Modeling morphological production with an algorithmically specified InflACT-R

Jane Li (sli213@jhu.edu) Grusha Prasad (gprasad@colgate.edu)
Johns Hopkins University *Colgate University*

Introduction: Prior speeded inflection studies on the English past tense found that regular inflections (*love→loved*) are produced more rapidly than irregulars (*think→thought*). Additionally, while frequent irregular inflections are produced more rapidly than their infrequent counterparts (e.g., *thought* is produced faster than *kept*), such a frequency effect is absent for regular inflections[1-2]. The dual-route theory[3] accounts for this by postulating that irregular inflected forms are stored in the lexicon, and thus subject to frequency effects that impact retrieval[4]; regular inflections are not retrieved from the lexicon but instead computed on the fly. This account describes the *computations* involved in inflection but does not specify details to understand verb inflection at a Marr's *algorithmic* level. For example, the theory assumes that the on-the-fly rule computation is suppressed for irregular inflections, but does not specify *why* rules are suppressed for rule-governed irregular inflections (*ring→rang*) or *how* the processor suppresses these rules. The theory also assumes that some regular inflected forms can be retrieved from memory if it is faster than the on-the-fly computation[5] but does not specify *which* regular inflections are stored in memory and *why*. In this work we use the ACT-R[6] framework to specify the components of an algorithmic level model of past tense inflection (InflACT-R). Unlike previous ACT-R models that focus on the *learning* of irregular morphology[7], InflACT-R focuses on online processing. As a starting point, we focus on one of the components of the model and run simulations to illustrate how this component can account for some but not all of the existing empirical evidence. We then use this gap to motivate and discuss the other components.

Model: InflACT-R has two components: declarative memory (stores stems and morphological operations) and procedural memory (specifies the steps for producing the inflected form). When presented with a stem **s**, the inflector retrieves the operation **o** with the highest activation from the set of possible transformations that can apply to **s**. Then, it applies **o** to produce the inflected form. We propose that the activation of an operation given a stem is influenced by several factors [Eqn1]. We focus on the base-level activation of transformations because it is the factor most directly influenced by frequency of inflected forms: a transformation with high base-level activation will be retrieved more quickly than one with lower activation [Eqn2].

Method: The operations we included in our simulations ranged from ones that correspond to the regular rule ([d]-suffixation, as in *love*-*love*[d]), minor rules (ɪ → æ, *ring*-*rang*), and strong suppletion (*go*-*went*). We generated a dataset with 35 operations obtained from [8]'s implementation of the Minimal Generalization Learner[9] and 3308 stems (with frequency information) and used this to compute base-level activations. For an operation ***o***, we estimate B_O by sampling inflected forms weighted by their frequencies (n = 100,000) and track the instances where *o* is the operation applied (Fig. 1).

Results [Tab1]: The base-level activations for regular operations are much higher than irregular ones, thus accounting for the result that regular inflections are produced more rapidly than irregular ones. Additionally, the high and low frequency inflections differ in orders of magnitudes relative to the regular case, thus accounting for the observed frequency effects for irregular but not regular inflections. Base-level activations by themselves are insufficient, however: the base-level activations of the regular competitors for any irregular stem is too high for the correct irregular operation to overcome. Thus, in order to account for all the empirical patterns, an inhibition term that decreases the activation of the incorrect regular term is necessary. Adding an inhibition term is conceptually related to the dual-route theory's proposal that rules are suppressed for irregular inflections. However, by implementing this within the ACT-R framework, there are fewer degrees of freedom on *why* and *how* a rule is suppressed. More broadly, this work suggests that it might be possible to account for the computational level differences that dual-route proposes between regular and irregular inflections without necessarily assuming that these differences emerge from separate algorithmic level processes.

References. **[1]** Prasada, S., Pinker, S., & Snyder, W. (1990). *Bulletin of the Psychonomic Society*. **[2]** Beck, M. L. (1997). *Second Language Research*, 93-115. **[3]** Ullman, M. T. (2001). *Nat. Rev. Neurosci.* **[4]** Jescheniak, J. D., & Levelt, W. J. (1994). *JEP: LMC*. **[5]** Pinker, S., & Ullman, M. T. (2002). *TiCS*. **[6]** Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). *Psychological review.* **[7]** Taatgen, N. A., & Anderson, J. R. (2002). *Cognition*. **[8]** Wilson, C., & Li, J. (2021). *SIGMORPHON*. **[9]** Albright, A., & Hayes, B. (2002). *ACL*. **[10]** Vasishth, S., & Engelmann, F., (2021). *Sentence comprehension as a cognitive process*: *A computational approach.* **[11]** Cohen-Shikora, E. R., Balota, D. A., Kapuria, A., & Yap, M. J. (2013). *BRM*.

Equation 1[adapted from 10]. Activation of an operation *o*: *s* is the stem under consideration, *B* for base-level activation, *L* for lexical activation (not modelled in this abstract), and *I* for inhibition.

$$A_{so} = B_o + L_{so} + I_{so} + \varepsilon$$

Equation 2[10]. Relating activation to retrieval time: *F* is participant-specific latency factor, *f* is a latency exponent, and *A* is the activation of Equation 1.

$$RT = Fe^{-f(A_{so})}$$

Stem	Frequency	Regular operation	Base activation (retrieval time; ms)	Irregular operation	Base activation (retrieval time;ms)
love	High	∅ → d	5.159 (0.0013)	∅ → ∅	4.106 (0.0038)
age	Medium	∅ → d	5.159 (0.0013)	∅ → ∅	4.106 (0.0038)
implement	Low	∅ → ɨd	4.821 (0.0018)	∅ → ∅	4.106 (0.0038)
think	High	∅ → t	4.235 (0.0033)	ɪŋk → ɔt	1.532 (0.0497)
sing	Medium	∅ → d	5.159 (0.0013)	ɪ → ɑ	0.472 (0.1435)
wring	Low	∅ → d	5.159 (0.0013)	ɪ → ʌ	0.656 (0.1193)

Table 1. Simulation results of a few example stems. Green cells are the correct inflection. Based on [10] time calculated by setting f to 1 and F to 0.32 (median of Beta(2,6) used by [10]).

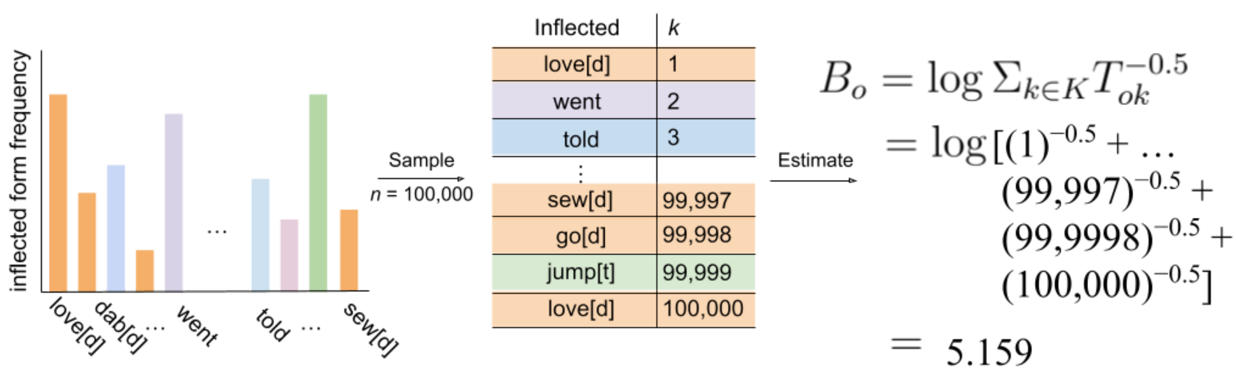

Figure 1. Estimation procedure for the base-level activation of the [d]-suffixation rule, generalizable to any operation. Sampling (with replacement) is done over the weighted frequency of inflected forms. Some samples are distorted with errors (e.g., *go-ed* in the table), error proportions are estimated by error rates from the Past Tense Inflection Project[11].