Amanda Klassen COGS200

Article Summary of Rumelhart and McClelland (1986) "On Learning the Past Tenses of English Verbs

In their paper, Rumelhart and McLelland explore the nature of rules in language and the cognitive processes from which it arises. Diverging from the traditional view that the grammar is explicitly and inaccessibly represented, they develop a model in which these patterns are only the result of the system which produces them. While rules still play a crucial descriptive role within linguistics, the cognitive mechanism itself does not expressly represent them. An interactive activation model of letter and word units had already demonstrated a simple version of this idea, successfully replicating aspects of the word superiority effect. To prove that this can account for language learning, a parallel distributed processing (PDP) model was used to show behaviours analogous to that seen in the acquisition of the past tense.

To succeed, the model had to mirror three stages of past tense usage. First, only a few verbs are used in past tense, many of them irregular, all represented as discrete items. Next, rule-based behaviour emerges, the past tense ending *-ed* not only applied to a much larger number of regular verbs, but also overgeneralised to both present and past irregular forms; children can also correctly append it to a nonsense word. Finally, both regular and irregular forms are used properly, although regularisation of low-frequency weak verbs continues into adulthood. Transition between stages is gradual. The PDP model was ideal for this, as distributed representation allows past patterns to strengthen new ones, emulating rule-learning, and can accommodate regularities and irregularities simultaneously, but not until the former are sufficiently strengthened; this is in line with data from the second stage of past tense learning.

The PDP model consisted of two parts: a simple pattern associator, which created weighted connection strengths between unit activation-patterned representations of present and past tense forms; and input and output systems for conversion between strings of phonological representations and the context-sensitive featural ones used by the pattern associator. Reflecting the real-world conditions, the model received only positive input, root forms shown alongside their past tense. The model generated internal predictions and altered connection strengths after errors, strengthening inputs and lowering activation thresholds for output units showing misses, doing the opposite for false alarms. A corpus-generated set of the 506 most frequent verbs were presented in three stages, the first consisting of the 10 most common, the second these plus 410 medium-frequency (roughly parallel with the word burst seen at 18 months), a third using free-generation response to 86 low-frequency verbs to test responses to novel inputs.

After initial learning trials, the model performed equally well on all verbs. In the second phase, regular forms had better-than-chance featural output on first presentation, but performance dropped sharply for irregular forms, regularised forms outnumbering correct ones. This was corrected gradually, more quickly for high-frequency verbs, but these items never again matched the accuracy for regular forms. Results varied between irregular verb classes, worse for past tenses with the most phonological difference, such as vowel change, best for verbs which end in *-t/-d* and require no change. This created a second generalisation, where regular verbs with *-t/-d* endings incorrectly resulted in a no-change output. Outputs during the third stage generated a high proportion of correct responses, particularly for regular forms.

PDP model simulations successfully demonstrated the same three stages as children, including slow transition from the second to third. Differences in performance between verb classes, such as no-change irregulars and the second generalisation they cause, are also observed in real-world studies. These, along with the ability of the model to make correct predictions for novel verbs, proves that formal rules can be instantiated by a system with no representation of them.