1. Phonological bootstrapping

The hypothesis is that phonetics provide reliable cues to group words into coarse lexical categories, such as nouns and verbs. Thus, the sound of words is not entirely arbitrary but there are regularities that provide category level information to bootstrap the acquisition of syntax.

We used Naive Discriminative Learning to run computational simulations of lexical category acquisition based on phonological cues, evaluating how effective category learning was.

2. Discriminative learning

Naive Discriminative Learning (NDL) implements a very simple form of error driven learning through a flat neural network that learns cue-outcome associations using the Rescorla-Wagner equations:

\[
\Delta V_{ij} = \begin{cases} 
\alpha \beta_i (\lambda - \sum_{c \in t} V_c) & \text{if } c_i \in t \text{ and } o_j \in t \\
\alpha \beta_i (0 - \sum_{c \in t} V_c) & \text{if } c_i \in t \text{ and } o_j \notin t \\
0 & \text{if } c_i \notin t 
\end{cases}
\]

Cue-outcome associations are:
- **strengthened** if cue and outcome are present
- **weakened** if the cue is present and the outcome isn’t
- **left unchanged** when the cue is not present

The change in association depends on the number of cues in a learning trial and the amount of error in predicting each outcome given the present cues. Associations are updated independently for every outcome.

3. Experimental design & results

A grid search over parameters was carried out, using **four different test sets**, obtained crossing ambiguousness (ambiguous, unambiguous) and novelty (known, new), resulting in four test sets. The other parameters that were investigated are:

- **Training regime:** words or utterances
- **Vowels:** full or reduced (unstressed) vowels
- **Activation values:** sum or frequencies
- **K top active words:** 20, 50, 100, 200

4. Conclusions & future work

Under certain conditions, tagging words based on localist phonological features using the NDL model is feasible. The grid search shows that:
- **unambiguous** words are easier
- **known** words are easier
- **Training on single words** is easier than on full utterances
- **Syllables** perform better, but may not scale to non-words
- **Summing activation values** is better
- **Comparing to baseline** only helps when training on utterances and considering frequencies
- **reduced/full vowels and stress/no stress** don’t make a difference

NEXT: Correlate model’s decisions and confidence with behavioral results from children and adults to see whether the model is sensitive to the same information and uses it like humans do.

Acknowledgements

This research was supported by a BOF/TOP grant (ID 29072) of the Research Council of the University of Antwerp. Big thanks to Stéphan Tulkens for the help on implementing the NDL model and to Marco Marelli for the discussions about discriminative learning and the useful advice on the experimental design of this work. The poster was designed on Overleaf using the baposter template by Brian Amberg.