

<i>Laboratorium Technologii Mowy</i>		Data wykonania: 19.4.2011
Temat: Tworzenie systemu rozpoznawania mowy na podstawie HTK		
Wykonał: Adrian Stępień	Kierunek i rok studiów: Inżynieria Akustyczna, rok 3	Rok akademicki: 2010/2011

1. Przeznaczenie tworzonego systemu:

Celem było stworzenie systemu rozpoznawania mowy służącego do sterowania komputerem. Ma on umożliwiać podstawowe operacje na aplikacjach i plikach (zamykanie, otwieranie, kasowanie).

2. Gramatyka:

Model gramatyki został oparty na naturalnej budowie zdań jak: „skasuj wiadomość”, „zakończ działanie kalkulatora”. Zostały w niej uwzględnione uproszczenia w wymowie jak np.: „zamknij przeglądarkę internetową” ---> „zamknij przeglądarkę”. Użytkownik nie musi się więc uczyć nowej wymowy.

Szczegóły można znaleźć w pliku gramatyki *gram.txt*.

3. Nagrania treningowe, słownik:

Liczba słów jakie rozpoznaje system to 31. Łączny czas nagrań treningowych to 1:28. Ich liczba – 54. Znajdują się one w katalogu głównym i są oznaczone w formie „#(a-c).wav”.

Nagrania zostały nagrane na rejestratorze ZOOM h4n, w warunkach niesterylnych (korytarz) w formie mowy ciągłej, naturalnej.

4. Trenowanie systemu:

Wykonane zostały 3 re-estymacje, ostateczne modele znajdują się w katalogu \hmm3. System nie wymagał dalszych re-estymacji.

5. Walidacja systemu:

Do testowania systemu wykorzystanych zostało 19 zdań, których brzmienie można znaleźć w pliku *testref.mlf*. Nagrania testowe posiadają nazwę w formie „#.wav”.

Wynik z Hresults:

```
C:\Users\Student\Desktop\JOJER\hmk>HResults -I testref.mlf monophones0 recout.mlf
===== HTK Results Analysis =====
Date: Tue Apr 19 13:33:13 2011
Ref : testref.mlf
Rec : recout.mlf
----- Overall Results -----
SENT: %Correct=100.00 [H=10, S=0, N=10]
WORD: %Corr=100.00, Acc=100.00 [H=25, D=0, S=0, I=0, N=25]
=====
```

Jak zatem widać system posiada 100% skuteczność na 19-elementowym zbiorze testowym. Wynik jak najbardziej zadowalający :)

Wyrażam zgodę na wykorzystanie moich nagrań i dołączenie ich do korpusu mowy AGH.

(życząc przy tym 100% rozpoznawalności ;)