# Title.
# Effects of lexical neighbourhood density and phonotactic probability studied with a new database of matched pairs of real signs and modelled pseudosigns in the Swedish Sign Language

Erik Witte[1], Krister Schönström[2], Thomas Björkstrand[2], Henrik Danielsson[3] & Emil Holmer[4]
[1]Faculty of Medicine and Health, Örebro University and Department of Behavioural Sciences and Learning (IBL), Linköping University, erik.witte@oru.se, [2] Department of Linguistics, Stockholm University, [3] IBL, Linköping University, [4] IBL, Linköping University, emil.holmer@liu.se

**Keywords:** Sign language, Lexical database, Pseudosign generation, Language processing, Lexical access, Neighbourhood density, Phonotactic probability, Lexical decision task

The combined effects that the presence of competing lexical neighbours and the typicality of phonotactic patterns have on the accuracy and speed of which the human brain accesses items in its mental lexicon have been studied for spoken languages for well over twenty years. Typically, for items in dense phonological neighbourhoods lexical access is slower and less accurate than for items in sparse neighbourhoods, and items with highly probable phonotactic patterns are easier and faster to perceive. Phonotactic probability (PP) and neighbourhood density (ND) thus forms two fundamental processes in human word recognition (Vitevitch & Luce, 2016). In comparison, sign language phonotactics allows for more ways in which phonological features may be combined – both simultaneously and sequentially – than what is possible in spoken language. Consequently, the roles that ND and PP play in lexical access and related processes, such as learning new lexical items, may partly differ between spoken and signed languages. Only a few studies have investigated such effects in sign languages (Cf. Williams, Stone & Newman, 2017; Caselli, Emmorey & Cohen-Goldberg, 2021).

To investigate the effects of PP and ND on lexical access, we created a database of signs and pseudosigns for use in a lexical decision task (LDT). Real signs were drawn from the Swedish Sign Language Dictionary (https://teckensprakslexikon.su.se), which contains over twenty thousand phonologically transcribed signs in the Swedish Sign Language (Svenskt Teckenspråk; STS). Based on these transcriptions and frequency data from the STS-korpus (Öqvist, Riemer Kankkonen & Mesch, 2020), metrics of neighbourhood density (ISD20, mean distance to the 20 closest neighbours) and phonotactic probability (CPP, the probability of the combination of different sign features in a given sign) were devised. Pseudosigns were created by weighted sampling from sign features based on such probabilities. 400 real signs and 400 pseudosigns were closely matched into pairs with similar values of ISD20 and CPP (see figure 1), video recorded, and then presented in an LDT.

Accuracy results were analysed with logistic regression, modelling the probability of responding “real sign”, given ISD20, CPP, and their respective interactions with sign type (see figure 2). Preliminary results from a single fluent L2 signer indicate that, a) the accuracy in identifying real signs is independent of ND. Higher ND, however, increases the risk of mistaking a pseudosign for a real sign. b) High PP increases the chance of correctly identifying real signs. The accuracy in identifying pseudosigns, however, seems to be independent of PP. In the final presentation, both response time and accuracy analyses will be included, based on LDT results from L1 signers of STS.

Our preliminary results indicate that PP shows a clear facilitatory effect upon lexical access. However, in contrast to the results of Caselli et al. (2021), we see no evidence of lexical competition among real signs. This suggests that sign languages suffer less from lexical competition, and possibly that mental representations of signs are more distinct than for words. This ongoing study will deepen the understanding in this area.

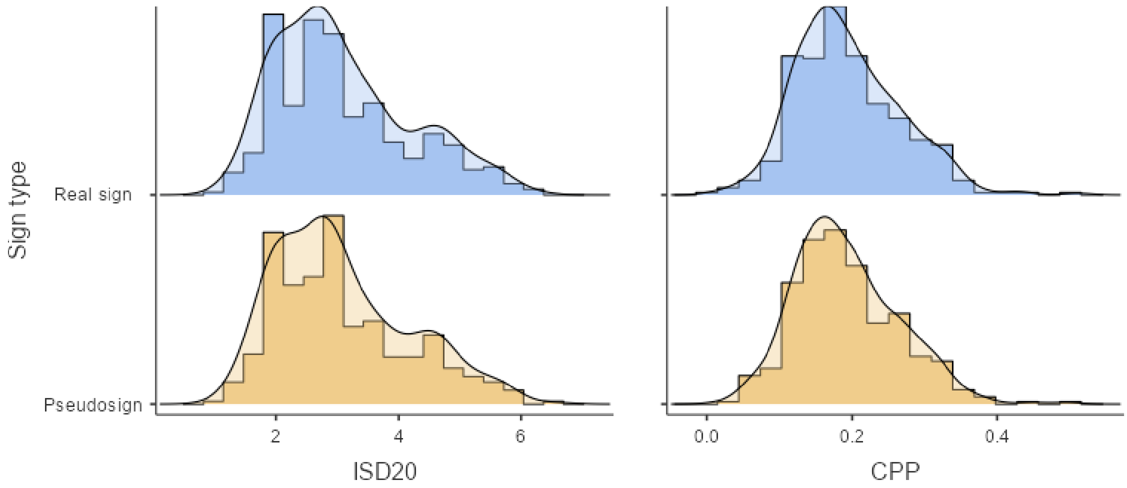


*Fig. 1: Histograms and density plots showing the distributions of the neighbourhood density metric ISD20 and the phonotactic probability metric CPP for the real signs compared to the pseudosigns used in the current study. Note that low ISD20 values indicate high neighbourhood density.*

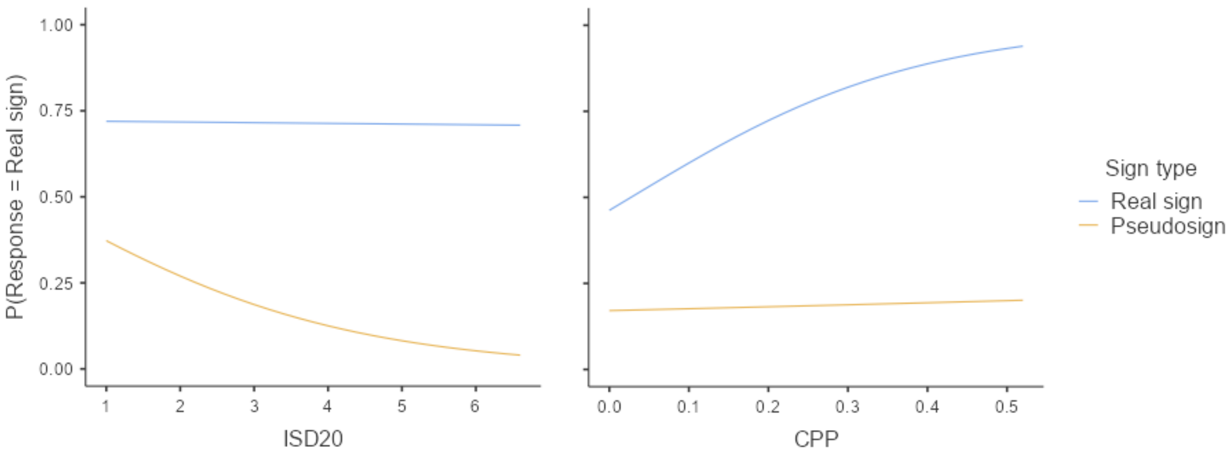


*Fig. 2: Marginal means for the probability of responding “real sign” in the lexical decision task given the presented sign type and the neighbourhood density metric ISD20 (left plot) and the phonotactic probability metric CPP (right plot). Note that low ISD20 values indicate high neighbourhood density.*

**References**


Caselli, Naomi K., Karen Emmorey, Ariel M. Cohen-Goldberg. (2021). The signed mental lexicon: Effects of phonological neighborhood density, iconicity, and childhood language experience. *Journal of Memory and Language* 121. 104282.

Williams, Joshua T., Adam Stone, Sharlene D. Newman. (2017). Operationalization of Sign Language Phonological Similarity and its Effects on Lexical Access. *The Journal of Deaf Studies and Deaf Education* 22(3). 303-315.

Vitevitch, Michael S., & Paul A. Luce. (2016). Phonological Neighborhood Effects in Spoken Word Perception and Production. *Annual Review of Linguistics* 2(1). 75-94.

Öqvist, Zrajm, Nikolaus Riemer Kankkonen, & Johanna Mesch. STS-korpus: A Sign Language Web Corpus Tool for Teaching and Public Use. *European Language Resources Association (LREC 2020).* 177-180.