

Abstrakt

Tato diplomová práce zkoumá průnik komputačních přístupů k jazyku, vědy o sítích a psycholingvistického výzkumu produkce slov. Práce představuje vědu o sítích spolu s jejím formalismem a aplikací v lingvistickém výzkumu ve formě fonologických a sémantických sítí. Představuje relevantní psycholingvistický výzkum zpracování slov, konkrétně *lexical decision task*, který vypovídá o efektivitě zpracování slov. Nakonec jsou představeny velké jazykové modely a vektory slov. Cílem této práce je zkonstruovat sémantickou síť angličtiny na základě slovních vektorů vytvořených jazykovým modelem BERT ze vzorku z The TV Corpus. Struktura výsledné sémantické sítě je analyzována ve světle výsledků z *lexical decision task* čerpaných z databáze MALD, které odrážejí efektivitu zpracování slov. Výsledná sémantická síť má strukturu malého světa (*small-world structure*), což znamená, že slovní vektory transformované do sémantické sítě mohou zachytit kognitivně salientní sémantické vztahy mezi slovy. Lineární regresní analýza mezi síťovými proměnnými *degree centrality*, *closeness centrality*, and *clustering coefficient* pro jednotlivá slova v sémantické síti a reakčním časem pro stejná slova z databáze MALD neprokázala statisticky významný vztah. Zdá se, že *clustering coefficient* má mírně negativní vztah k reakční době, který se blížil statistické významnosti, což znamená, že slova z hustších částí sítě jsou zpracovávána rychleji. Současné výsledky dovolují opatrný optimismus pro využití sémantických sítí založených na slovních vektorech pro výzkum kognitivních procesů, které jsou základem jazyka.

Klíčová slova: sémantická síť, zpracování slov, slovní vektory, strojové učení