

---

# **Validating Bundled Gap Filling – Empirical Evidence for Ambiguity Reduction and Language Proficiency Testing Capabilities**

Niklas Meyer, Michael Wojatzki & Torsten Zesch

# Types of Gap-fill Exercises

The students have to \_\_\_\_ the test.

The students have to \_\_\_\_ the test.

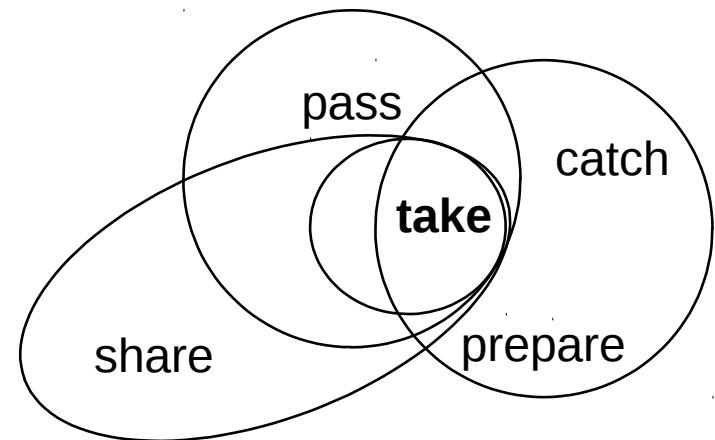
- a) take
- b) fold
- c) entertain
- d) fry

- high ambiguity
- production task

- low/moderate ambiguity
- recognition task

# Bundled Gap-fill [Wojatzki et al. 2016]

The students have to \_\_\_\_ the test.  
Their cook will \_\_\_\_ three salmons.  
All passengers should \_\_\_\_ their seats.  
Both authors \_\_\_\_ credit for this.



- low ambiguity
- production task

# Approach

$g_1$ : "The kids have to \_\_\_\_ their lunch."

For any given context, find additional contexts that resolve as much ambiguity as possible

$$D(b) = \log \frac{P(F(b) = t)}{\max_{w \in V \setminus \{t\}} P(F(b) = w)}$$

- Evaluation so far:
  - bundles decrease error rates
  - $r = .66$

	Average Success Rate	Average Disambiguation Measure
Single Gap	.27	-0.50
Bundle <sub>2</sub>	.59	4.00
Bundle <sub>3</sub>	.68	7.75
Bundle <sub>4</sub>	.78	11.06

# Ambiguity Reduction & Language Proficiency Testing

- RQ1: Ambiguity reduction due to algorithm or bundles?
- Compare 3 conditions:
  - 1.maximize:
  - 2.minimize:
  - 3.random: randomly created bundles

RQ2: Relation to language proficiency tests (*cTest*)?

# Results on the poster



**ASK ME ANYTHING**