

THE INTERACTION OF STATISTICS AND SELECTIONAL CONSTRAINTS ON SITUATED WORD LEARNING

JUDITH KÖHNE (JUDITH@COLI.UNI-SAARLAND.DE) & MATTHEW W. CROCKER



References
 Altmann, G.T.M., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, 73, 247-264.
 Yu, C. & Smith, L.B. (2007). Rapid word learning under uncertainty via cross-situational statistics. *Psychological Science*, 18, 414-420.

ADULTS RELY ON STATISTICS

To learn word meanings, they keep track of co-occurrences of words & referents (cross-situational word learning, CSWL, Yu & Smith, 2007).

ADULTS RELY ON THEIR KNOWLEDGE ABOUT LANGUAGE STRUCTURES AND THE WORLD

They use this knowledge to make predictions and inferences when processing spoken sentences (Altmann & Kamide, 1999)

How do statistics & knowledge interact in word learning?

→ Experiment 1

Which cue dominates when they are in conflict?

→ Experiment 2

EXPERIMENTAL PARADIGM & BASIC PROCEDURE

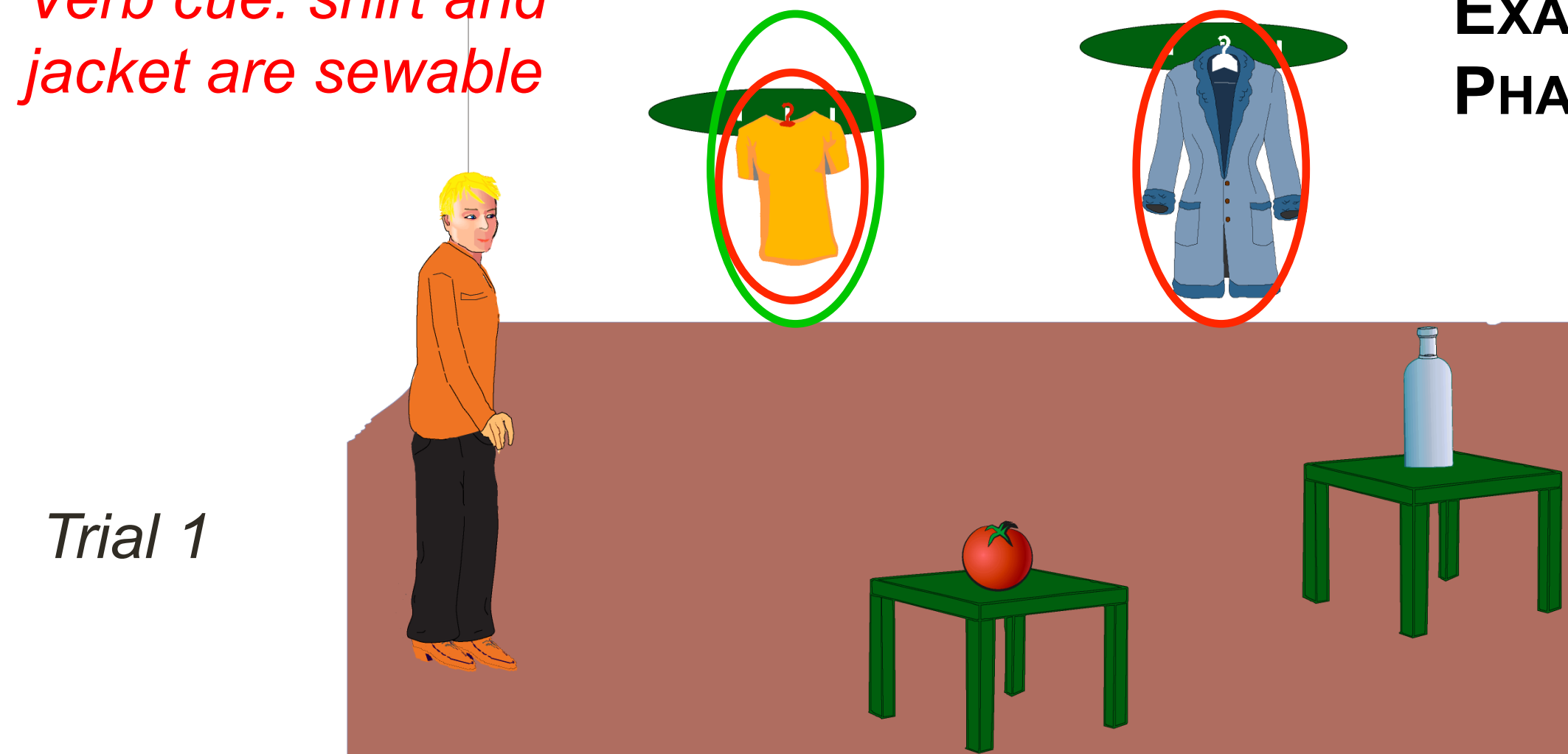
German adults learn a semi-natural language (based on Indonesian) in 3 basic steps:

1. **Verbs:** Familiarization with verbs (restrictives: e.g. eat & non-restrictives: e.g. take).

2. **Noun learning:** Visual scenes & auditory SVO-sentences presented. Task: understand sentences & learn nouns.

3. **Vocabulary test** (Forced choice) & confidence rating ('How sure are you about your choice?', 1(not sure)-9(sure))

Verb cue: shirt and jacket are sewable

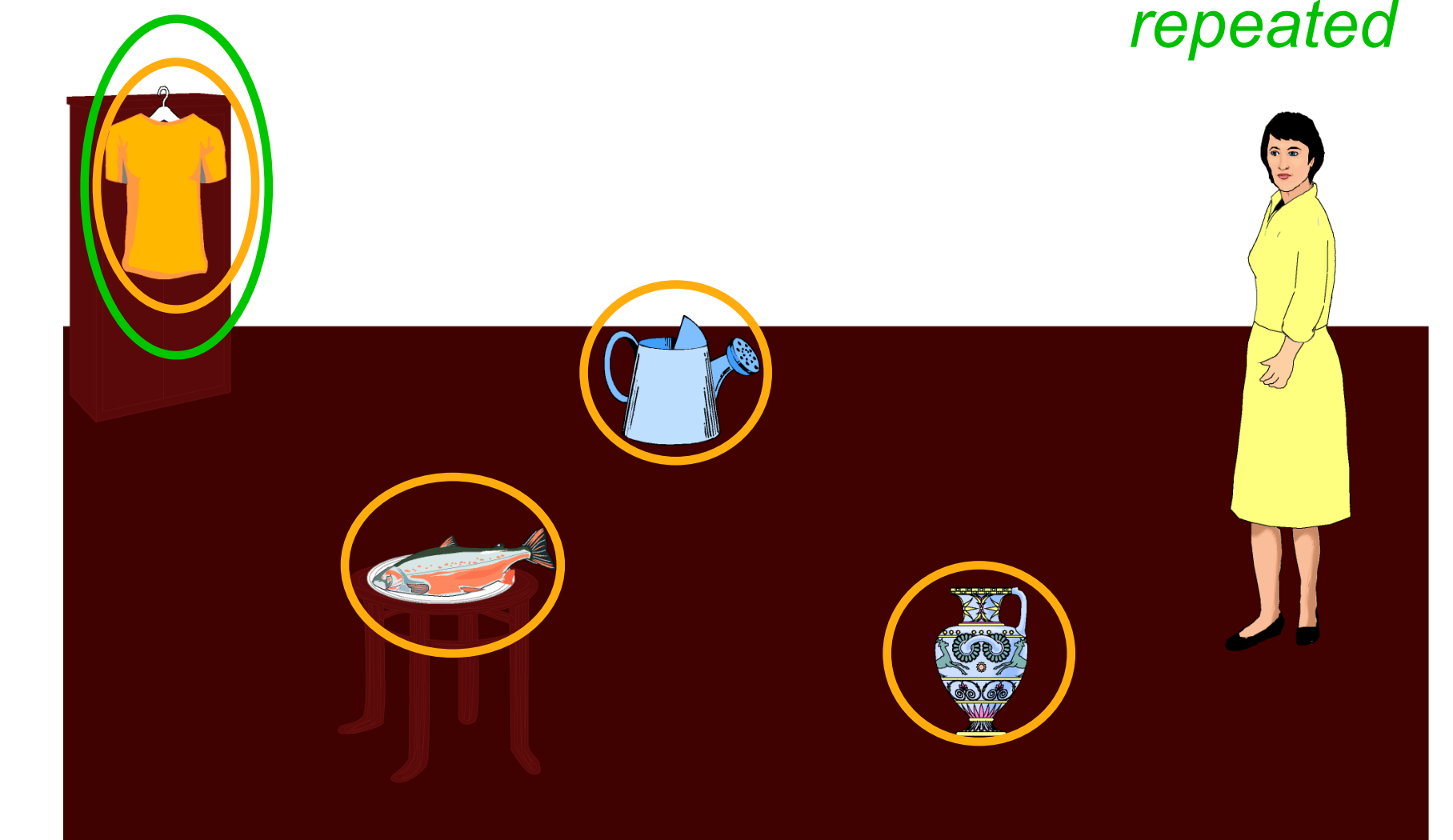


Trial 1

[Si laki]_{NP1} [mankemema]_{verb} [si badut]_{NP2}
 DET man sew-MEMA DET shirt
 'The man will sew the shirt.'

EXAMPLE NOUN LEARNING PHASE

CSWL: shirt repeated, 'badut' repeated



Trial 2

[Si gadis]_{NP1} [tambamema]_{verb} [si badut]_{NP2}
 DET woman take-MEMA DET shirt
 'The woman will take the shirt.'

Remember: All verbs are known already in this phase!

EXPERIMENT 1

Manipulation of the degree of verb restriction to study the interaction of CSWL and verb-derived inference learning

CONDITIONS

The nouns (Phase 2) were in one of three conditions:

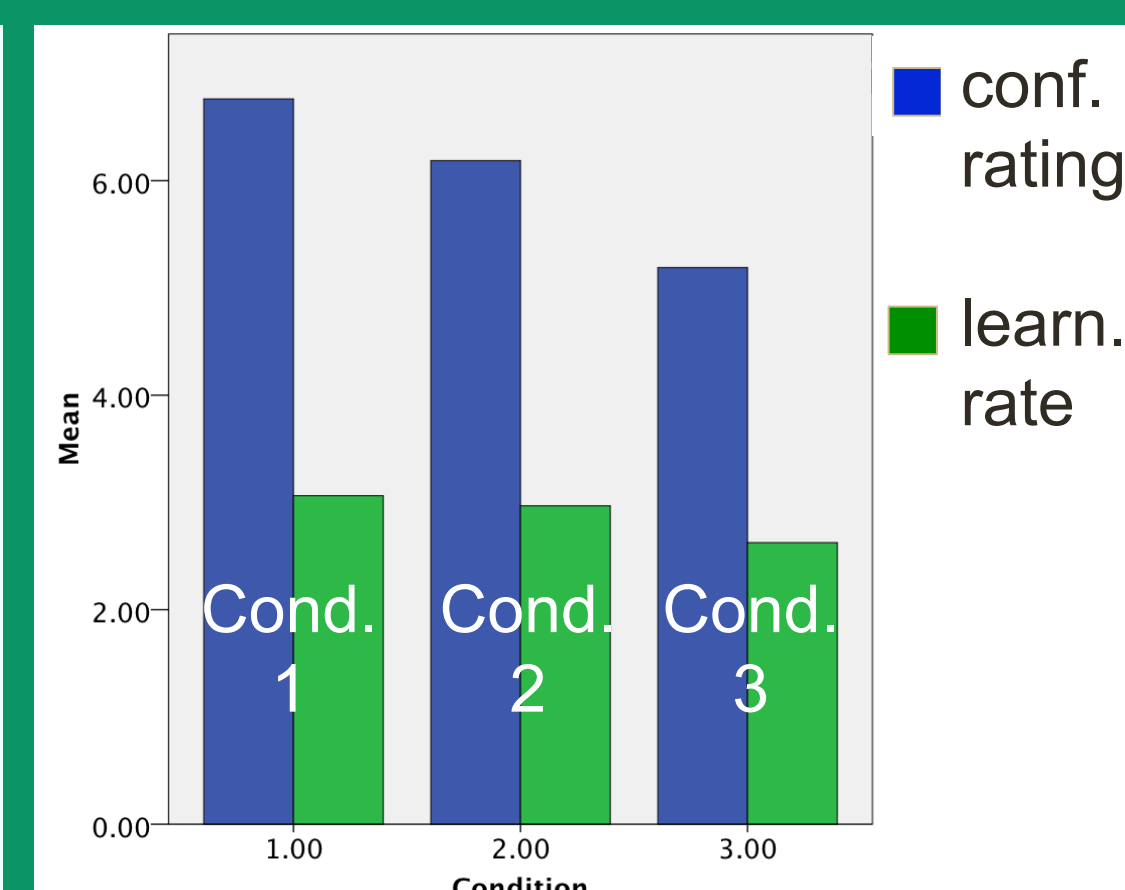
- Noun occurs with **restrictive verb** and there is only **1 object** in scene matching the verb category (+3 distractor objects): **Si laki bermamema si sonis** ('The man eats the tomato'); scene: **tomato, shirt, bottle, jacket, man**
- Noun occurs with **restrictive verb** but there are **2 objects** in scene matching the verb category (+2 distractor objects): **Si laki mankemema si badut** ('The man sews the shirt'); scene: **shirt, jacket, tomato, bottle, man**
- Noun occurs with **non-restr. verb** > **4 potential referent objects**: **Si gadis tambamema si badut** ('The woman takes the shirt'); scene: **shirt, fish, vase, can, woman**

RESULTS

✓ Noun learning was significantly above chance for all conditions.

✓ Nouns were learned best (77%) and decisions were rated highest (7.0) in Condition 1. > Direct verb cue (with the visual context and word knowledge) helped learners to learn noun meanings.

✓ Learning & rating were higher for Condition 2 (74%; 6.4) than 3 (66%; 5.4). > Verb cues worked together with CSWL.



- Confidence ratings: all differences significant
- Learning: difference between Cond. 1 & 3 sig

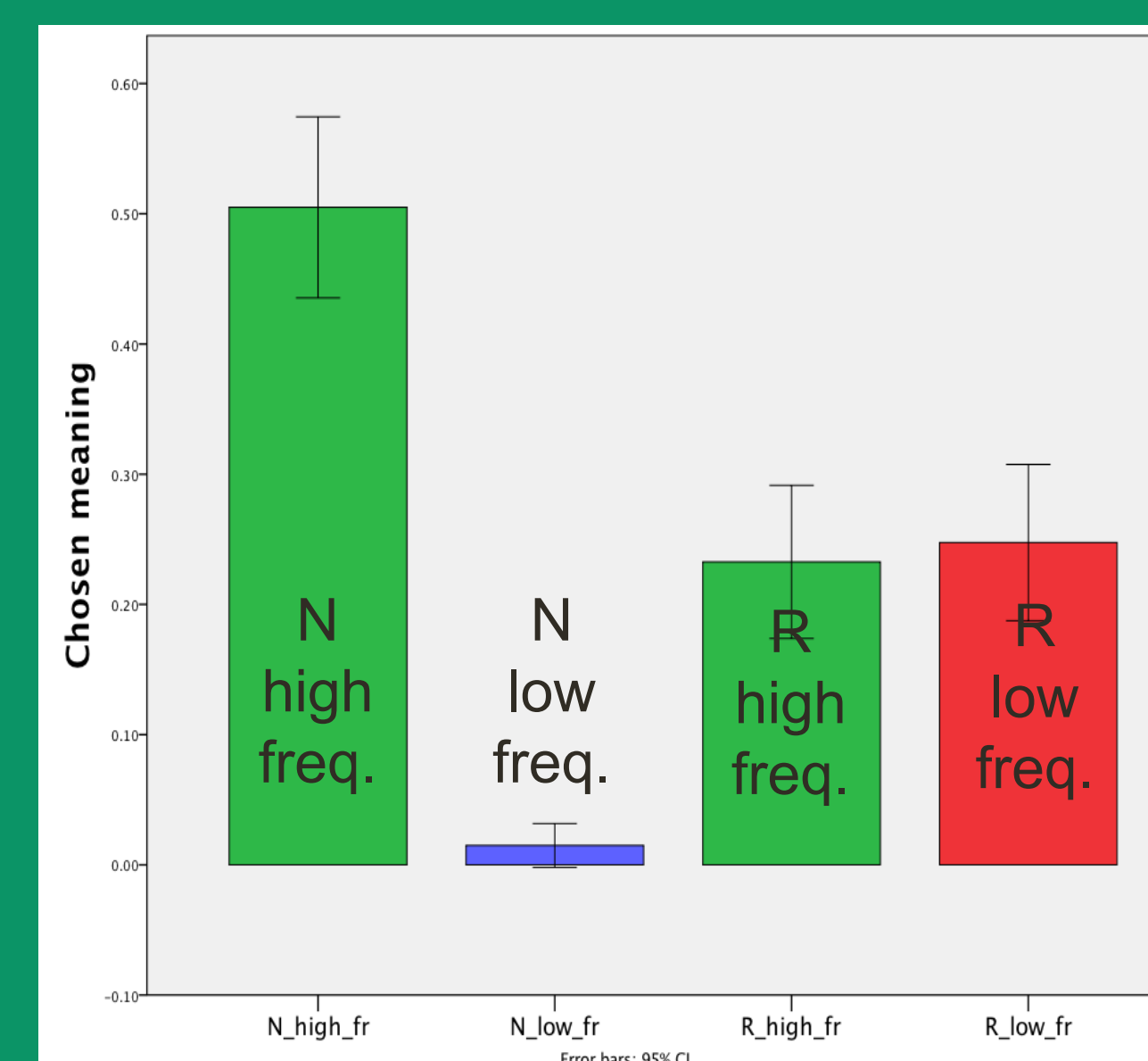
EXPERIMENT 2

Each noun has 2 potential meanings (= 2 potential referents) - 1 favored by frequencies, 1 favored by verb information

CRUCIAL DESIGN CHARACTERISTICS

- Each noun has two potential conflicting meanings
 - Low frequency meaning (co-occurrence noun & object: 50%)
 - High frequency meaning (co-occurrence noun & object: 83%)
- AND each noun is in one of 2 Conditions
 - R**(estrictive): Noun occurs with restrictive verb which favors the low-frequency meaning!
 - N**(on-restrictive): Noun occurs with non-restrictive verbs, i.e. there is no cue except statistics applicable.

>>> Nouns in **Condition N** just have a **frequency-favored** and a **frequency-disfavored** meaning.
 Nouns in **Condition R** have a **frequency-favored** and a **verb-favored** meaning (= conflict)!



RESULTS

✓ Noun learning was significantly above chance for both conditions, no significant difference (N: 87.5%, R: 80.8%)

✓ **N**: High-frequency meaning chosen 97% & **R**: High-frequency meaning chosen: 48.5%

>>> Interaction of chosen meaning and condition

✓ No difference in confidence ratings for **N** vs. **R** BUT sig. higher rating for high-frequency choices than low-frequency choices in **R** (7.5 / 6.5)

>>> In **R**, learners' choices are clearly biased by verbs! However, when learners chose the verb-favored meaning, their confidence was lower > frequencies still influenced them!

OVERALL CONCLUSIONS

- CSWL & verb information (+visual context +world knowledge) interact in foreign language word learning.
- When in conflict, both cues are considered.

- Statistical information cannot be fully ignored but overridden.
- Prior knowledge can be ignored, however, probably even stronger influence in real life.