Explaining differential performance on academic vocabulary assessments for English language learners (ELLs) using explanatory item response models

Autumn McIlraith, Paulina Kulesz, David Francis

University of Houston

Joshua Lawrence, Rebecca Knopf

University of Oslo

SSSR, 2019

Acknowledgments

- This work is supported by an ongoing IES Measurement Grant: R305A170151, *Improving the Accuracy of Academic Vocabulary Assessment for English Language Learners*, David Francis, PI.

- Data for the current study comes from the *Word Generation* project (Catherine Snow, https://wordgen.serpmedia.org/).

- Thanks to Martin Walczak for help with item coding

Background

- English Language Learners (ELLs) are a growing demographic in U.S. schools.
- ELLs frequently struggle to meet benchmarks for reading proficiency (U.S. Dept of Education, 2017).
- A challenge for assessment
  - Standardized assessments: norming sample typically English-only
  - Differential item functioning?
- For middle school students: Academic Vocabulary is a key ingredient in success across academic disciplines (e.g., Anderson & Freebody, 1981; Coxhead, 2000)

Word Generation study

- Focus: **"all-purpose" academic vocabulary words**, often less explicitly taught, but important for comprehension of discipline-specific texts (Anderson & Freebody, 1981; Coxhead, 2000)
- See Snow, Lawrence, & White (2009) for details of the intervention

Journal of Research on Educational Effectiveness, 2: 325–344, 2009

ISSN: 1934-5747 print / 1934-5739 online
DOI: 10.1080/19345740903167042
Routledge
Taylor & Francis Group

Generating Knowledge of Academic Language Among Urban Middle School Students

Catherine E. Snow, Joshua F. Lawrence, and Claire White
Harvard Graduate School of Education and Strategic Education Research Partnership, Cambridge, Massachusetts, USA

Word Generation study

- Focus: **“all-purpose” academic vocabulary words**, often less explicitly taught, but important for comprehension of discipline-specific texts (Anderson & Freebody, 1981; Coxhead, 2000)
- See Snow, Lawrence, & White (2009) for details of the intervention
- Students in Grades 6, 7, and 8, in 13 middle schools in a large urban district in California
- **The current study used data from the pre-test (fall semester), prior to the start of the Word Generation intervention.**

Current study

Pre-test data were collected from middle school students, classified into the following categories by the California school district:

(N represents number included in the current analyses)

	N
• **EO:** English-only speakers	3,600
• **LEP:** limited English proficiency, continued to qualify for language support	1,851
• **IFEP:** initially fluent: proficient in English at the start of the study	1,034
• **RFEP:** reclassified as fully English proficient, started as limited	3,793
	10,278

Current study: Items

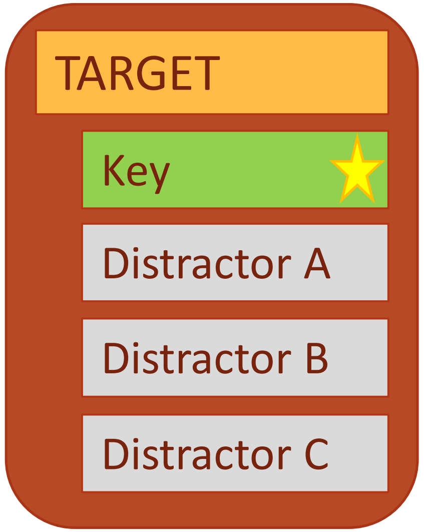
- Synonym task
- 50 items on each of two forms, some unique and some shared across forms (81 unique items)
- **Distractors:** semantically, phonologically, or orthographically related, or unrelated

XX. He <u>acquired</u> a pet.

a. got
b. trained
c. lost
d. adored

XX. We had <u>sufficient</u> food at the party.

a. delicious
b. too much
c. standard
d. enough



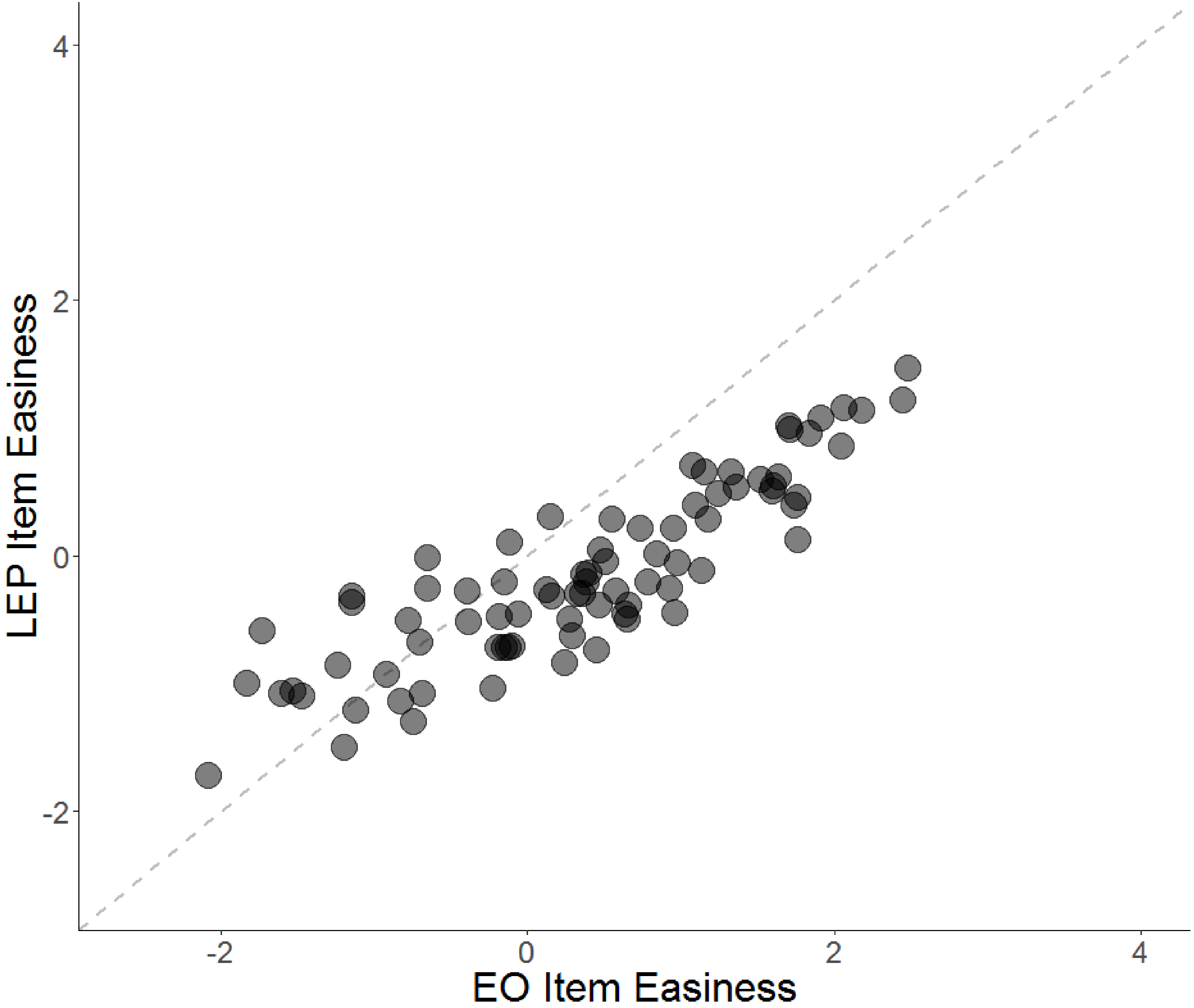
Differential Item Functioning (DIF) analyses

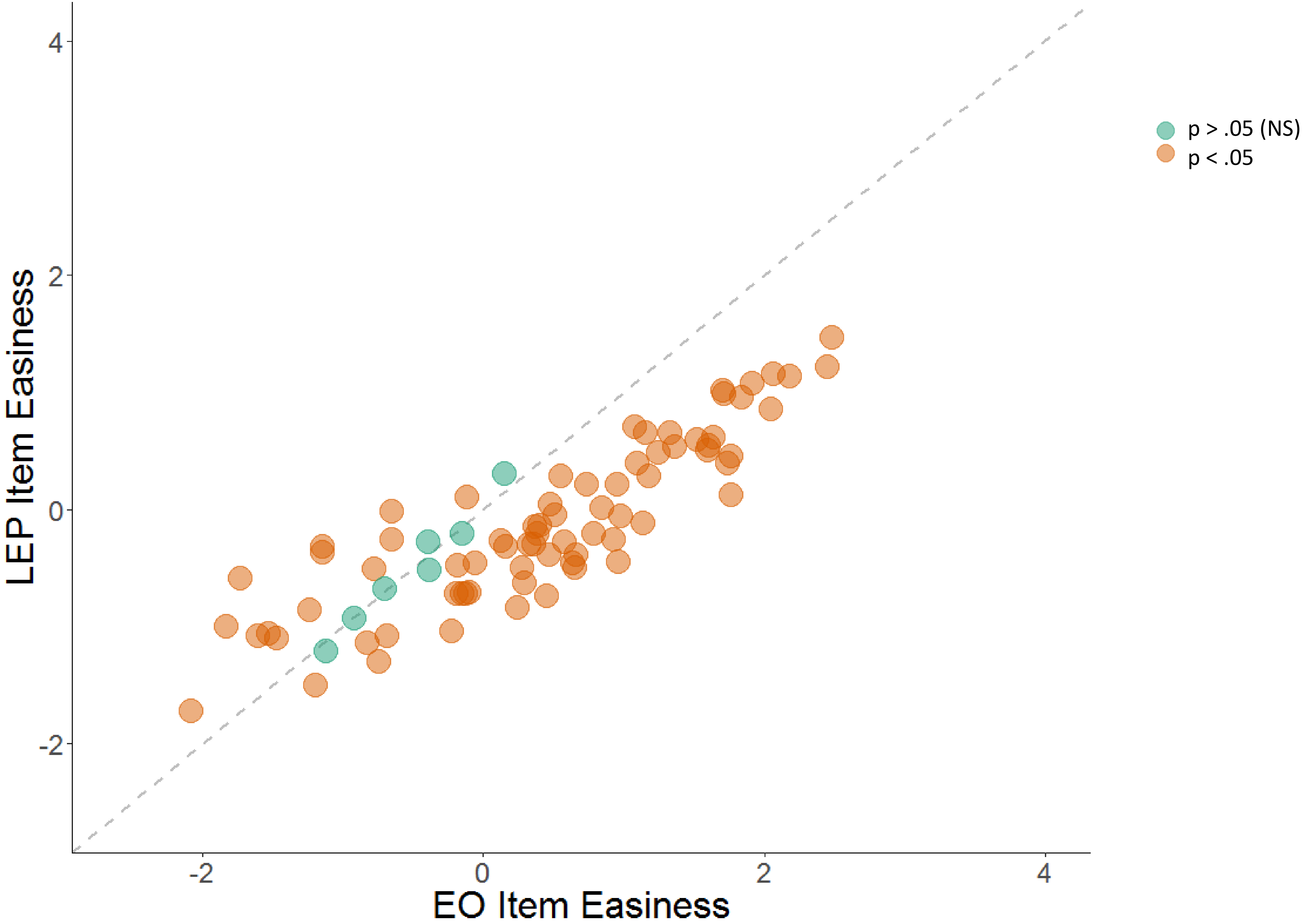
- Item response theory (IRT) approach
- 1PL IRT models (uniform DIF), models estimated in GLIMMIX

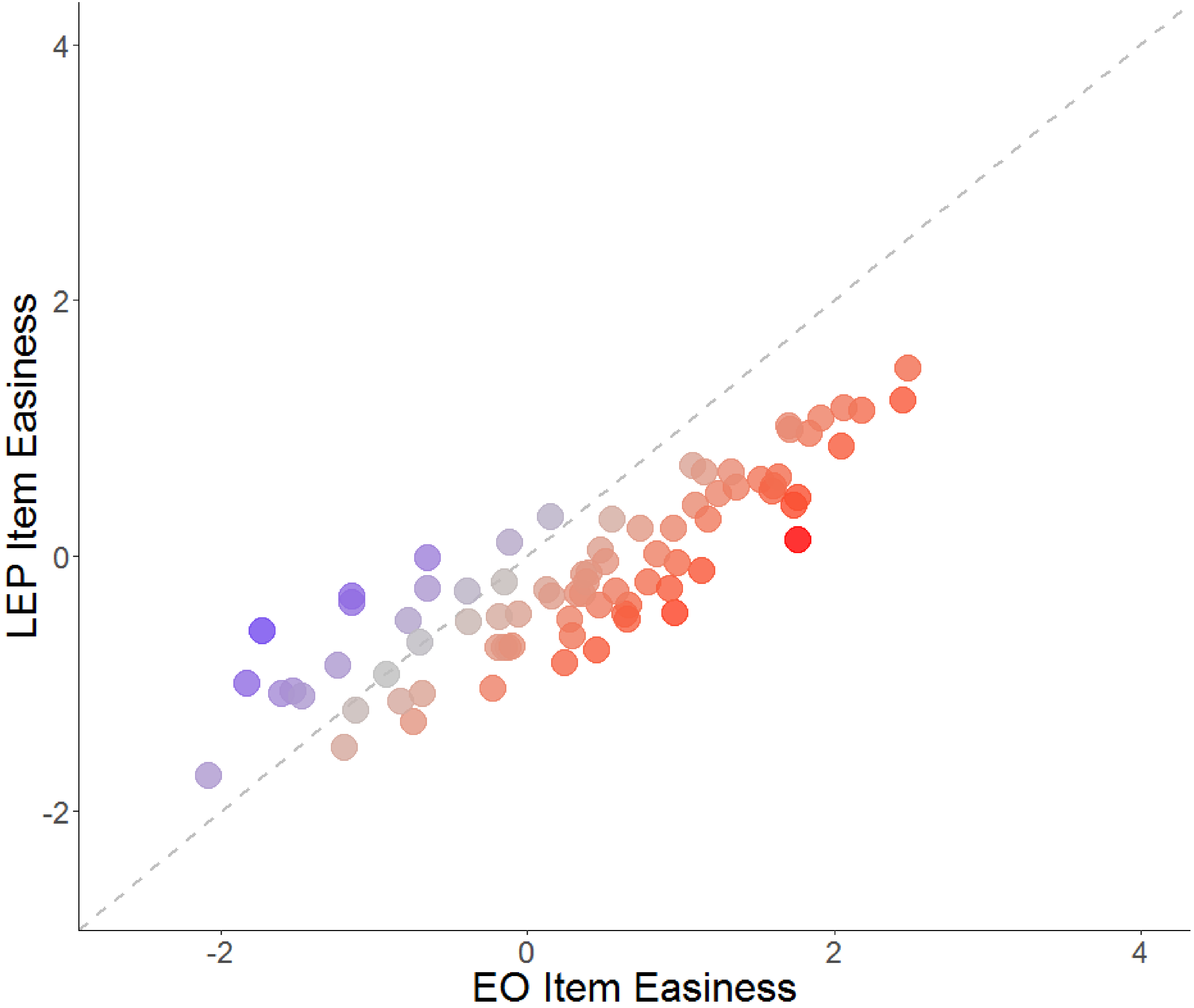
- Comparing groups: do two students with the same underlying ability have equal probabilities of getting an item correct?
  - Reference group: EO (English-only)
  - Focal group: LEP (Limited), IFEP (Initially fluent), or RFEP (Reclassified)

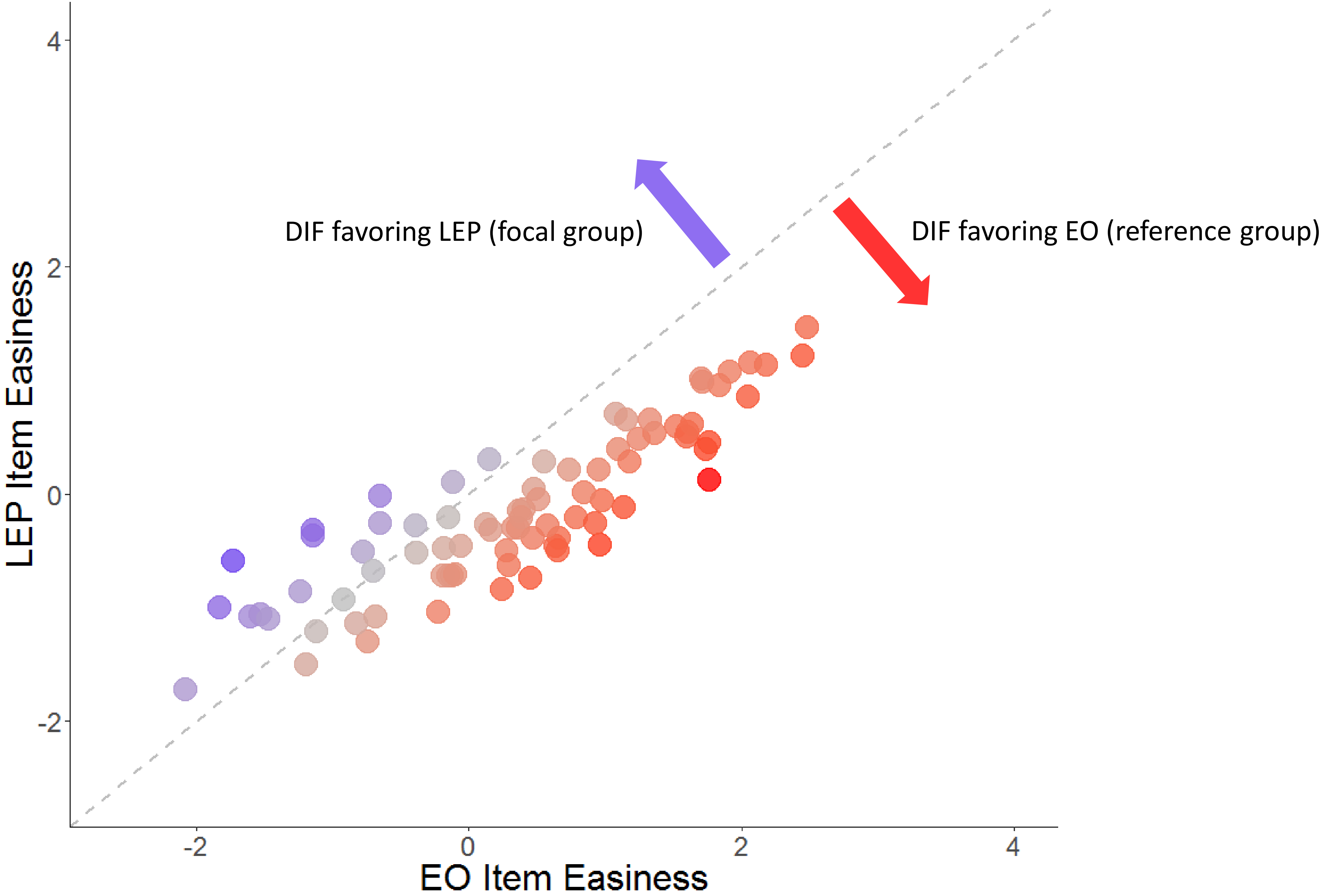
- Predicted item “easiness” (higher values = easier items) for each group
- DIF test: difference between the easiness estimates, statistical significance?

DIF Results: EO vs. LEP


LEP Item Easiness
4
2
0
-2
-2
0
2
4
EO Item Easiness

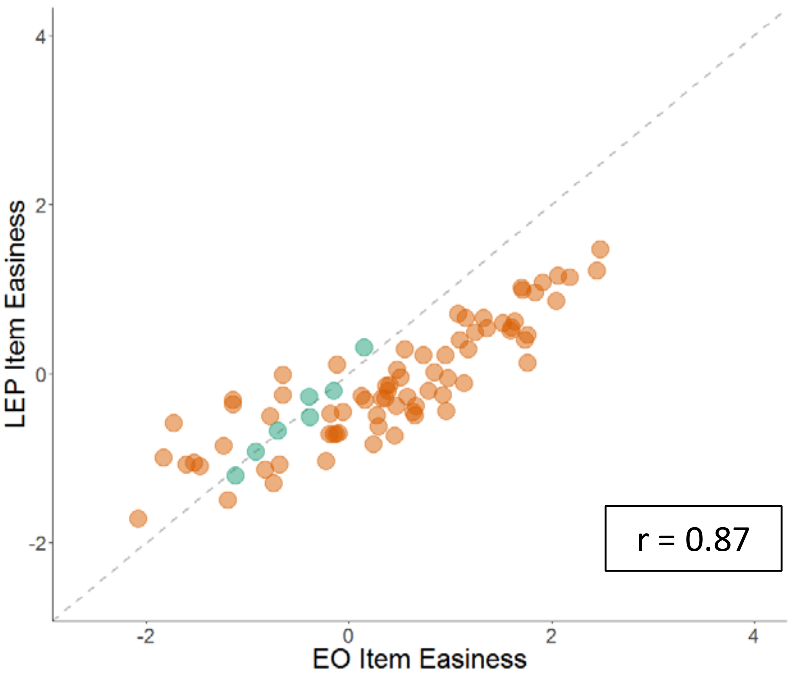

4
2
0
-2
LEP Item Easiness
-2
0
2
4
EO Item Easiness
p > .05 (NS)
p < .05


LEP Item Easiness
4
2
0
-2
-2
0
2
4
EO Item Easiness

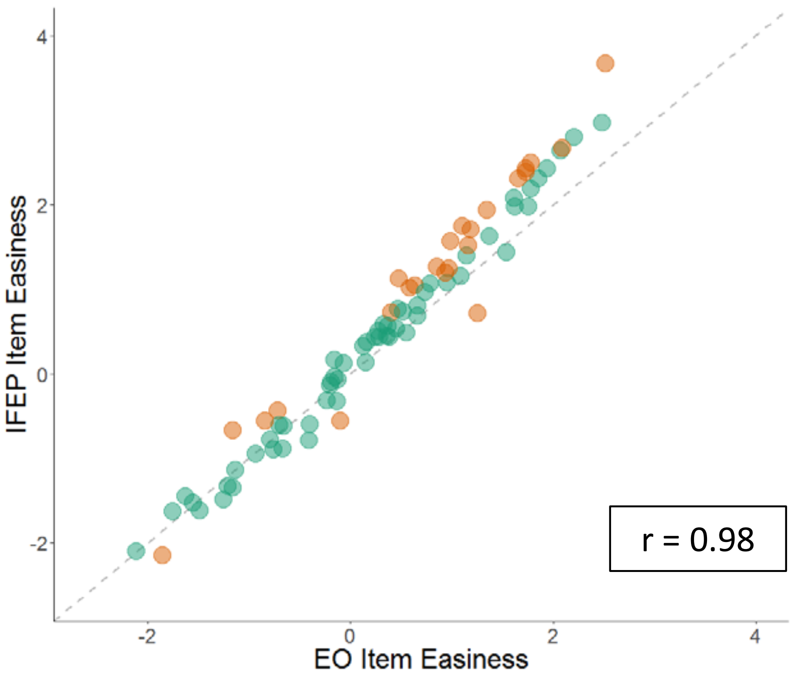

DIF favoring LEP (focal group)
DIF favoring EO (reference group)
LEP Item Easiness
EO Item Easiness
4
2
0
-2
-2
0
2
4

DIF Results

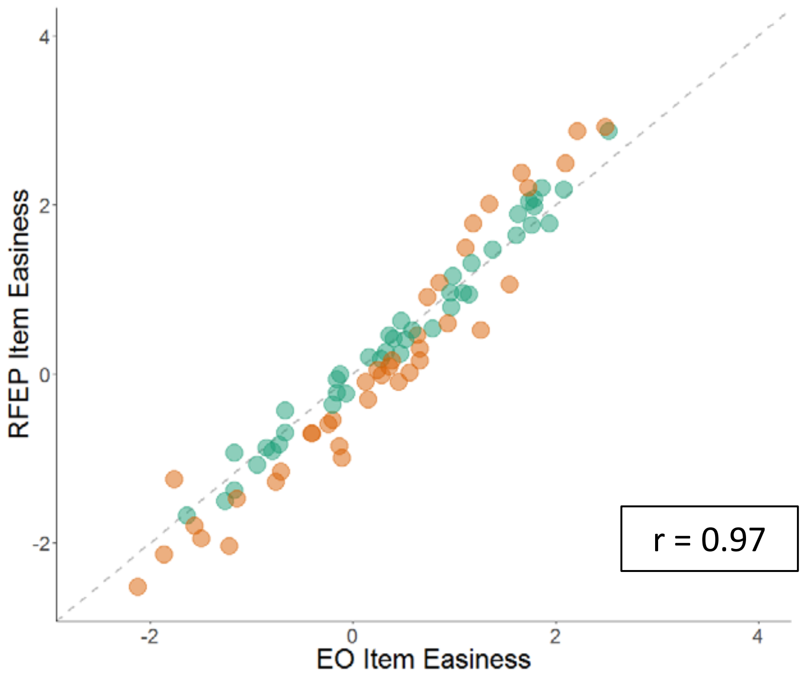
EO vs. LEP



EO vs. IFEP

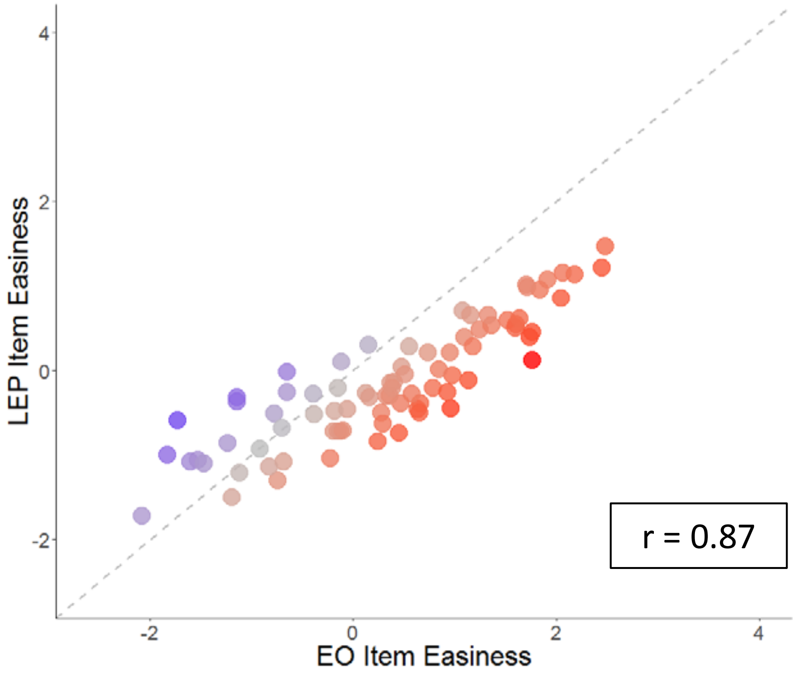


EO vs. RFEP

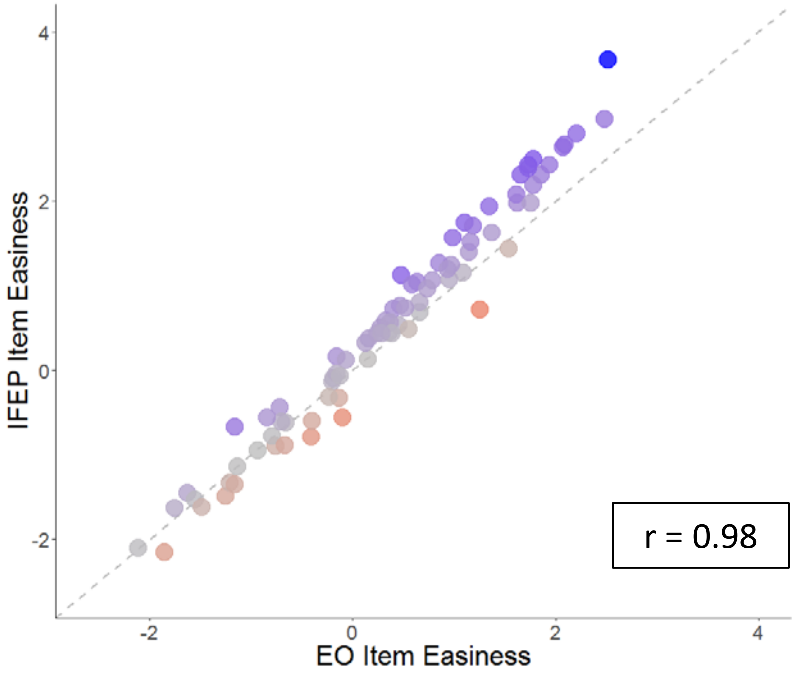


DIF Results

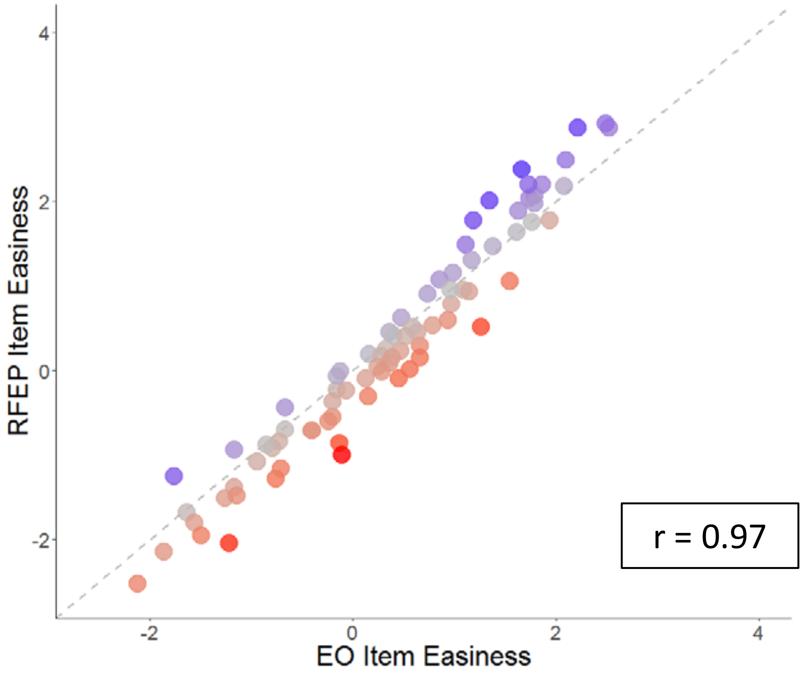
EO vs. LEP



EO vs. IFEP

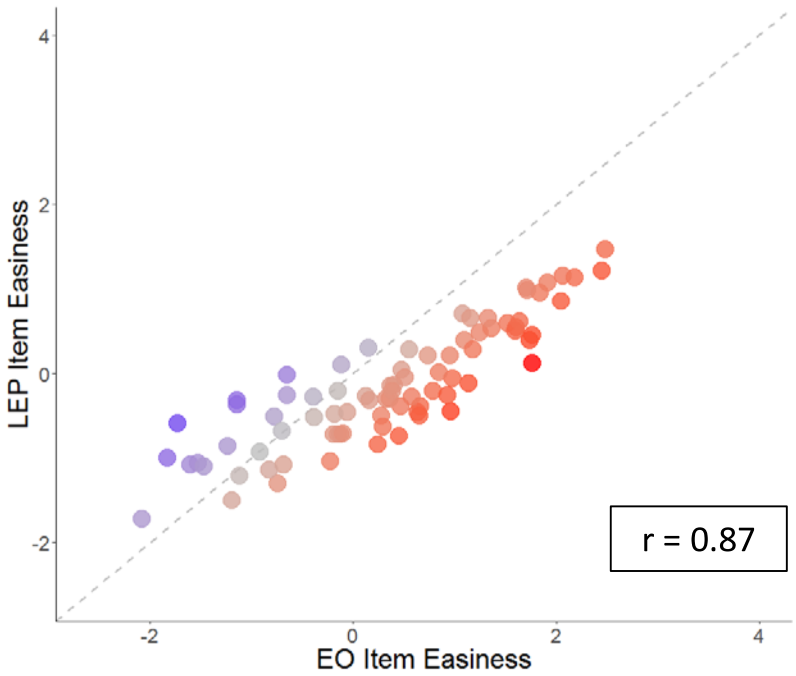


EO vs. RFEP

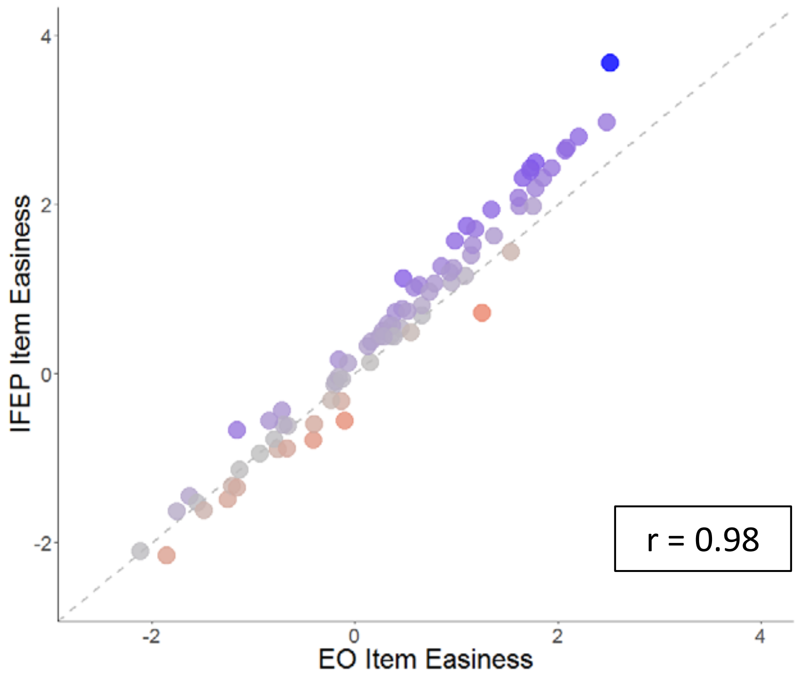


DIF Results

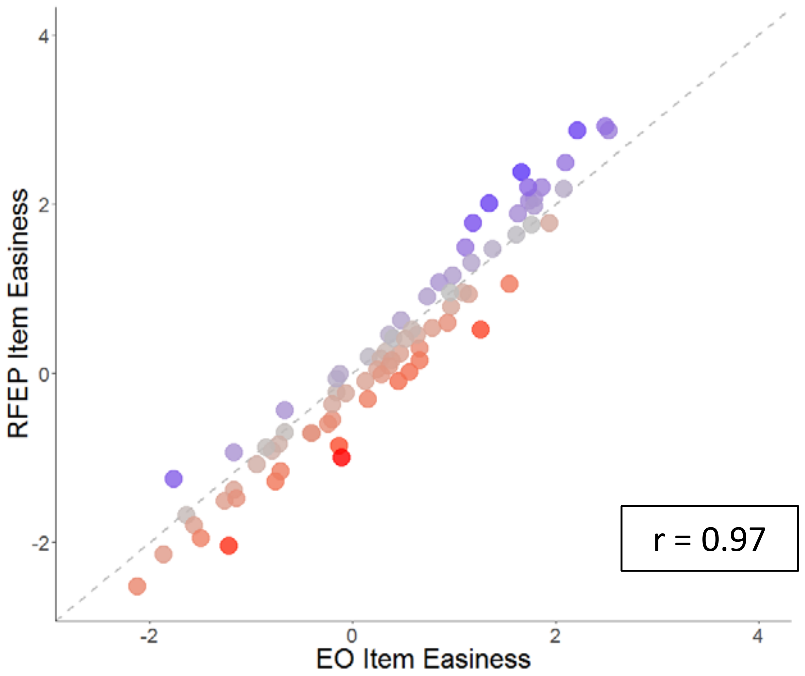
EO vs. LEP



EO vs. IFEP

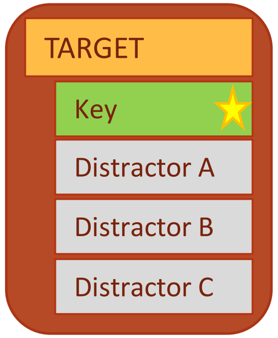


EO vs. RFEP



Explanatory IRT (eIRT) and other analyses

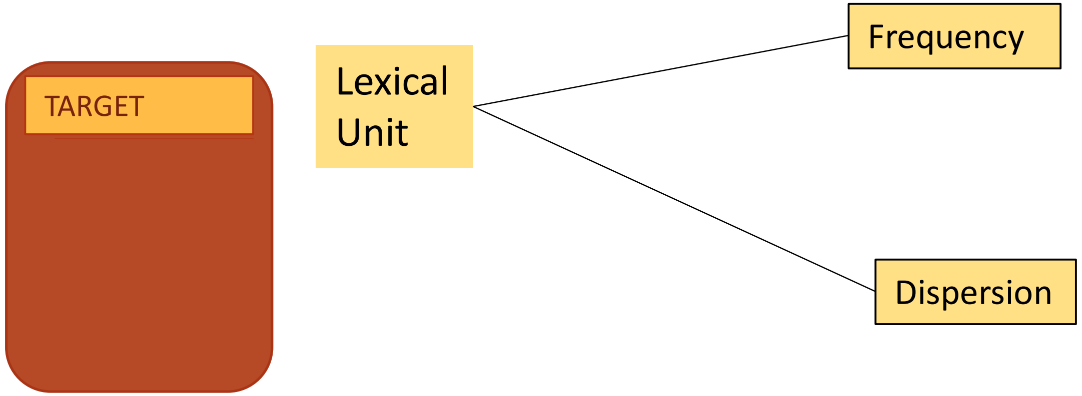
Item characteristics used as predictors

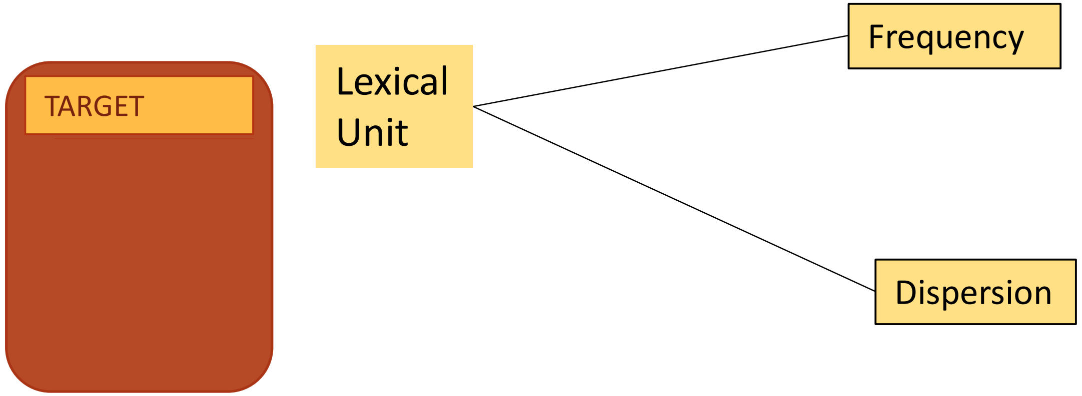


XX. He acquired a pet.

a. got
b. trained
c. lost
d. adored

TARGET

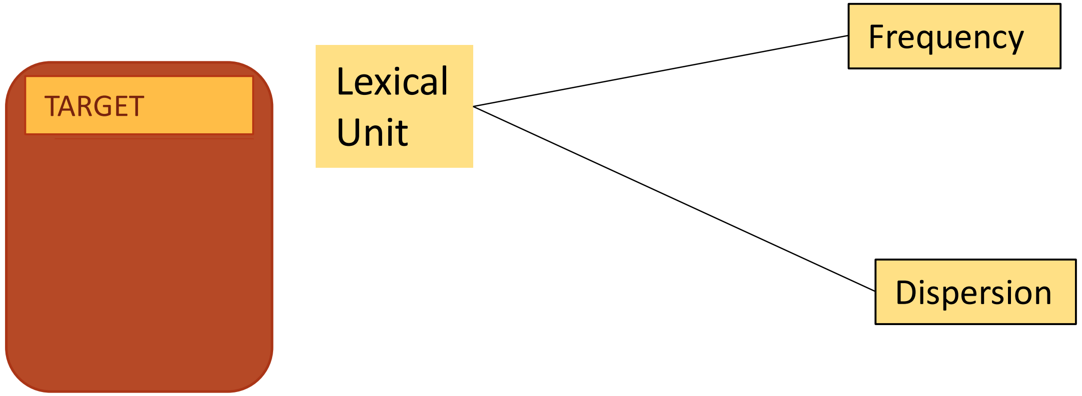

TARGET
Lexical Unit
Frequency
Dispersion



Word frequency (Zeno, zipfian transform)

Word Frequency (Brown corpus, log transform)

Word Frequency (HAL corpus, log transform)



Word frequency (Zeno, zipfian transform)

Word Frequency (Brown corpus, log transform)

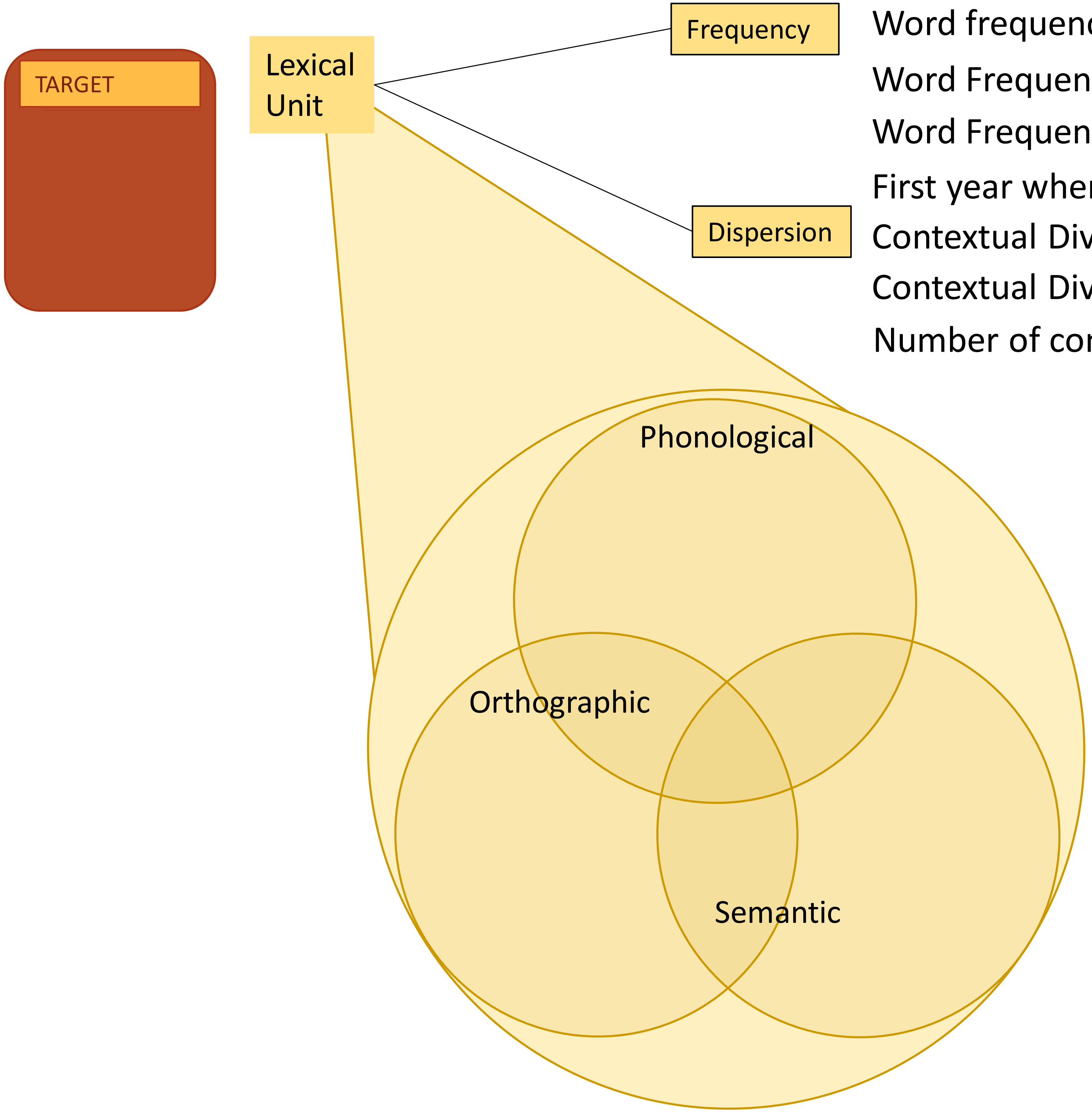
Word Frequency (HAL corpus, log transform)

First year when word appeared

Contextual Diversity (Adelman)

Contextual Diversity (Subtlex)

Number of content areas where word appears (Zeno)



Word frequency (Zeno, zipfian transform)

Word Frequency (Brown corpus, log transform)

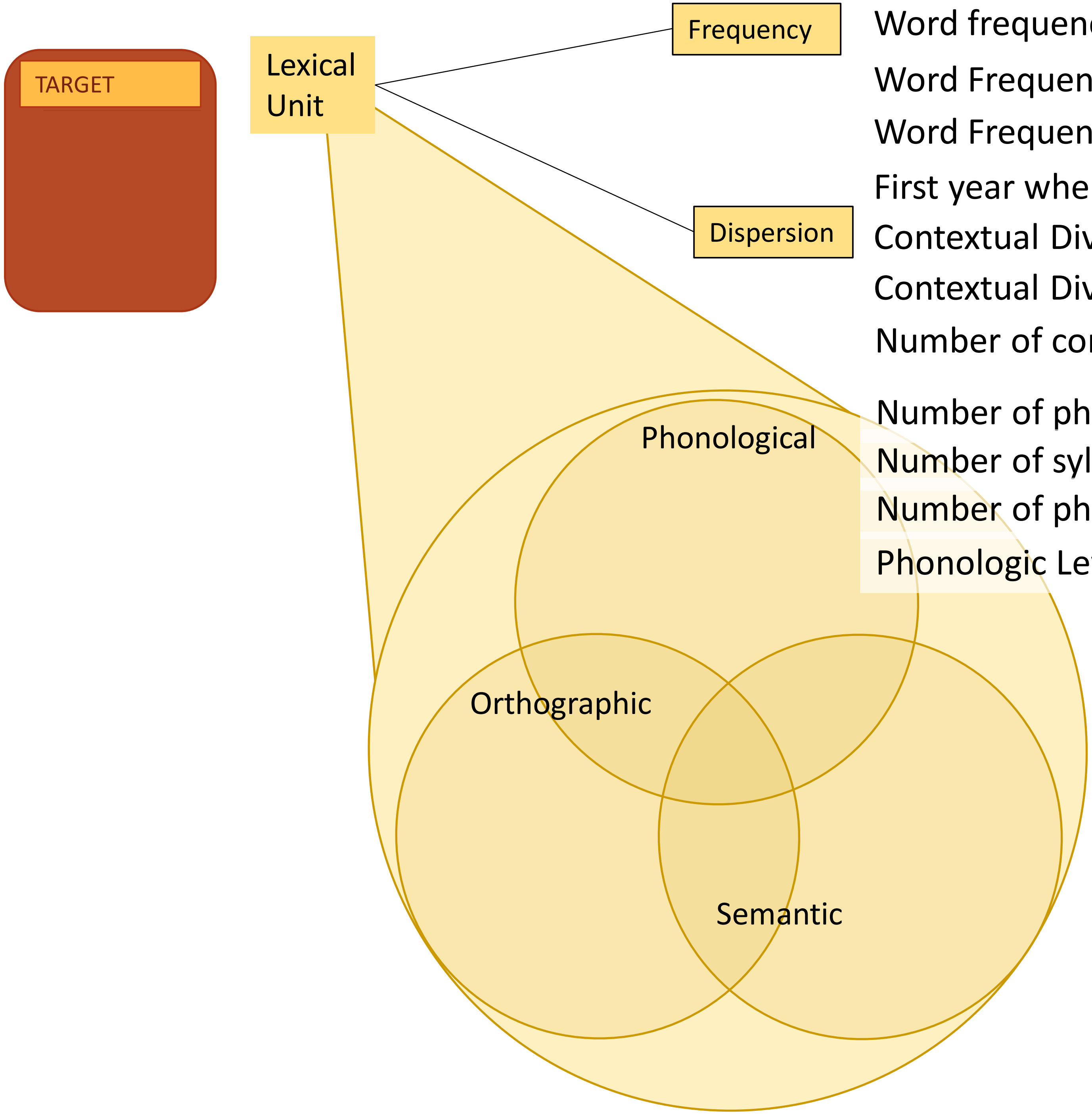
Word Frequency (HAL corpus, log transform)

First year when word appeared

Contextual Diversity (Adelman)

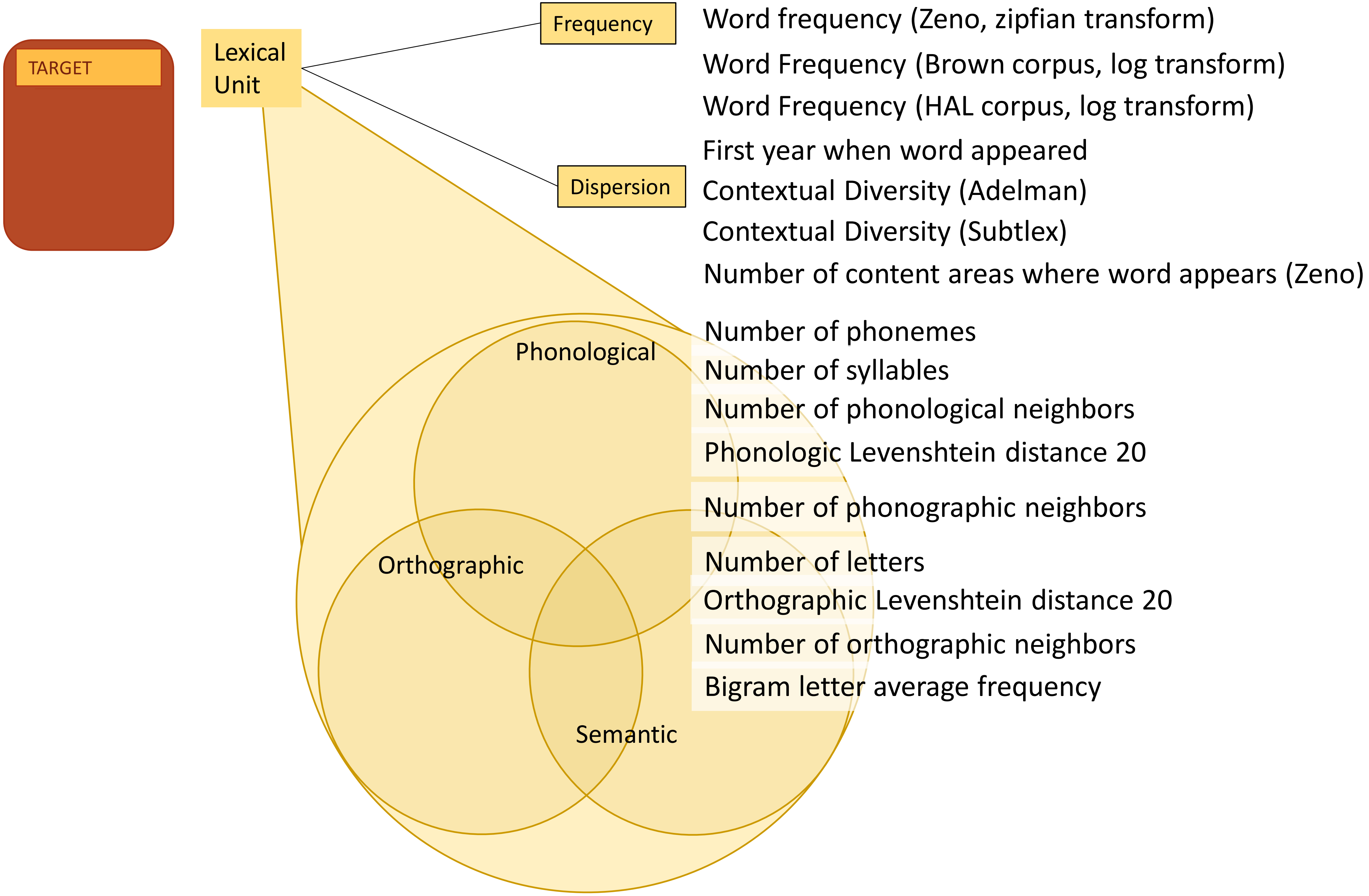
Contextual Diversity (Subtlex)

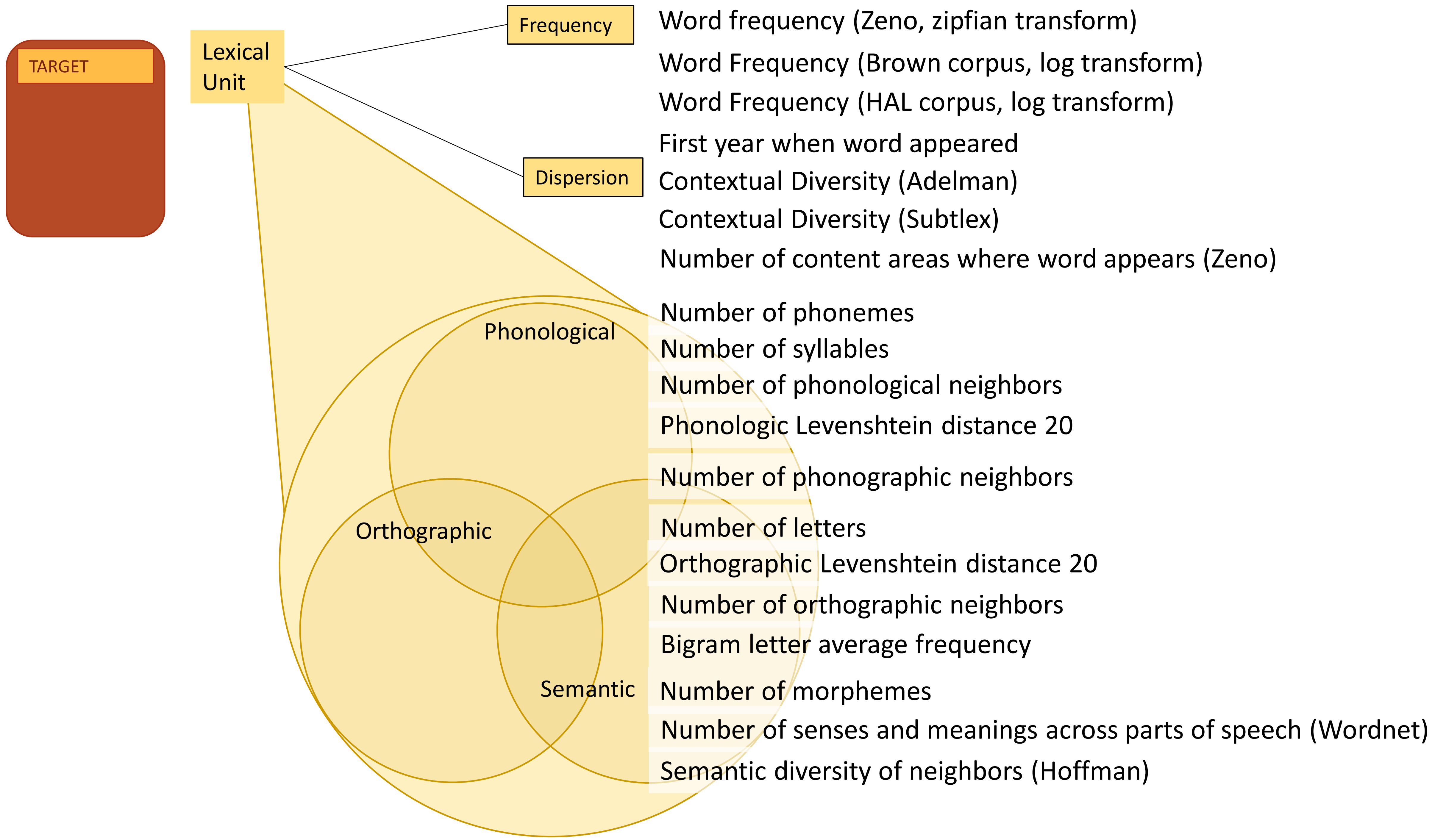
Number of content areas where word appears (Zeno)



Word frequency (Zeno, zipfian transform)
Word Frequency (Brown corpus, log transform)
Word Frequency (HAL corpus, log transform)
First year when word appeared
Contextual Diversity (Adelman)
Contextual Diversity (Subtlex)
Number of content areas where word appears (Zeno)

Number of phonemes
Number of syllables
Number of phonological neighbors
Phonologic Levenshtein distance 20


TARGET
Lexical Unit
Frequency
Dispersion
Word frequency (Zeno, zipfian transform)
Word Frequency (Brown corpus, log transform)
Word Frequency (HAL corpus, log transform)
First year when word appeared
Contextual Diversity (Adelman)
Contextual Diversity (Subtlex)
Number of content areas where word appears (Zeno)
Phonological
Orthographic
Semantic
Number of phonemes
Number of syllables
Number of phonological neighbors
Phonologic Levenshtein distance 20
Number of phonographic neighbors
Number of letters
Orthographic Levenshtein distance 20
Number of orthographic neighbors
Bigram letter average frequency


TARGET
Lexical Unit
Frequency
Dispersion
Phonological
Orthographic
Semantic
Word frequency (Zeno, zipfian transform)
Word Frequency (Brown corpus, log transform)
Word Frequency (HAL corpus, log transform)
First year when word appeared
Contextual Diversity (Adelman)
Contextual Diversity (Subtlex)
Number of content areas where word appears (Zeno)
Number of phonemes
Number of syllables
Number of phonological neighbors
Phonologic Levenshtein distance 20
Number of phonographic neighbors
Number of letters
Orthographic Levenshtein distance 20
Number of orthographic neighbors
Bigram letter average frequency
Number of morphemes
Number of senses and meanings across parts of speech (Wordnet)
Semantic diversity of neighbors (Hoffman)

Explanatory IRT (eIRT) and other analyses

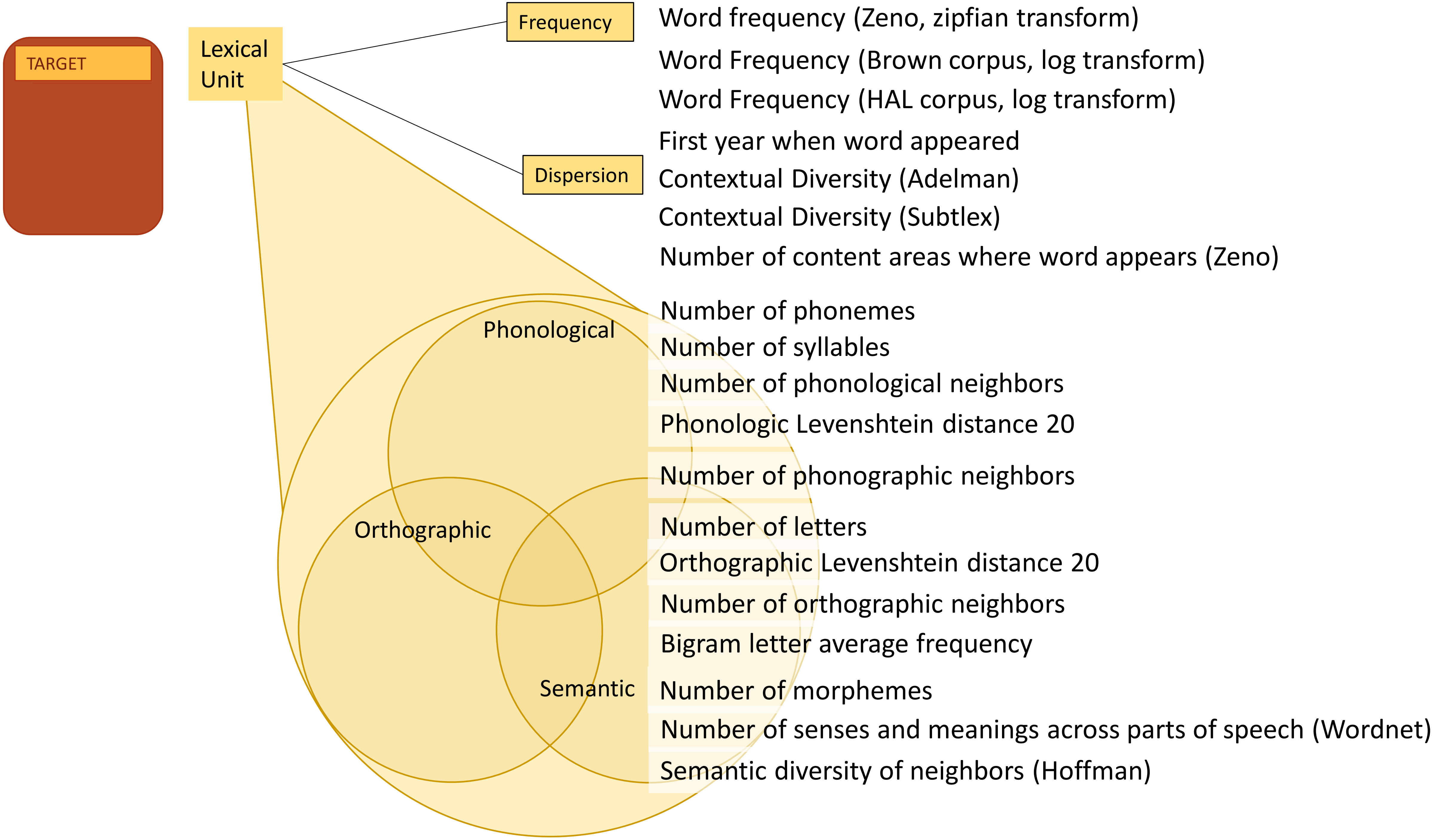
Item characteristics used as predictors

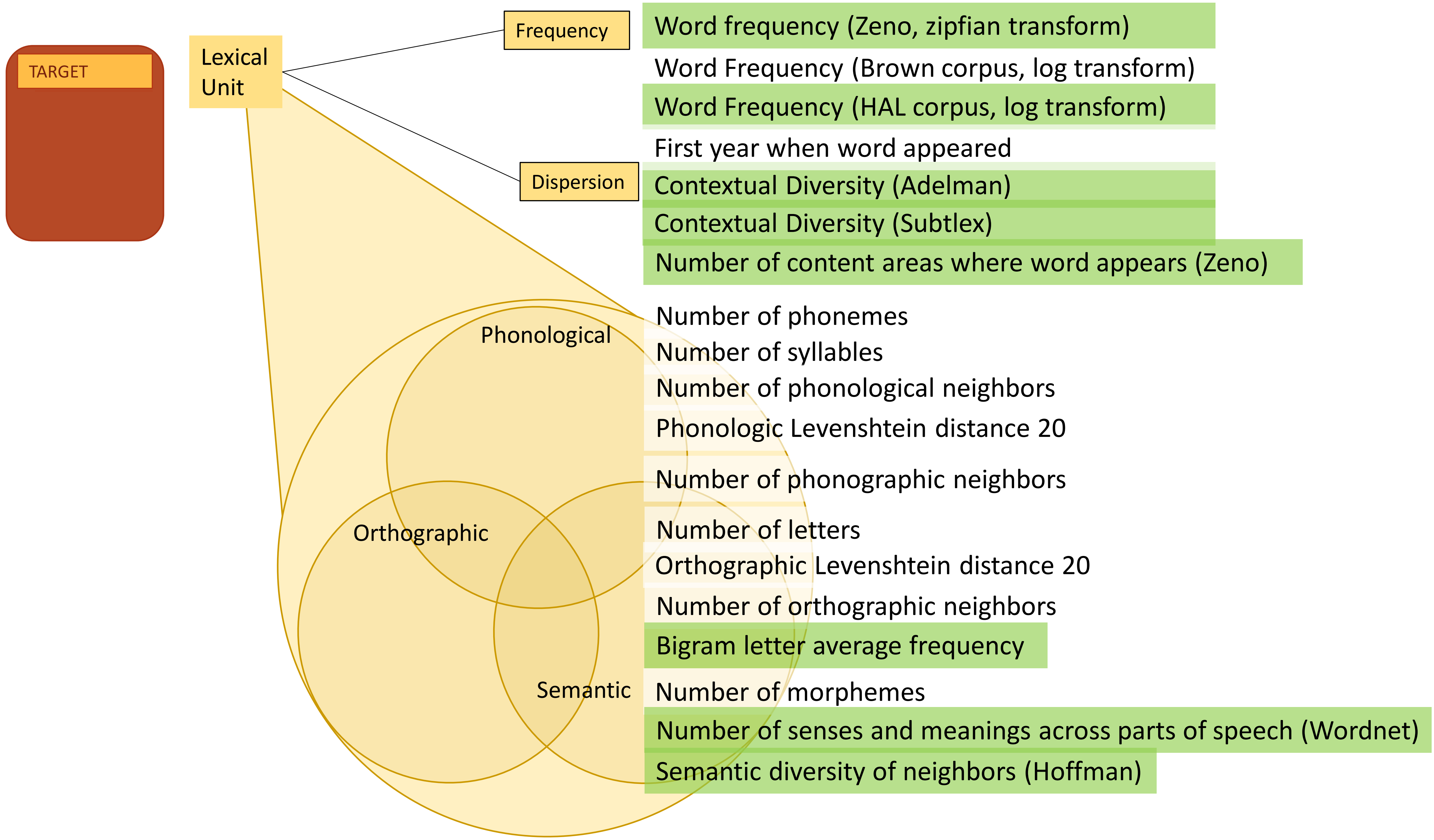
eIRT approach: binary 1/0 (correct/incorrect) responses modeled as outcome

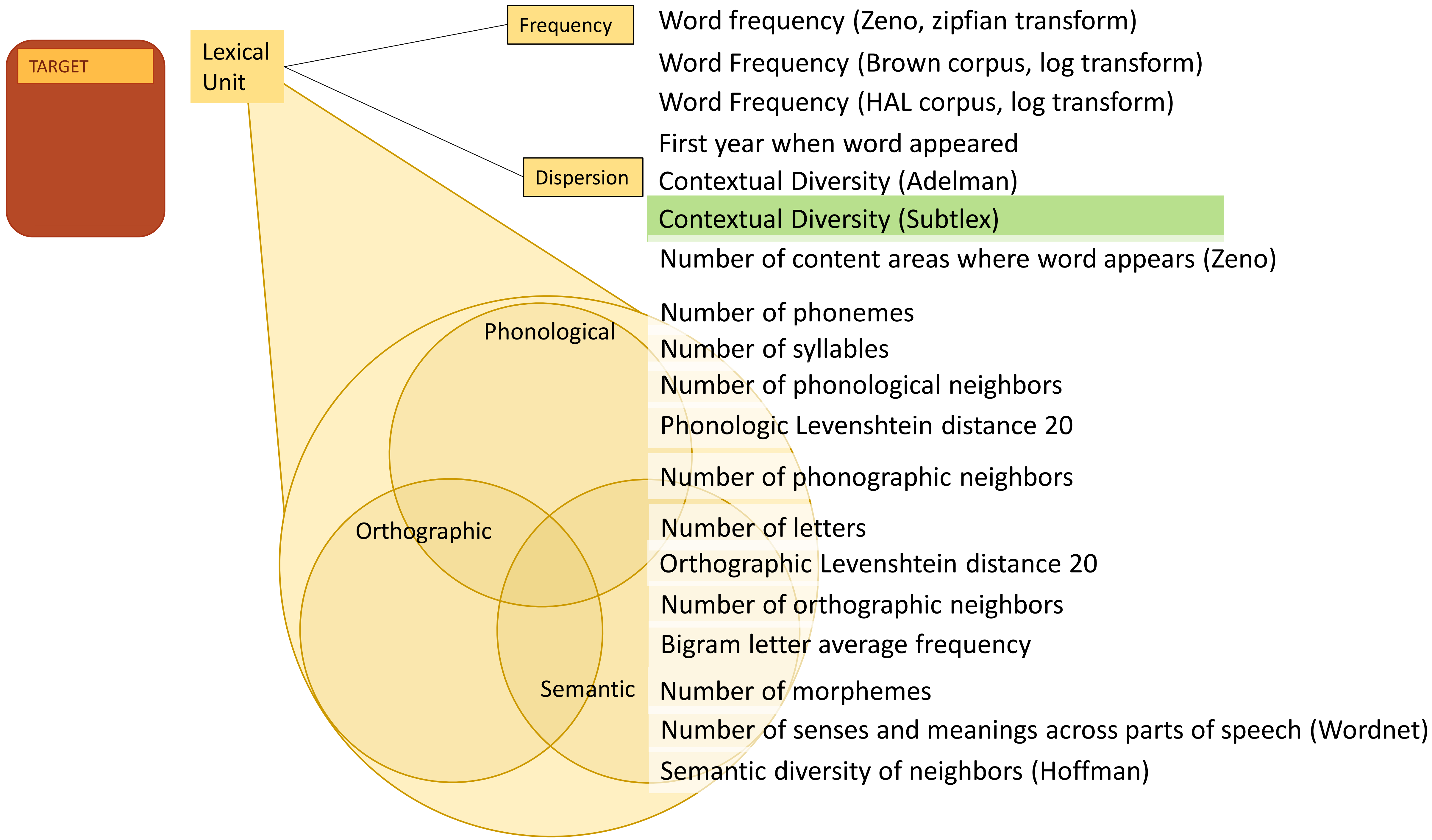
Supplemented with **regression approach:** Model parameters from GLIMMIX as outcomes

- EO item easiness
- Focal group item easiness
- DIF

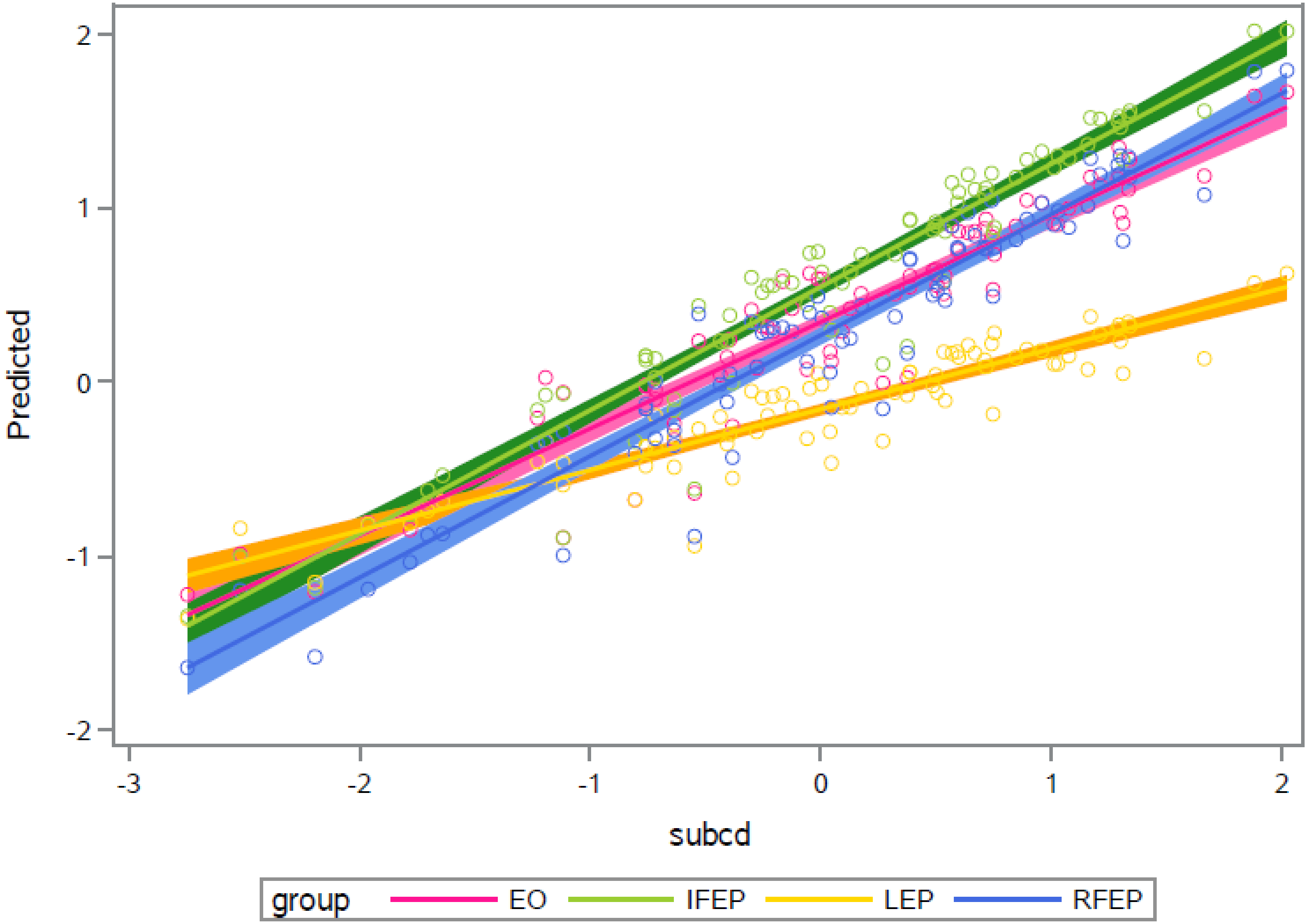
- **Stepwise regression**
- **Best subsets regression**

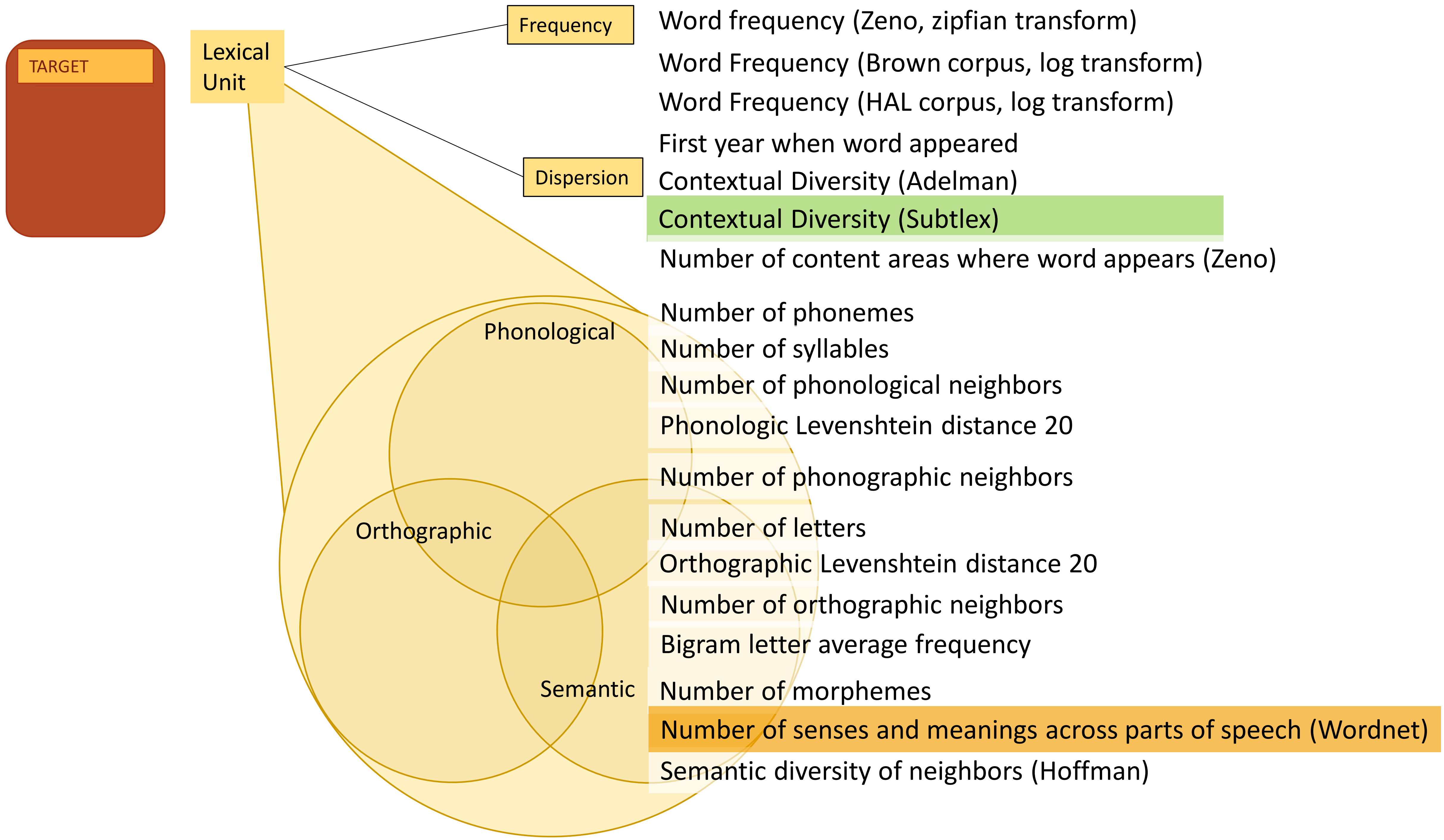

TARGET
Lexical Unit
Frequency
Dispersion
Phonological
Orthographic
Semantic
Word frequency (Zeno, zipfian transform)
Word Frequency (Brown corpus, log transform)
Word Frequency (HAL corpus, log transform)
First year when word appeared
Contextual Diversity (Adelman)
Contextual Diversity (Subtlex)
Number of content areas where word appears (Zeno)
Number of phonemes
Number of syllables
Number of phonological neighbors
Phonologic Levenshtein distance 20
Number of phonographic neighbors
Number of letters
Orthographic Levenshtein distance 20
Number of orthographic neighbors
Bigram letter average frequency
Number of morphemes
Number of senses and meanings across parts of speech (Wordnet)
Semantic diversity of neighbors (Hoffman)


TARGET
Lexical Unit
Frequency
Dispersion
Word frequency (Zeno, zipfian transform)
Word Frequency (Brown corpus, log transform)
Word Frequency (HAL corpus, log transform)
First year when word appeared
Contextual Diversity (Adelman)
Contextual Diversity (Subtlex)
Number of content areas where word appears (Zeno)
Phonological
Orthographic
Semantic
Number of phonemes
Number of syllables
Number of phonological neighbors
Phonologic Levenshtein distance 20
Number of phonographic neighbors
Number of letters
Orthographic Levenshtein distance 20
Number of orthographic neighbors
Bigram letter average frequency
Number of morphemes
Number of senses and meanings across parts of speech (Wordnet)
Semantic diversity of neighbors (Hoffman)

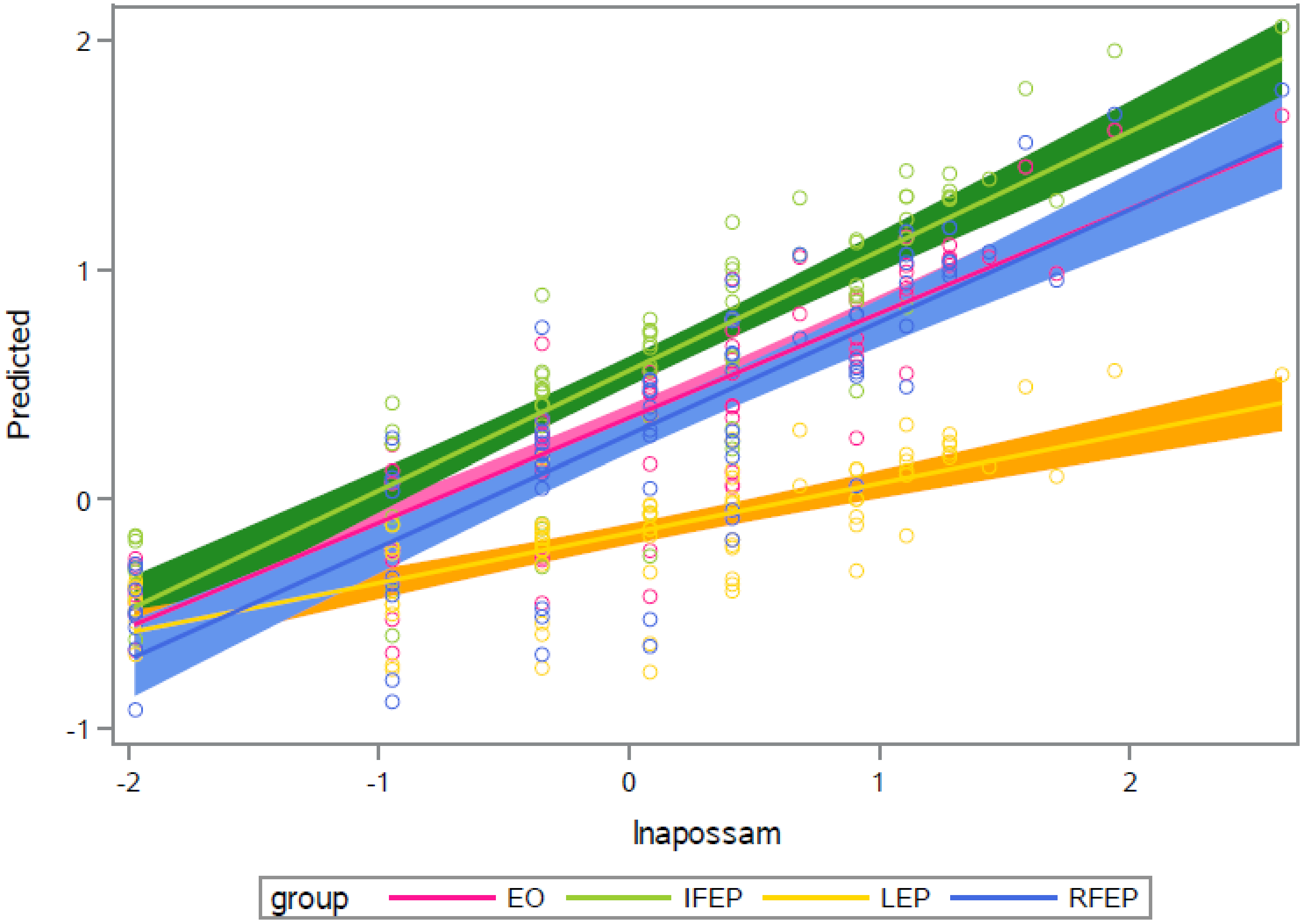

TARGET
Lexical Unit
Frequency
Dispersion
Word frequency (Zeno, zipfian transform)
Word Frequency (Brown corpus, log transform)
Word Frequency (HAL corpus, log transform)
First year when word appeared
Contextual Diversity (Adelman)
Contextual Diversity (Subtlex)
Number of content areas where word appears (Zeno)
Phonological
Orthographic
Semantic
Number of phonemes
Number of syllables
Number of phonological neighbors
Phonologic Levenshtein distance 20
Number of phonographic neighbors
Number of letters
Orthographic Levenshtein distance 20
Number of orthographic neighbors
Bigram letter average frequency
Number of morphemes
Number of senses and meanings across parts of speech (Wordnet)
Semantic diversity of neighbors (Hoffman)

eIRT: contextual diversity*group




TARGET
Lexical Unit
Frequency
Dispersion
Word frequency (Zeno, zipfian transform)
Word Frequency (Brown corpus, log transform)
Word Frequency (HAL corpus, log transform)
First year when word appeared
Contextual Diversity (Adelman)
Contextual Diversity (Subtlex)
Number of content areas where word appears (Zeno)
Phonological
Orthographic
Semantic
Number of phonemes
Number of syllables
Number of phonological neighbors
Phonologic Levenshtein distance 20
Number of phonographic neighbors
Number of letters
Orthographic Levenshtein distance 20
Number of orthographic neighbors
Bigram letter average frequency
Number of morphemes
Number of senses and meanings across parts of speech (Wordnet)
Semantic diversity of neighbors (Hoffman)

eIRT: meanings & senses*group



Converging results across analyses

- **Contextual diversity (subtitles): number of documents in the corpus containing a given word (corpus of subtitles from movies and television)**
  - Single strongest predictor of item easiness for all groups.
  - Positive relation: greater contextual diversity is associated with easier items
  - Correlated with some other predictors (e.g., frequency), often the only significant unique predictor in models with multiple predictors
  - eIRT: significant interaction of contextual diversity*group

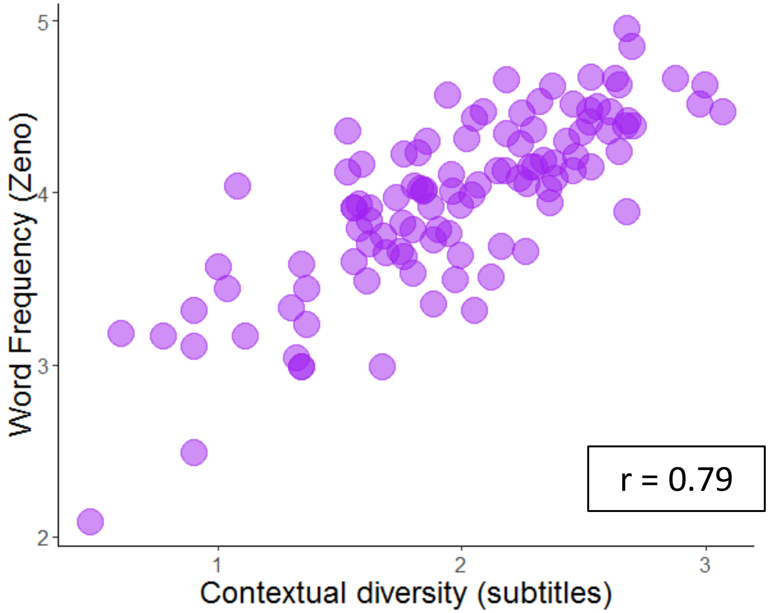
- **Number of meanings and senses (WordNet): combined meanings and senses across parts of speech**
  - Greater number of meanings and senses associated with easier items
  - eIRT: significant interaction of meanings&senses*group

- **Word frequency:** higher frequency associated with easier items

Discussion

- Frequency vs. Contextual diversity: where we encounter a word vs. how often
  - Adelman, Brown, & Quesada, 2006
- More meanings & senses: easier items
- How many predictors are enough?

- IFEP & RFEP students: show similar pattern to EO students
- LEP: weaker relations between predictors & item easiness
- Heterogeneity among ELLs



Still exploring…

- Part of speech
- Word concreteness
- Differential distractor functioning
- Carrier sentence
- Characteristics of Key
- Similarity between Key & Target
- Similarities between responses
- Student characteristics

XX. We had <u>sufficient</u> food at the party.

a. delicious
b. too much
c. standard
d. enough

Thank you!

http://172.27.244.67//sample-apps/alm/

Autumn McIlraith

autumn.mcilraith@times.uh.edu