

mgr inż. Magdalena Igras, dr inż. Bartosz Ziółko
Katedra Elektroniki, Akademia Górniczo-Hutnicza, Kraków

„Rodzaje pauz akustycznych i ich konotacje z interpunkcją w transkrypcjach mówionego języka polskiego”

Zespół Przetwarzania Sygnałów Katedry Elektroniki AGH prowadzi badania statystyczne nad językiem polskim na potrzeby automatycznego rozpoznawania mowy. Jednym z obecnie analizowanych zagadnień jest możliwość wnioskowania na temat quasi-interpukcji w mówionym języku polskim na podstawie różnych parametrów, w tym pauz akustycznych, w celu automatycznego wstawiania interpukcji w transkrypcjach wypowiedzi.

Badania wykazują trzy możliwe typy pauz akustycznych w języku mówionym. Pierwszym, najbardziej intuicyjnym, ale zarazem nie najczęstszym, jest cisza. W zależności od osoby mówiącej i kontekstu sytuacyjnego, może charakteryzować się różną długością. Kolejnym typem są pauzy wypełnione, a więc pseudo-słowa, frazy nie mające wpływu na znaczenie zdania, takie jak *yyy*, *eee*, *hmm*, *mmm*, zaburzające ciągłość wypowiedzi. Często mogą wskazywać na potrzebę wstawienia przecinka lub czasami kropki w transkrypcjach wypowiedzi, w których się pojawiają. Brzmienie pauz wypełnionych jest zależne od języka (np. w języku angielskim najczęstsze jest *um*). W wypadku języka polskiego najczęstszymi okazują się *yyy/yh* oraz *mmm*. Trzecim badanym przez nas zjawiskiem w języku mówionym są pauzy oddechowe. Są one silnym wskazaniem na wstawienie kropki w transkrypcji.

Ze względu na genezę użycia pauz wyodrębniliśmy: 1) regularne, naturalne pauzy - spowodowane oddechem przed kolejną wypowiedzianą frazą (pauzy oddechowe), 2) nieregularne pauzy intencjonalne - celowe użycie pauzy jako środka stylistycznego, zwłaszcza przez profesjonalnych mówców (pauzy ciche) oraz 3) nieregularne, nieintencjonalne zaburzenia ciągłości, efekt niepewności, namysłu lub wahania, u niedoświadczonych mówców nawet kilkanaście takich nieciągłości na minutę (zjawiska akustyczne typu cisza i pauzy wypełnione).

Przeanalizowaliśmy materiał badawczy złożony z nagrań monologów zarówno mówców niedoświadczonych, jak i profesjonalnych, w sytuacjach o charakterze formalnym. Pośród analizowanych nagrań znalazły się prezentacje studentów, wypowiedzi profesorów z posiedzenia uczelnianej Rady Wydziału, wystąpienia z obrad Parlamentu Europejskiego oraz tłumaczenia w czasie rzeczywistym wypowiedzi obcojęzycznych z obrad Parlamentu Europejskiego. Pierwszym krokiem badań było dokonanie manualnej transkrypcji nagrań oraz, na bazie obserwacji z tego procesu, wyznaczenie czynników determinujących nieprecyzyjność bądź utrudniających jednoznaczną transkrypcję uwzględniającą interpukcję.

Następnie zbadaliśmy ilościowo częstość używania poszczególnych typów pauz oraz znaków interpukcyjnych w ich transkrypcjach. Sprawdziliśmy statystycznie, jak często wystąpienie każdego z wyróżnionych typów pauz było nośnikiem informacji o interpukcji. Wnioski z tej analizy są jedną z kluczowych informacji dla projektowania algorytmów automatycznego wstawiania interpukcji.

Badając interpunkcję przy transkrypcji języka mówionego należy zwrócić uwagę również na inne zjawiska nieciągłości, powszechnie występujące w mowie spontanicznej, które zaburzają granice zdań. Do najważniejszych zaliczamy przerwanie wypowiedzenia początku zdania i rozpoczęcie od nowa, zmianę szyku zdania w trakcie wypowiedzi oraz wtrącenia zdań. Natomiast w obrębie wyrazów najczęściej występującymi nieciągłościami są powtórzenia, poprawienia oraz prolongacje niektórych spójników, przyimków i końcowych sylab. Zliczenia występowania poszczególnych nieciągłości wskazują, że są to zjawiska na tyle częste w mowie spontanicznej, że powinny być uwzględniane w systemach rozpoznawania mowy.

Badania nad interpunkcją w mowie spontanicznej służą wzbogaceniu automatycznych transkrypcji mowy spontanicznej o interpunkcję, co podwyższy czytelność i przydatność automatycznych transkrypcji, a zarazem dostosuje je do przetwarzania przez modele językowe, które operują na pełnych zdaniach oraz nie uwzględniają nieciągłości mowy.

Projekt został sfinansowany ze środków Narodowego Centrum Nauki przyznanych na podstawie decyzji numer DEC-2011/03/D/ST6/00914.