Proceedings of the Society for Computation in Linguistics

Volume 2 Article 54

2019

Transient blend states and discrete agreement-driven errors in sentence production

Matthew Goldrick
Northwestern University, matt-goldrick@northwestern.edu

Laurel Brehm
Max Planck Institute for Psycholinguistics, laurel.brehm@gmail.com

Pyeong Whan Cho
University of Michigan, pyeongwhan.cho@gmail.com

Paul Smolensky
Johns Hopkins University; Microsoft Research, smolensky@jhu.edu

Transient blend states and discrete agreement-driven errors in sentence production

Errors in subject-verb agreement are common in everyday language production. This has been studied using a preamble completion task ([1]) in which a participant hears or reads a preamble containing inflected nouns and forms a complete English sentence ("The key to the cabinets" could be completed as → *The key to the cabinets is gold.*) Existing work has focused on errors arising in selecting the correct verb form for production in the presence of a more 'local' noun with different number features (*The key to the cabinets <u>are</u> gold*; [3]-[4]). However, the same paradigm elicits substantial numbers of *preamble* errors (*The key to the <u>cabinets</u>* → *The key to the <u>cabinet</u>*; [1]) that existing theories have largely failed to address.

We propose a Gradient Symbolic Computation (GSC; [2]) account of agreement and preamble errors. Sentence processing is modeled as a continuous-time, continuous-state stochastic dynamical system (as in [4]). Within this continuous representational space, a subset of states reflect discrete symbolic structures. The remainder are *blend* states where multiple symbols are simultaneously partially active. Initial phases of computation prefer blend states; an additional dynamic control parameter, *commitment strength*, pushes the model to discrete structures. This process, combined with stochastic gradient ascent dynamics respecting grammatical constraints on syntactic structures, yields discrete sentence outputs. We propose that transient blend states allow portions of target and non-target syntactic structures to interact, yielding both verb and preamble errors.

Model of sentence generation

A GSC model implemented a probabilistic context-free grammar G (1), limiting the sentence-length to 4. Preverbal noun phrases such as N_s N_p correspond to a noun phrase with a singular head and a plural 'local' prepositional phrase complement (*The key to the cabinets*). Probabilities for noun number followed English biases for singular. Note singular nouns can take plural agreement, as in pseudo-partitives or collectives (*A number of problems <u>are</u>*; [4]).

In this model, agreement errors may arise when there is partial activation of a structure where the verb agrees with the local noun. We included one grammatical parse with this structure, $N_{s/p}$ RC: a noun phrase with a relative clause complement N_iV_i (*The key <u>the cabinets</u> <u>were</u> locked by*). As this stands in for a range of structures where the local noun and the immediately following verb agree in number, we assigned it a relatively high probability (0.4).

Representational similarity was introduced among the vectors encoding symbols in the grammar: (a) Assuming the symbols differing only in number are similar, the dot product of pairs of vectors encoding singular vs. plural (e.g. NP_p and NP_s) was set (arbitrarily) to 0.1. (b) Following [2], rules with non-terminals that have multiple expansions (e.g. $NP_s \rightarrow N_sN_s \mid N_sN_p$) are encoded by multiple distinct non-terminal representations (e.g. NP_s and NP'_s). The similarity of these representations was set to an (arbitrarily) higher value, 0.5. (c) The dot product of all other pairs was set to 0. From these constraints, 29 *filler* vectors were randomly chosen as vector encodings of the terminal and non-terminal symbols. Ten orthonormal *role* vectors were randomly chosen as vector encodings of the structural position of the symbols. These filler and role vectors were composed by the outer product [5] to generate 290 binding vectors, e.g. $S \otimes r_{root}$.

(1) *A probabilistic context-free grammar G yielding 3 sentence types:* $[N_iV_i]$*;* $[[N_iN_j]V_i]$*;* $[N_i[N_jV_j]V_i]$. *Note: subscripts denote grammatical number of the associated symbol.*

- $S \rightarrow 0.22\ N_s\ V_s \mid 0.11\ N_s\ V_p \mid 0.167\ N_p\ V_p \mid 0.22\ NP_s\ V_s \mid 0.11\ NP_s\ V_p \mid 0.167\ NP_p\ V_p$
- $NP_s \rightarrow 0.4\ N_s\ N_s \mid 0.2\ N_s\ N_p \mid 0.4\ N_s\ RC$ $\qquad NP_p \rightarrow 0.4\ N_p\ N_s \mid 0.2\ N_p\ N_p \mid 0.4\ N_p\ RC$
- $RC \rightarrow 0.44\ N_s\ V_s \mid 0.22\ N_s\ V_p \mid 0.33\ N_p\ V_p$

Proceedings of the Society for Computation in Linguistics (SCiL) 2019, pages 375-376.
New York City, New York, January 3-6, 2019

Sentence generation was modeled by initializing the system to a random point near the equilibrium state of the system at commitment strength 0. Commitment strength was then increased, pushing the system to select a discrete structure. Prior to testing, the grammar (1) was implemented by initializing the model as in [2] and then updating the grammatical constraint parameters to minimize the difference between model and target output probabilities. Post-training, the model generated grammatical structures on more than 80% of trials; on these trials, the model approximated the grammar's probability distribution over full parse trees.

Results: Simulation of preamble task

The preamble was encoded by weak external input which decayed over time. Sentence generation then followed the normal procedure above.

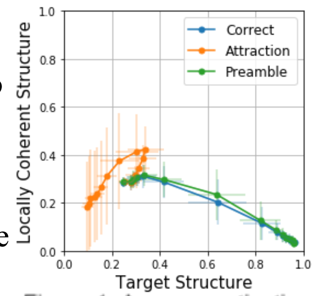

Figure 1. Average activation (separated by outcome) of target (local N as PP complement) vs competitor (local N as head of RC controlling V) non-terminal nodes

Attraction errors. Replicating previous studies ([1],[3],[4]), preambles with a singular head and plural local noun were more likely to yield verb completion errors (*the key to the cabinets are*: 17%) than control preambles with only singular nouns (*the key to the cabinet are*: 8%) or preambles with a plural head and singular local noun (*the keys to the cabinet is*: 1%). As shown in Figure 1, these errors appeared to derive from a transient blend representation in which there is partial activation of the target (X axis) as well as a locally-coherent structure in which the local noun controls the number of the verb (Y axis). To test this hypothesis, we trained a new model in which the frequency of parses containing the locally-coherent structure was decreased by 50%, reducing its presence in blend states. Consistent with our hypothesis, the rate of attraction errors dramatically decreased (17% → 7%).

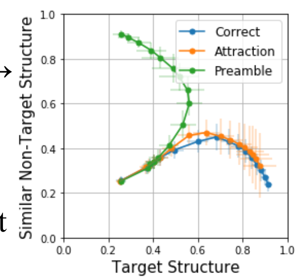

Figure 2. Average activation of target (plural PP) vs. competitor (singular PP) non-terminal nodes

Preamble errors. Errors on the local noun (*the key to the cabinets* → *the key to the cabinet*) occurred at a significant rate (9% for head singular, local plural). As shown in Figure 2, these errors derive from a blend state containing the target structure ($NP_s \rightarrow N_s\ N_p$; X axis) as well as a highly similar noun phrase containing a singular prepositional phrase complement (i.e., a preamble error; $NP_s \rightarrow N_s\ N_s$; Y axis). To test this hypothesis, we trained a new model in which the vectors encoding the two noun phrases were not similar (orthogonal). Consistent with our hypothesis, the rate of preamble errors dramatically decreased (9% → 1%).

Conclusions

In the GSC model, agreement and preamble errors arise due to blend representations. Attraction errors reflect interactions between the target and a portion of non-target parses, while preamble errors reflect interactions between highly similar non-terminal nodes. The performance of the model with closer approximations of English parse probabilities will be examined.

References [1] Bock, K., & Miller, C. A. (1991). Broken agreement. *Cognitive Psychology*, *23*, 45-93. [2] Cho, P. W., Goldrick, M., Lewis, R., & Smolensky, P. (2018). Dynamic encoding of uncertainty in gradient symbols. In *CMCL 2018*, 19–28. [3] Eberhard, K. M., Cutting, J. C., & Bock, K. (2005). Making syntax of sense: number agreement in sentence production. *Psychological Review*, *112*, 531-559. [4] Smith, G., Franck, J., & Tabor, W. (2018). A self-organizing approach to subject–verb number agreement. *Cognitive Science*. [5] Smolensky, P. (1990) Tensor product variable binding and the representation of symbolic structures in connectionist networks. *Artificial Intelligence, 46,* 159–216.