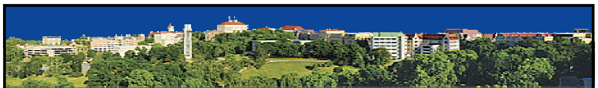


Lexical & Semantic Similarity in Word Learning

Holly L. Storkel, Ph. D.
Speech-Language-Hearing: Sciences & Disorders

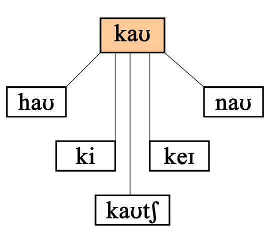
Model of the Lexicon: Assume 2 Representations

/kaʊ/

- Lexical Representation
- Semantic Representation

Similar Representations Grouped Together: Lexical

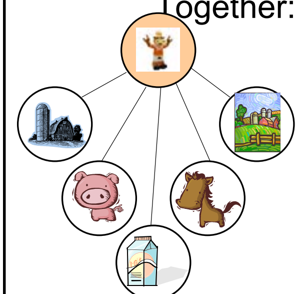
/kaʊ/



Quantifying Lexical Similarity: Neighborhood Density

- Number of phonologically similar words
  - Differing by 1 sound
  - Sparse /dcg/ (6) vs. dense /kæt/ (27)
- Influences recognition, production, & memory

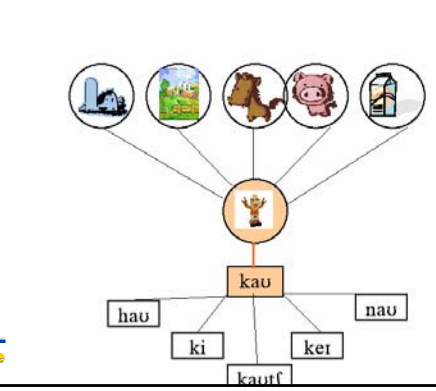
Similar Representations Grouped Together: Semantic



Quantifying Semantic Similarity: Multiple Methods

- Number of meaningfully related words
  - Small (cat - 3) vs. large (bird - 23)
- Methods vary in how similarity determined
  - Co-occurrence
  - Feature analysis
  - Association data
- Influences memory
  - Emerging studies of recognition

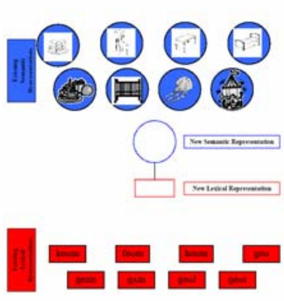
Similarity Matters for Known Word Retrieval



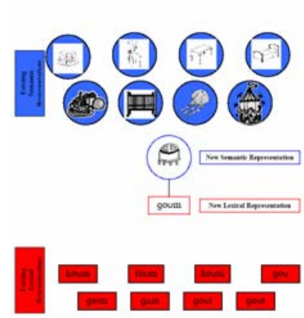
What about word learning?

/goʊm/

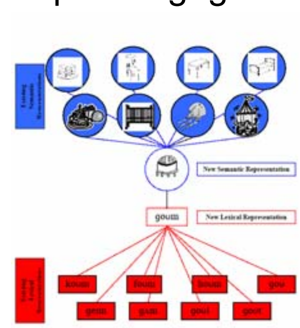
Step 1: Triggering



Step 2: Configuration



Step 3: Engagement



KU THE UNIVERSITY OF KANSAS

Question

Does similarity to existing representations influence word learning?

Overall Purpose

- How to determine similarity for novel words?
  - Study 1: Neighborhood density
  - Study 2: Set size
- Does similarity influence word learning?
  - Study 3: Neighborhood density
  - Study 4: Set size

Study 1: Neighborhood Density (Storkel, Hoover, & Kiewig)

- Relatively well established density algorithm
  - Count the number of words in a corpus that differ by one phoneme
  - Works for real words and novel words
- Question: What corpus?

Adult Corpus

- Hoosier Mental Lexicon (HML)
- Dictionary of ~20,000 words
- Computer readable phonemic transcription
- Used in most past research of children & adults
  - Is this appropriate?

Study 1: Questions

- How do child and adult neighborhood density values compare?
  - Methodological issues: Do we need child values?
  - Theoretical issues: How does the lexicon grow from child to adult?

Growth of the Lexicon

- Hypothesis 1:
  - Words added to lexicon are equally distributed across neighborhoods
    - Number of neighbors increases equally for all words

	Child Density	Adult Density
cat	0	3
bird	20	23

Growth of the Lexicon

- Hypothesis 2:
  - Words added to lexicon are asymmetrically distributed across neighborhoods
    - Number of neighbors increases differently for different words

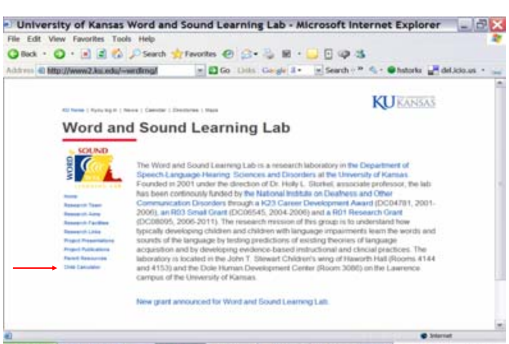
	Child Density	Adult Density
cat	0	3
bird	13	23

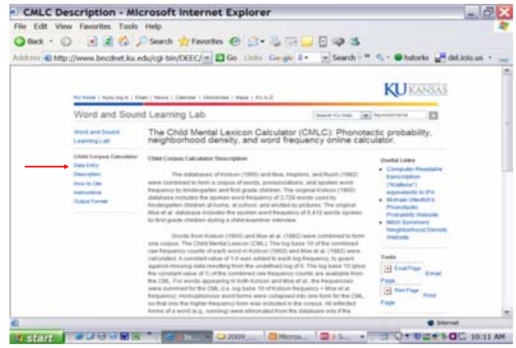
Methods

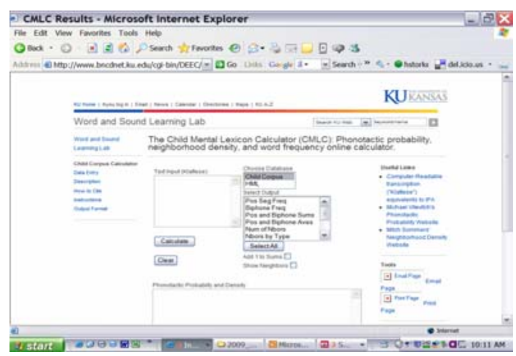
- Set of real words
- Compute two neighborhood densities
  - One based on adult corpus (HML)
  - One based on a child corpus (next slide)
- For both computations (adult & child), use 1 phoneme change algorithm

Child Corpus

- Based on spoken data from Kindergarten and 1st grade children
  - Moe, Hopkins, & Rush (1982)
  - Kolson (1960)
- Calculator created by Storkel, Hoover, & Kiewig
  - Provided phonemic transcriptions using HML conventions
  - Eliminated homonyms and morphologically related words
    - ~5,000 words
  - Calculator interface: http://www2.ku.edu/~wrdlrng/



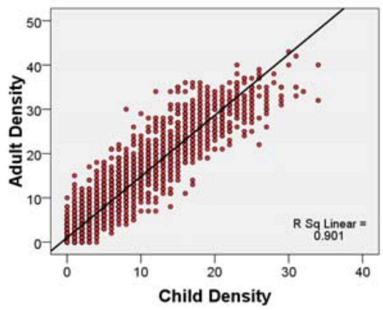




Question

How do child and adult densities compare?

Results: Correlation



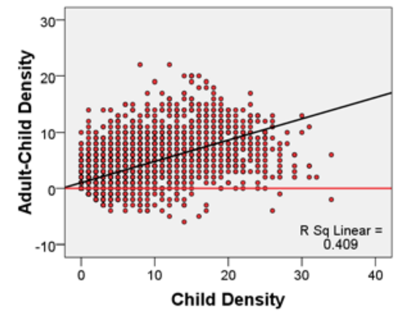
Results: T-test

- Adult density values were significantly larger than child density values
- This is the expected pattern if the lexicon is growing

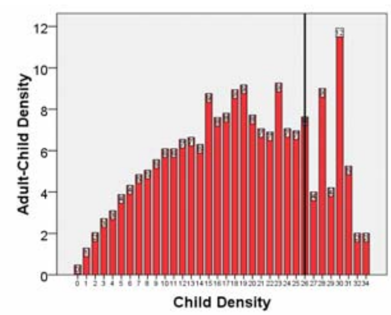
Question

Is the growth equivalent or asymmetric?

Asymmetric Growth



Asymmetric Growth



Study 1 Conclusion

- Adult and child density values highly correlated
  - Words that are dense in one corpus will tend to be dense in the other
- However, growth is asymmetric
  - Few words added to sparse neighborhoods
  - Many words added to dense neighborhoods
  - Difference between sparse and dense increases

Study 2: Set Size (Storkel & Adlof, in press a)

- How to determine semantic similarity for nonobjects?
- How do child and adult values compare?
  - Methodological issues: Do we need child values?
  - Theoretical issues: How does the lexicon grow from child to adult?

Determining Semantic Similarity

- Many methods used with real objects won't work for nonobjects
  - Ex. – frequency of co-occurrence
- Only viable method = discrete association
  - Developed by Nelson, McEvoy, & Schreiber for real words

Discrete Association Method

- Show picture (or word)
- Participant reports first word that comes to mind
- Word reported by two or more participants = neighbor
- Number of neighbors = set size

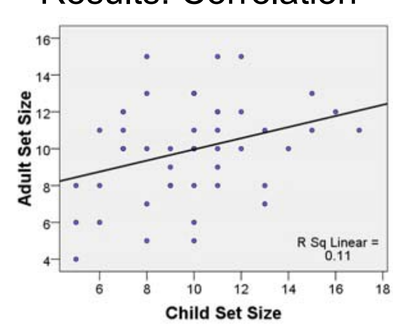
Our Study

- Participants
  - 82 undergraduate college students
  - 92 preschool children
- Stimuli
  - Nonobjects from Kroll & Potter (1984)
  - Only 47 of 88 tested on adults & children

Question

How do child and adult set sizes compare?

Results: Correlation



Results: T-test

- No significant difference
- Adult values tended to be similar to child values
  - No growth?!?
- Actual words reported as neighbors differed across children and adults (see article)

Study 2 Conclusion

- Adult and child set size values similar
  - (Although actual semantic neighbors differ)
- No semantic growth?
  - Unlikely
  - Probably an artifact of the stimuli (i.e., nonobjects)
  - Need to examine real words to determine growth pattern

Question

Do these variables influence word learning?

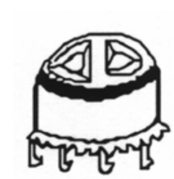
Study 3 & 4

Study 3 & 4: Basic Word Learning Method

- Participants
  - Typically developing preschool children
- Stimuli
  - Nonwords paired with nonobjects
  - Manipulate one variable
  - Hold other variables constant
- Exposure in game format on computer
- Test referent identification & naming
  - Prior to exposure (baseline)
  - End of exposure (learning)
  - 1-week post-exposure (retention)

Exposure Example

This is a goʊm. Say goʊm. [goʊm]. Yes, that's a goʊm. Remember, it's a goʊm. We're going to play a game. Find the goʊm. That's the goʊm. Say goʊm. [goʊm]. Don't forget the goʊm.



Naming Example

goʊm

Referent Identification Example

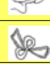
goʊm

2

Study 3: Neighborhood Density (Storkel & Rosales, in progress)

- Participants: 25 preschool children
- Vary neighborhood density of nonwords
  - Low vs. High
- Hold semantic set size of nonobjects constant
  - Mid

Stimuli

Low		High	
dɔɪk		jʌt	
gɪf		boʊg	
paɪb		tab	
jam		faʊn	
nɛp		wæd	

Stimuli Characteristics

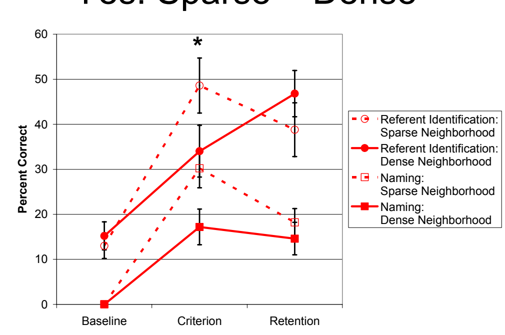
	Low	High
Neighborhood Density	5 (1)	12 (2)
Semantic Set Size	10 (3)	11 (3)

THE UNIVERSITY OF KANSAS

Question

Does neighborhood density influence word learning?

Yes: Sparse > Dense



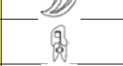
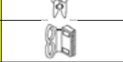
Study 3 Summary

- Sparse learned better than dense
  - Only at criterion test (i.e., immediate learning)
  - Not at retention test

Study 4: Set Size (Storkel & Adlof, in press b)

- Participants: 36 preschool children
- Hold neighborhood density of nonwords constant
  - Mid
- Vary semantic set size of nonobjects
  - Low vs. High

Stimuli

Low		High	
goʊm		boʊg	
maɪf		wun	
hɛg		pig	
jeɪm		jʌt	
fip		meɪb	

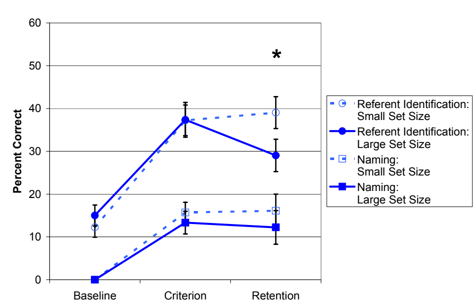
Stimuli Characteristics

	Low	High
Neighborhood Density	10 (2)	11 (2)
Semantic Set Size	8 (2)	13 (3)

Question

Does semantic set size influence word learning?

Yes: Small > Large



Study 4 Summary

- Small set size learned better than large set size
  - Not at criterion test (i.e., immediate learning)
  - Only at retention test

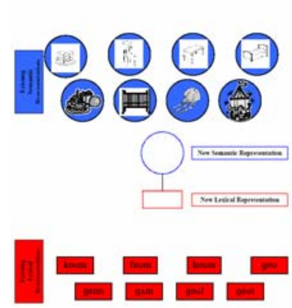
Study 3 & 4 Summary

- Lexical and semantic similarity both influence learning
  - Less similar items learned better than more similar
- Timing of influence varies
  - Lexical = immediate learning
  - Semantic = retention

Hypothesis 1

- Lexical similarity influences early components of word learning
  - Triggering of learning
    - Low similarity = more unique
    - More obvious that lexical representation does not exist for new word
    - Learning immediately triggered

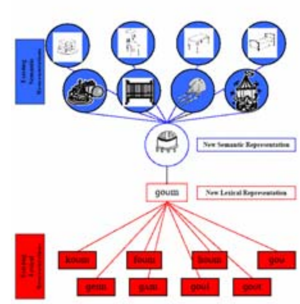
Step 1: Triggering



Hypothesis 1

- Semantic similarity influences later components of word learning
  - Engagement of new representation with existing representations
    - Low similarity = less confusability with existing representations
    - Less degradation over time

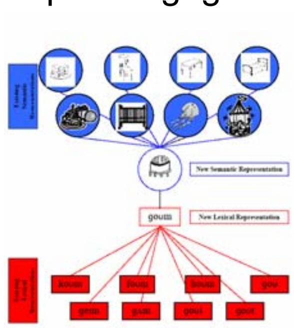
Step 3: Engagement



Hypothesis 2

- Lexical and semantic similarity both affect engagement but timing of engagement varies
  - Early engagement for lexical representations
    - Overlap in word learning "steps"
  - Late engagement for semantic representations
    - Less overlap in word learning steps

Step 3: Engagement



Future Directions

- Identify measures that are sensitive to
  - Triggering
  - Configuration
  - Engagement
- Test the role of lexical and semantic similarity in each process

Acknowledgements

DC 008095 (Storkel)
DC 009133 (Adlof)
DC 009135 (Hoover)
DC 000052 (T-Grant)
DC 005803 (BNCD)

Staff of the Word & Sound Learning Lab