

Testowanie hybrydowych modeli reprezentacji fonologicznej za pomocą korpusów mowy

Cel projektu

Celem niniejszego projektu jest sprawdzenie zasadności tzw. hybrydowych modeli przechowywania głosek w umyśle. W tym celu poszukiwane będą dowody na zasadność postulowania, iż użytkownicy języka przechowują zarówno ‘abstrakcje’ (czyli najistotniejsze cechy głosek, konieczne do rozróżniania wyrazów), jak i ‘egzemplarze’ (z angielskiego ‘exemplars’, ślady pamięciowe wyrazów zawierające także o wiele bardziej szczegółowe informacje dotyczące brzmienia wyrazów). Jako źródło danych użyte zostaną korpusy mowy, czyli duże zbiory nagrań rozmów zsynchronizowane z transkryptami.

Poprzez poszukiwanie dowodów na istnienie abstrakcji i egzemplarzy, projekt dąży do wsparcia integracji dwóch tradycji w fonologii (dziedzinie językoznawstwa zajmującej się organizacją systemów dźwięków w językach) w kwestii przechowywania głosek w umyśle. Z jednej strony mamy ujęcia formalne, zwykle skupiające się na abstrakcjach, a z drugiej ujęcia wywodzące strukturę fonologiczną z użycia języka (‘usage-based approaches’), zwykle skupiające się na roli egzemplarzy. Za korzystne należy uznać zwiększenie stopnia w jakim te ujęcia się porozumiewają. Większy zakres wymiany idei i wyników badań wpłynąłby korzystnie na rozwój całej dziedziny naukowej. Krokiem w tą stronę są modele hybrydowe, unifikujące te dwa ujęcia.

Badania realizowane w projekcie

Korpusy zawierające mowę w języku angielskim (w odmianie amerykańskiej i szkockiej) oraz polskim będą przeszukiwane specjalistycznym oprogramowaniem w celu zlokalizowania wyrazów, których cechy akustyczne mogą dowieść zasadności abstrakcji lub egzemplarzy. Do badanych zjawisk zaliczać się będą: wariacja w długości głosek trących (takich jak ‘s’ w ang. *speak* ‘mówić’ czy ‘th’ w ang. *thwart* ‘udaremnić’) zależna od słów, w których te głoski występują, redukcja długości słów zależna od tego, czy dane słowo zazwyczaj występuje w przewidywalnym kontekście, osłabienie głoski ‘r’ występującej po samogłoskach w szkockim akcencie języka angielskiego zależne od interakcji między stopniem znajomości między rozmówcami (znany/obcy) i słowami, których używają, i inne. Cechy nagrań audio będą mierzone automatycznie. Dodane zostaną dalsze informacje, które opisano w literaturze jako istotne (np. jak często używany jest dany wyraz, jak szybko mówi dana osoba) i włączone do modeli statystycznych. Dzięki takiemu doborowi zjawisk i sposobowi ich modelowania, zbadać będzie można rolę abstrakcji i egzemplarzy.

Na przykład, jeśli głoska ‘r’ jest osłabiona bardziej gdy rozmówcy się znają niż jeśli są sobie obcy, ale jeśli ta różnica wpływa na niektóre słowa bardziej niż na inne, będzie to dowodem na przechowywanie egzemplarzy (rozmówcy nie przełączają się z jednego stylu mowy na inny, lecz każde słowo zachowuje się indywidualnie). Jeśli natomiast głoska ‘r’ jest osłabiona gdy występuje po grupie samogłosek, o których można powiedzieć, iż tworzą klasę, gdyż mają ze sobą coś wspólnego (e.g. są wymawiane przez uniesienie języka), będzie to dowodem na przechowywanie abstrakcji (o wspólnych cechach głosek można mówić dopiero po abstrahowaniu od szczegółów wymowy). Jeśli odkryte zostaną oba efekty, będzie to dowodem na słuszność modeli hybrydowych.

Powody podjęcia danej tematyki badawczej

Kwestia przechowywania głosek w umyśle jako abstrakcji lub egzemplarzy jest kluczowa dla fonologii. Niedostateczny stopień integracji wyników badań przeprowadzanych w tych dwóch ujęciach spowalnia postęp dziedziny. Obecny projekt przyczyni się do owej integracji poprzez odkrycie, przy użyciu tej samej metody zastosowanej do podobnych zbiorów danych, dowodów na potrzebę postulowania zarówno abstrakcji i egzemplarzy. Co więcej, poprzez demonstrację jak dane pozyskane z mowy konwersacyjnej mogą zostać wykorzystane do argumentacji w teorii fonologicznej, projekt może przyczynić się do dalszej popularyzacji tego innowacyjnego sposobu prowadzenia badań.