Distributional characteristics of accurate second language English inflectional morpheme use

What this research was about and why it is important

Usage-based theories hold that (second) language acquisition is influenced by the distributional characteristics of learners' input including such factors as frequency and formulaicity. Our study empirically tested this claim by examining the accuracy of four inflectional morphemes (past tense *-ed*, progressive *-ing*, third person *-s*, plural *-s*) in a large-scale corpus of learner writing. The study demonstrated that second language learners use the morphemes more accurately in reliable forms, i.e., when the frequency of the inflected word form (e.g., *arrived*) relative to the frequency of the lemma of the form (e.g., *arrive*) is high. Contrary to a previous study that investigated the issue experimentally in laboratory online processing tasks, however, our study did not consistently identify the effects of the frequency of the inflected form itself or the formulaicity of the words surrounding the inflected form on morpheme accuracy.

What the researchers did

- We drew data from a large-scale corpus of learner writing (EF-Cambridge Open Language Database [EFCAMDAT]).
- Based on the coding of learners' errors (available in the corpus), we identified both the errors and accurate uses of each target morpheme in each sample of writing.
- The information about how these morphemes and lemmas are typically distributed in the language as a whole (e.g., frequency of the inflected form, formulaicity of the context in which the inflection occurred) was calculated based on a large-scale corpus of American English (Corpus of Contemporary American English).

What the researchers found

- We found a consistent positive association between reliability (i.e., the frequency of the inflected form relative to the frequency of the lemma) and morpheme accuracy.
- This suggests that the lemma functions as a cue for its inflected forms and that learners use (tally) such an association in processing inflectional morphemes when they read or hear input.
- On the other hand, the study did not find a consistent pattern between the raw frequency of inflected forms and morpheme accuracy or between the formulaicity of the context and morpheme accuracy.
- The association between the distributional characteristics and morpheme accuracy did not differ much across learners' proficiency levels or their longitudinal development.

Things to consider

- Both our study and the previous experimental study on the same topic found a consistent effect of reliability, indicating the robust effect of lemma-morpheme association strength on morpheme accuracy.
- This robust effect can be interpreted as evidence for the phenomenon known as associative learning, a theory of learning that can explain (at least some parts of) second language acquisition.
- Contrary to the previous experimental study, our study did not find the effect of the raw frequency of the inflected form or formulaicity on morpheme accuracy.
- This difference is possibly due to methodological differences between the experimental task (elicited imitation, demanding online processing of predetermined stimuli) and the more conscious, written compositions in EFCAMDAT analyzed for the current study, in which learners decide which language they need, and so the language may be less prone to effects of formulaicity or the frequency of inflected forms.