Neural network and human cognition: A case study of grammatical gender in Swedish

Ali Basirat[1], **Marc Tang**[1]

[1]*Uppsala University, Department of Linguistics and Philology*

Presenter's e-mail address: ali.basirat@lingfil.uu.se ; marc.tang@lingfil.uu.se

We target grammatical gender in Swedish and compare the errors in grammatical gender recognition produced by the artificial neural network model versus errors of L1 and L2 Swedish speakers attested in the existing literature. This research thus follows the call of SweCog 2017 "creating a strong interdisciplinary cluster of cognitive science oriented research within Sweden". First, we combine cognitive science with computational linguistics and language acquisition. Second, our analysis is based on the Swedish language. Third, as mentioned by the call of SweCog: "describing on-going work are especially welcome". Our study is also an on-going project which will benefit from the feedback at the conference.

"Language offers a window into cognitive function, providing insights into the nature, structure and organisation of thoughts and ideas." (Evans & Green, 2009, p. 5). Thus, linguists are interested in systems of nominal classification, i.e. how languages classify nouns of the lexicon, due to their diverse lexical and pragmatic functions as well as cognitive and cultural correlates. As stated by Lakoff & Johnson (2003, pp. 162–163), "In order to understand the world and function in it, we have to categorize, in ways that make sense to us, the things and experiences that we encounter". This need is fulfilled in language via various means of categorization, one of the most prominent being grammatical gender (Aikhenvald, 2000; Corbett, 1991). In gender system (also known as noun class system) of languages, all nouns of the lexicon are assigned to a specific number of classes. Saying that a language has two genders implies that there are two classes of nouns which can be distinguished syntactically by the agreement they take with other syntactic units in their contexts (Bohnacker, 2004, p. 198; Senft, 2000). As demonstrated in Fig. 1, nouns in Swedish are divided into neuter and uter (common). The two categories are thus reflected on the determiners and adjectives respectively.

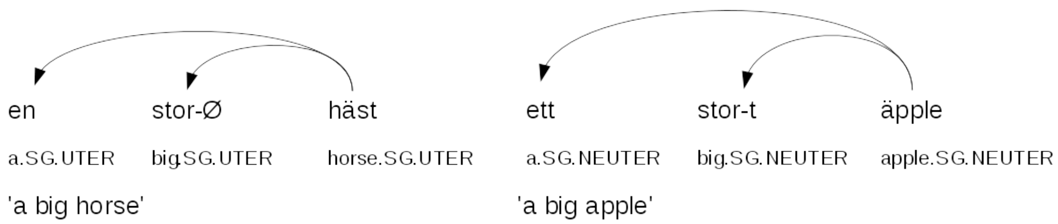

Figure 1: The grammatical gender agreement in Swedish

We selected Swedish due to the following reasons. Gender assignment across languages is underlined by cognitive and sociocultural principles (Aikhenvald, 2012; Corbett, 1991, p. 57; Kemmerer, 2017). By way of illustration, long-shaped objects tend to be affiliated to masculine grammatical gender and round-shaped objects are generally associated with feminine grammatical gender. However, contradictory observations are made in Swedish. First, grammatical gender assignment is generally viewed as arbitrary (Andersson, 1992; Teleman, Hellberg, & Andersson, 1999), but some semantic regularities are still attested in Swedish. Animate nouns, especially "all non-pejorative, classificatory nouns denoting adult human beings, a qualified majority of all other human nouns and a majority of all other animate nouns" strongly tend to be affiliated to the uter gender (Dahl, 2000, pp. 586–587), while other inanimate nouns are affiliated to the neuter gender. This distribution is also reflected in terms of count/mass division. Nouns referring to concrete and countable entities are more likely to be uter while abstract or collective meanings are associated to neuter, e.g. "possible people containers" such as nouns denoting location and organization are perceived as collective units. Thus, they tend to be neuter (Fraurud, 2000, pp. 191–203). Second, the L1 and L2 acquisition of Swedish grammatical gender are controversial and differ significantly from other languages. Monolingual children acquire the Swedish gender system with nearly no errors (Andersson, 1992; Bohnacker, 1997; Plunkett & Strömqvist, 1990), which is considered rare in comparison to other gender languages, for which "children's acquisitional paths have been reported not to be quite so error-free" (Bohnacker, 2004, pp. 214–217). Moreover, gender assignment on Swedish nouns via their phonological form or semantics is generally considered as unpredictable (Andersson, 1992; Teleman et al., 1999), which makes this observation even more unexpected. Nevertheless, while L1 acquisition display a lack of errors, L2 (child) learners do encounter difficulties, suggesting that different strategies are employed (Bohnacker, 2004, p. 218).

We chose a computational approach since the use of computer simulations of neurons and neural networks are one of the most important tools in computational cognitive neuroscience (Parks, Levine, & Long, 1998). We simulate the learning process of the brain with neural networks, which “have become a subject of intense interest to scientists spanning a broad range of disciplines including psychology, physics, mathematics, computer science, biology and neurobiology” (Gopal, 1996, p. 69). In this study, we apply the feed-forward neural network (Haykin, 1998) to classify between the Swedish grammatical genders of nouns. The architecture of the network is as follows: We first feed a 50-dimensional vector representation of words, called word vector, in the input layer of the network. The word vector is then processed in the only hidden layer of the network consisting of 100 neurons. In the output layer, the network generates two weight values corresponding to the grammatical gender of the input word. The input word vectors are extracted by the RSV (Real-valued Syntactic Word Vectors) word embedding model (Basirat & Nivre, 2017). Our model relies on two main sources of data, which both originate from the Swedish Language Bank (Språkbanken) located at the University of Gothenburg: a corpus of Swedish raw sentences and a list of nouns affiliated to grammatical genders. The corpus originates from Swedish Wikipedia available at Wikipedia Monolingual Corpora, Swedish web news corpora (2001-2013) and Swedish Wikipedia corpus collected by Språkbanken. The list of nouns and their affiliated grammatical gender was extracted from the SALDO (Swedish Associative Thesaurus version 2) dictionary. The dictionary, consisting of 21,670 words with uter and neuter genders, is partitioned in a standard way into three parts with no overlap, so that the results can be generalized to the entire lexicon of the language. We use 80% of words (17,338) to train the neural network, 10% of words (2,166) as the development set, and the remaining 10% as test set. The dictionary data is divided into three sets so that the performance of the neural network may be enhanced and re-measured between the development test and the test. All words are randomly selected in their base format with no morphological inflection and all three training sets contain an equivalent distribution of uter and neuter nouns.

Our preliminary results on the development set show that the neural network can correctly detect the grammatical gender of the nouns with the overall accuracy of 93.46%. In other words, when being presented a new word, the neural network can interpret correctly the grammatical gender of the noun in 93.46% of the time. It is worth mentioning that the only source of information used by the neural network is the contextual information encoded into the word vectors with no information about the possible morphological inflections of words. We have also observed that the precision for the uter nouns (97.1%) is significantly higher than the precision for the neuter nouns (84.6%), i.e. most of the difficulties encountered by the neural network are related to the fact that it could not identify the gender of neuter nouns. This correlates with the findings of human acquisition where learners tend to overgeneralise the uter gender (Bohnacker, 2004, p. 218) due to the lack of balance in terms of quantity between uter and neuter nouns in Swedish (71.06% vs 28.94% according to SALDO). As shown in Fig. 2, the errors generated by neural network are mostly cases where the noun is located in the space of the opposite gender, e.g. a neuter noun misinterpreted as uter nouns (red) is more likely to be found in the cloud of neuter nouns (green).

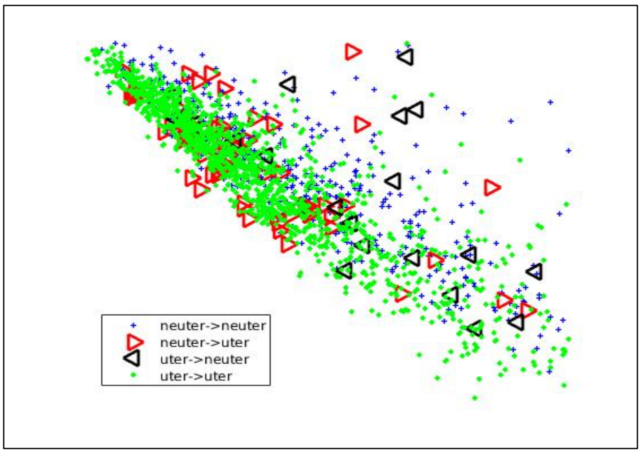

Figure 2: PCA representation of the word vectors classified by the neural network with respect to their grammatical genders. “X→Y” means the noun belonging to category X is classified as Y.

We will provide in our presentation a preliminary analysis, which further demonstrate that the distribution of errors made by both human and the artificial neural network is similar. We observe that both humans and the artificial neural network have difficulties interpreting the gender of nouns in cases of syntactic ambiguity and polysemy, e.g. participles such as *flyttande* 'moving' are attested as nouns in SALDO but they may also appear as participles in modifier structures. The same case occurs between verbs and nouns, e.g. *delta* can link to a verb 'participate' or a noun 'delta'. Moreover, nouns such as *kaffe* 'coffee' may be interpreted as uter or neuter depending on whether it is referring to the entity as count or mass. Thus, when two entries are competing for the same word vector, the vector is biased toward the most frequent reference which will negatively affect the decision making of the neural network. Finally, the actual limitations and future prospects of this study include an expansion of first-hand data from Swedish L1 and L2 speakers, i.e. we intend to gather data specific to our study via empirical experiments with human speakers. Furthermore, a cross-language analysis is also required to verify the stability of the neural network in various linguistic environments.

References

Aikhenvald, A. Y. (2000). *Classifiers: A Typology of Noun Categorization Devices*. Oxford: Oxford University Press.

Aikhenvald, A. Y. (2012). Round women and long men: Shape, size, and the meanings of gender in New Guinea and beyond. *Anthropological Linguistics*, *54*(1), 33–86.

Andersson, A.-B. (1992). *Second language learners' acquisition of grammatical gender in Swedish* (PhD dissertation). University of Gothenburg, Gothenburg.

Basirat, A., & Nivre, J. (2017). Real-valued Syntactic Word Vectors (RSV) for Greedy Neural Dependency Parsing. In *Proceedings of the 21st Nordic Conference on Computational Linguistics, NoDaLiDa* (pp. 21–28). Gothenburg: Linköping University Electronic Press.

Bohnacker, U. (1997). Determiner phrases and the debate on functional categories in early child language. *Language Acquisition*, *6*, 49–90.

Bohnacker, U. (2004). Nominal phrases. In G. Josefsson, C. Platzack, & G. Håkansson (Eds.), *The acquisition of Swedish grammar* (pp. 195–260). Amsterdam: John Benjamins.

Corbett, G. G. (1991). *Gender*. Cambridge: Cambridge University Press.

Dahl, O. (2000). Elementary gender distinctions. In B. Unterbeck & M. Rissanen (Eds.), *Gender in grammar and cognition* (pp. 577–593). Berlin: Mouton de Gruyter.

Evans, V., & Green, M. (2009). *Cognitive linguistics: an introduction* (Repr). Edinburgh: Edinburgh University Press.

Fraurud, K. (2000). Proper names and gender in Swedish. In B. Unterbeck, M. Rissanen, T. Nevalainen, & M. Saari (Eds.), *Gender in Grammar and Cognition* (pp. 167–220). Berlin: Mouton de Gruyter.

Gopal, S. (1996). Neural network models of cognitive maps. In J. Portugali (Ed.), *The Construction of Cognitive Maps* (pp. 69–85). Dordrecht: Kluwer Academic Publishers.

Haykin, S. (1998). *Neural networks: A comprehensive foundation*. New Jersey: Prentice Hall PTR.

Kemmerer, D. (2017). Categories of object concepts across languages and brains: the relevance of nominal classification systems to cognitive neuroscience. *Language, Cognition and Neuroscience*, *32*(4), 401–424.

Lakoff, G., & Johnson, M. (2003). *Metaphors we live by*. London: The University of Chicago Press.

Parks, R. W., Levine, D. S., & Long, D. L. (Eds.). (1998). *Fundamentals of neural network modeling: neuropsychology and cognitive neuroscience*. Cambridge, Mass: MIT Press.

Plunkett, K., & Strömqvist, S. (1990). *The acquisition of Scandinavian languages*. Gothenburg: University of Gothenburg.

Senft, G. (2000). *Systems of nominal classification*. Cambridge: Cambridge University Press.

Teleman, U., Hellberg, S., & Andersson, E. (1999). *Svenska Akademiens grammatik. Vol. 2: Ord. [The Swedish Academy Grammar, Part 2: Words]*. Stockholm: Norstedts.