predictors
  - Target region
    - Predictability (continuous), mean plausibility (continuous), COCA frequency (continuous)
  - Spillover region
    - Predictability (continuous), mean plausibility (continuous), length of target region
- Dependent variable
  - Residual reading time (Trueswell, Tanenhaus, & Garnsey, 1994)

Effect of predictability on target region

	Coef.	SE	*t*	*p*
(intercept)	20.61	9.88	2.09	0.04
Predictability	**-16.25**	**7.92**	**-2.05**	**0.04**
Plausibility	**-3.86**	**4.33**	**-0.89**	**0.37**
Frequency	5.12	5.38	0.95	0.34

* No significant predictability*plausibility interaction slope

Effect of plausibility delayed until spillover region

	Coef.	SE	*t*	*p*
(intercept)	-31.56	15.83	-1.99	0.06
Predictability	-8.28	8.24	-1.01	0,31
Plausibility	-13.11	5.04	-2.60	0,01
Target length	6.56	1.78	3.69	<.001

* No significant predictability*plausibility interaction slope

Effect of plausibility on reading

- Is not reducible to the effect of predictability:
  - Effect of predictability on reading time occurs on the region that is more or less predictable
  - Effect of plausibility on reading time occurs on the region that follows the region that is more or less plausible

Why is the effect of predictability on reading earlier than the effect of plausibility?

- Predictability is “pre-processed”
  - Higher activation of target word when it is predicted
- Plausibility is computed once the word is read
  - “Does the target word fit my event knowledge?”

Why can the effect of plausibility can be dissociated from the effect of predictability?

- Reichle, Warren, and McConnell (2009)
  - L_1 stage : Familiarity check stage
    - Affected by predictability
  - I stage: Post-lexical integration stage later than L_1 stage
    - Affected by plausibility

Conclusion

- Plausibility can be distinguished from predictability
- The effect of plausibility on processing cannot be entirely explained by predictability
  - Provided the range of the materials' predictability is broad enough
  - Provided the measures of predictability and plausibility are fine-grained enough (and treated as continuous variables)

So what is plausibility if not predictability?

- Plausibility measures how well does the participant fit in our event knowledge
- Hypothesis: Plausibility of a participant role filler is a function of how easy it is to relate it to “typical” events
  - We can approximate this “ease” by how similar a participant is to “typical" participants from events we have seen or heard about
  - There are more than one “typical” participant
  - Cloze task answers can be grouped into semantic domains and assume each domain has one “typical” participant

How we tested our hypothesis

- We automatically clustered all our cloze task answers using k-means clustering on the basis of pair-wise semantic similarity
- We defined the center of the cluster (the medoid) as the item with less distance to other members of clusters
- If our hypothesis is correct, distance from medoid should be smaller for high plausibility items than for low plausibility items

Condition		Sentence	Medoid	Distance to medoid	Mean plausibility
Pred.	Plau.				
H	H	The girl sent her boyfriend a present.	meal	0.86	6.9
H	L	The girl sent her boyfriend flowers.	chocolate	0.66	5.6
M	H	The girl sent her boyfriend brownies.	chocolate	0.60	6.6
M	L	The girl sent her boyfriend drugs.	chocolate	0.79	5.1
L	H	The girl sent her boyfriend a selfie.	N/A		6.85
L	L	The girl sent her boyfriend breakfast.	meal	0.33	5.35

Distance from center of semantic cluster is a predictor of plausibility judgment

	Estimate	Std. Error	Z value	Pr(>|z|)
Distance from medoid	-0,4492	0,1984	-2,264	0.0235*

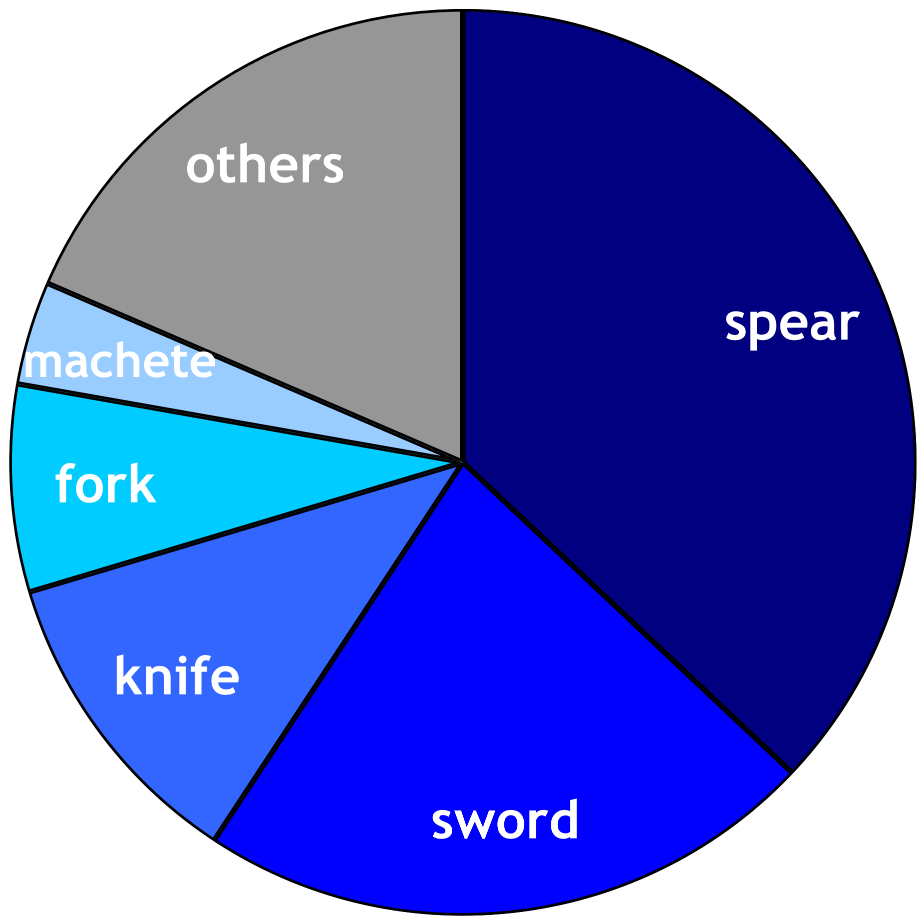
- As distance from medoid increases, plausibility rating decreases

Predictability and semantic similarity

- Plausibility does not reduce to predictability
- It has an independent effect on processing and that effect might be a function of semantic similarity to the centers of semantic clusters
- What about semantic similarity on its own?

Semantic predictability affects processing

*The aboriginal man **jabbed** the angry lion with a/n ____.*

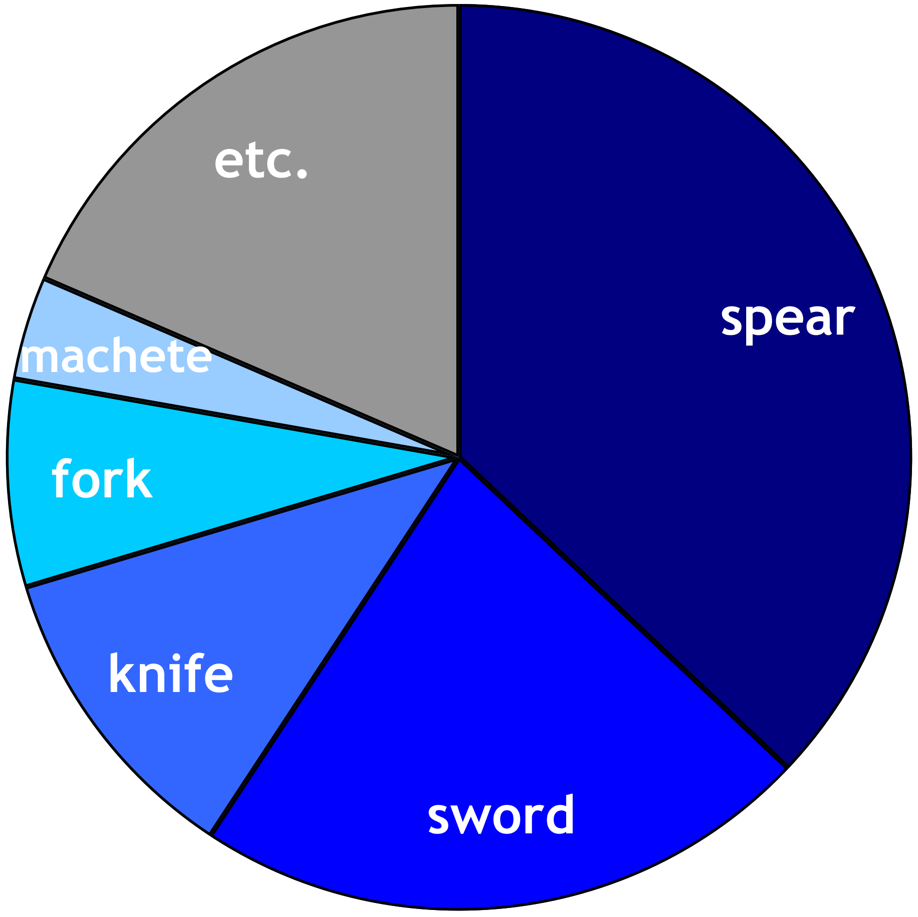


Spear is more likely than ***fork***, so it will be easier to process (e.g., Altmann & Kamide, 1999; 2003; McRae et al., 1997; Speer & Clifton, 1999).

Imagine a second verb has slightly different expectations…

attacked

The aboriginal man ~~jabbed~~ the angry lion with a/n ____.

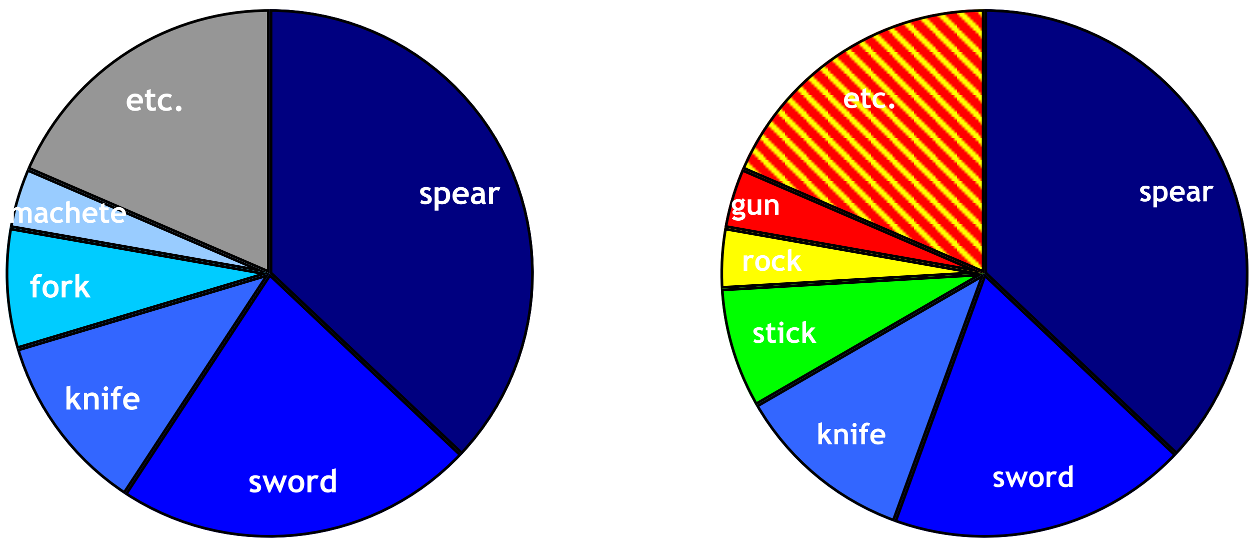


Note – the real numbers for these verbs aren't quite this perfect!

Imagine a second verb has slightly different properties…

attacked

The aboriginal man ~~jabbed~~ the angry lion with a/n _____.



Question 1:

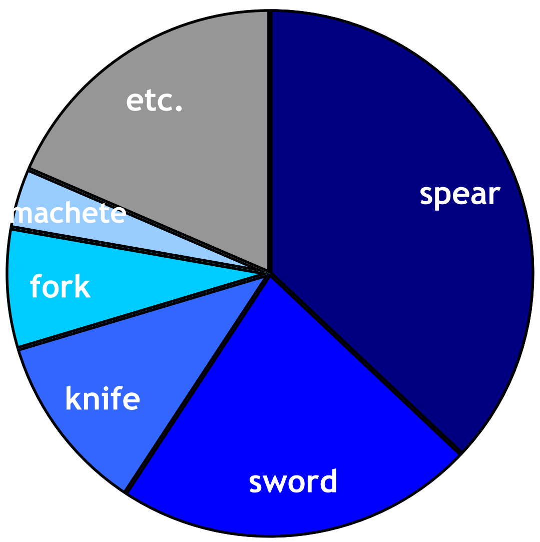
If ***spear*** is equally likely for both verbs, is it easier to process after ***jab***, due to the greater degree of shared similarities between the possible instruments?

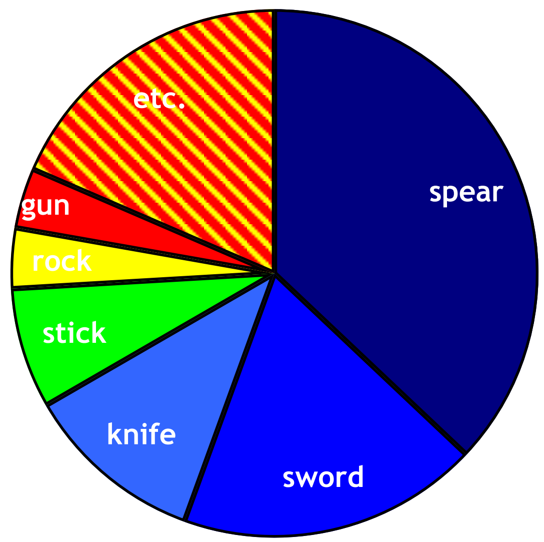
The *Shared Similarity* hypothesis predicts YES

Imagine a second verb has slightly different properties…

attacked

The aboriginal man ~~jabbed~~ the angry lion with a/n ____.





Question 2:

Is ***machete*** easier to process after ***jab*** than ***rock*** is after ***attack***, even though they are equally likely, due to the greater degree of shared similarities between ***machete*** and the other possible instruments of ***jab***?

The *Shared Similarity* hypothesis predicts YES

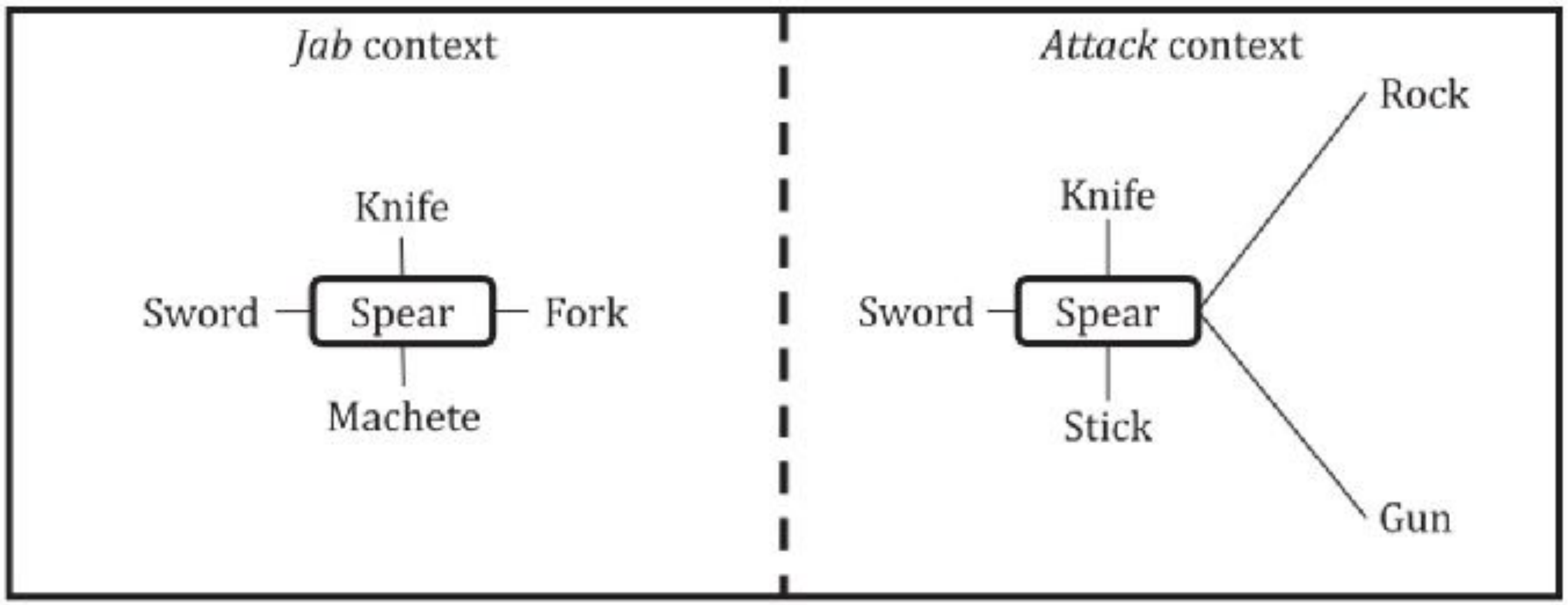


Fig. 2. Semantic similarity between spear and the other possible instruments in contexts (1) and (2).

Measuring predictability

- Sentence completion
  The aboriginal man jabbed the angry lion with _______________.
- List 5 possible things that fit in the blank
  The aboriginal man jabbed the angry lion with ____ near its prey.

Measuring shared semantic similarity

- Compare target instrument with sets generated via:
  - Sentence completion experiment
- Measure similarity using
  - Latent Semantic Analysis (LSA) (Deerwester et al., 1990)
  - Wordnet similarity (*vector pairs* similarity measure, Patwardhan & Pederson, 2006)

More similar → faster reading time

Shared similarity is not reducible to predictability

- Self-paced reading times with “makes sense” judgement

1. The gladiator |jabbed| the African tiger | with | a sword/spike | in | the Colosseum

Table 1
Summary of fixed effect predictors from the linear mixed-effects regression model for predicting reading times of the target word.

		Estimated coefficient	Standard error	*t* Value
(Intercept)	714.70	(714.70)	23.57	30.32
Predictability	−60.23	(−39.49)	10.74	−5.61
Similarity	−179.33	(−18.77)	79.75	−2.25
Length	21.52	(48.64)	3.28	6.57
Frequency	−14.55	(−14.55)	9.42	−1.54
Predictability × Frequency	35.39	(23.20)	10.41	3.40

Note: All predictors are centered, frequency predictor is residualized for length and predictability. Parenthetical values following the estimated coefficients are standardized coefficients from an alternate version of the model with standardized predictors.

Conclusions

- Shared semantic similarity between possible fillers plays a role in language comprehension
  - Knowledge of a verb must include very fine grained semantic information
  - … including information about which types of instruments can be used to perform the action
  - This information is (rapidly) used during language comprehension
- Structural aspect of verb meaning
  - Largely limited to a core set of semantic features
  - But include arbitrary properties
- If structural and non-structural aspects of verb meaning are both:
  - arbitrary
  - affect comprehension

  what is the difference?

Semantic similarity interacts with how constraining context is

- Some contexts constrain range of upcoming material more than others

1. The gladiator jabbed the African tiger with

2. The aborigine attacked the angry lion with

- Hypothesis: Effect of semantic similarity stronger in weaker contexts

- Entropy of instrument responses as measure of contextual strength

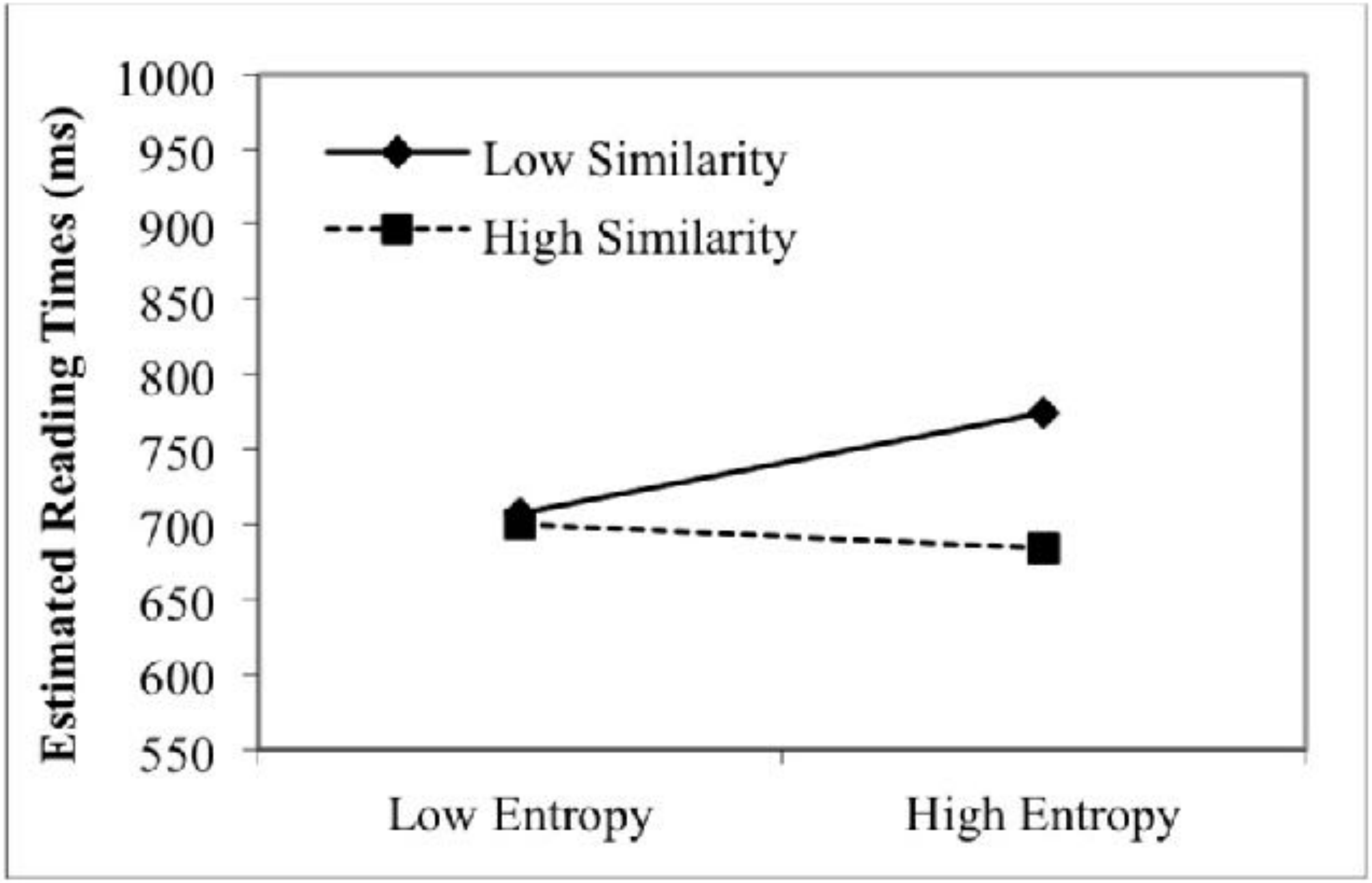


Figure 2: Interaction of Entropy and Similarity using standardized coefficients.

Two different cognitive tasks

- Our tentative hypothesis:
  - Processing and grammar induction are different tasks
  - Different demands on those tasks lead to sensitivity to distinct properties

Processing task

- *Comprehension*: Read word, integrate it with what was previously read and flesh out described event
  - Predicting next word and its various properties helps (= predictability)
  - How the described event fit our event knowledge helps (= plausibility)
  - How similar role fillers are helps (= filler similarity)
- *Production*: Choose lexical items, choose syntactic frames
  - How similar verb is to preceding verb or path-breaking verb affects frame activation and choice (= verb similarity)

Event grammar task

- Event grammar task: Sift 12,000 to 16,000 verb senses into 50 or so structural distinctions
    - Only sensitive to *obligatory* participants
    - Grouping into semantic natural classes helps (but grouping does not have to be semantic): That's why color does not recur, presumably
    - Token independence helps: Do not have to worry about checking properties of referents of NPs
- Some steps are missing in this argument