

Sprawozdanie z laboratoriów HTK

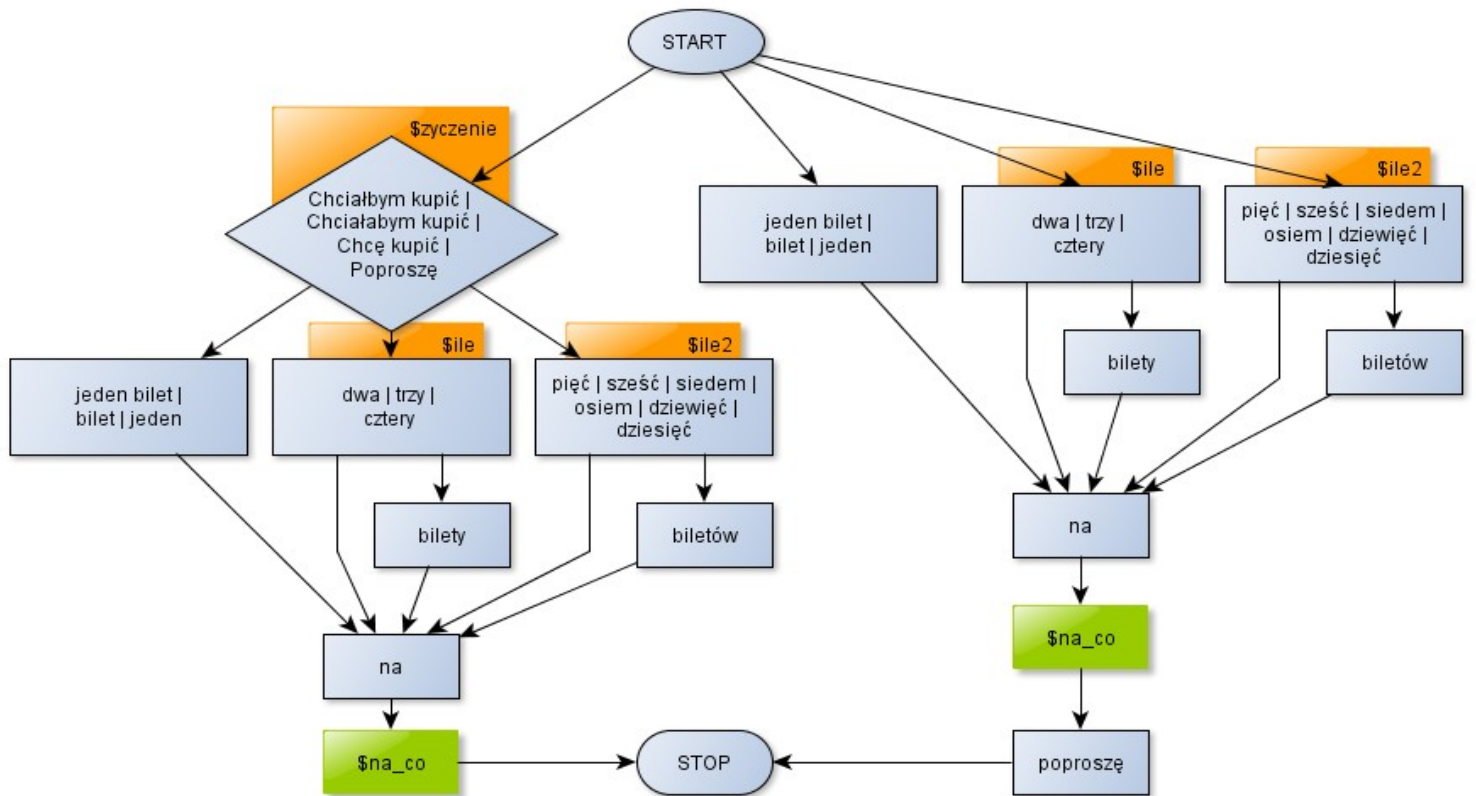
1. Opis gramatyki

System był projektowany jako automat sprzedający bilety na koncerty. Miał on za zadanie rozpoznawać dyspozycje kierowane przez klienta podczas zakupu biletów na określony koncert. Najistotniejsze elementy takiej dyspozycji to ilość biletów oraz nazwa zespołu. Naszym założeniem było, aby zaproponowana gramatyka była niezależna od płci klienta oraz definiowała możliwie jak najwięcej naturalnie brzmiących dyspozycji, jakie mógłby wydać potencjalny klient. Dla uproszczenia wprowadzonych zostało jedynie 5 nazw zespołów, jednak w łatwy sposób można dodać do systemu kolejne zespoły, pamiętając o odmianie ich nazw przez przypadki. 100 % zdań generowanych przez zaproponowaną gramatykę brzmi naturalnie i jest zrozumiałe w języku polskim, tym samym uznaliśmy, że napisana gramatyka spełniła swoje założenia, a sekwencje wygenerowane przez narzędzie HSGen posłużyły nam bez żadnych poprawek jako tekst do nagrań zbioru treningowego oraz zbioru testowego.

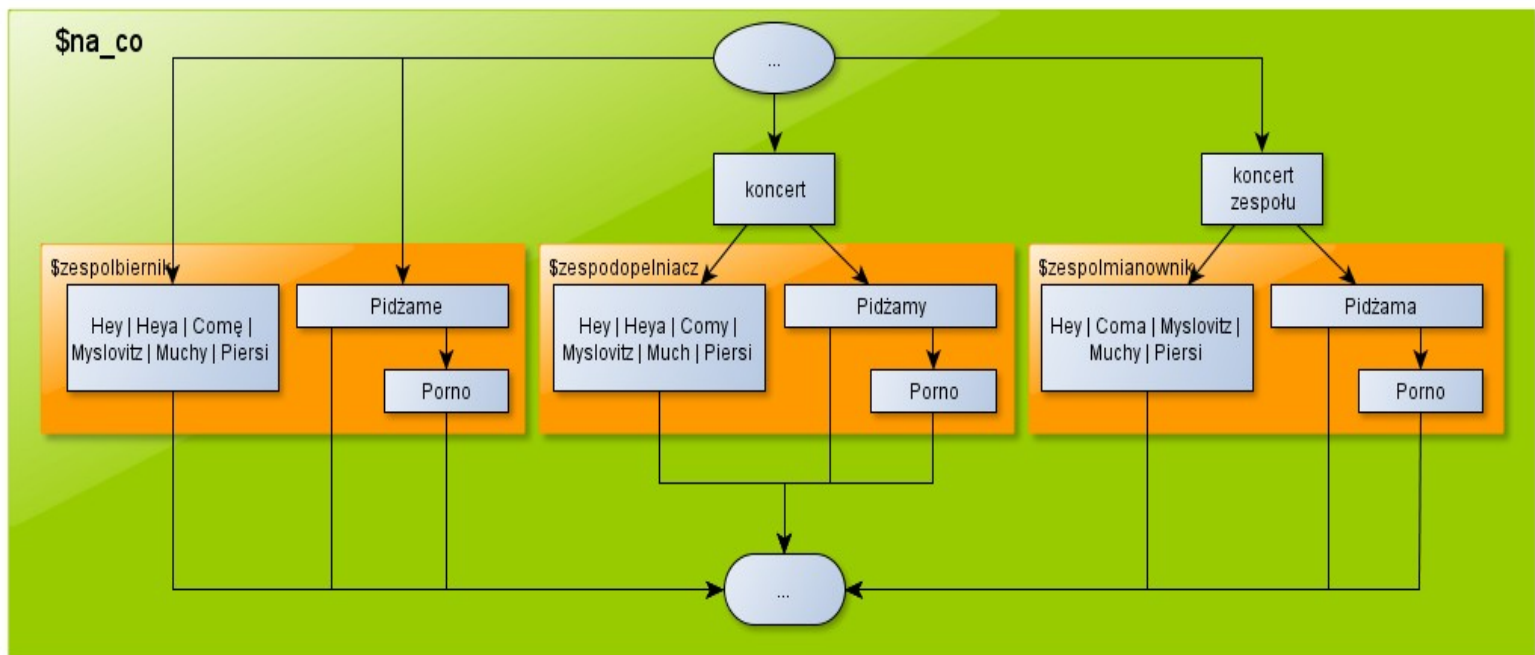
Plik opisujący gramatykę miał następującą składnię:

```
$zyczenie = (chciaLbym | chciaLabym | chcE) kupic | poproszE;  
  
$ile2 = piEc | szeSC | siedem | osiem | dziewiEc | dziesiEc;  
$ile = dwa | trzy | cztery;  
  
$zespolmianownik = Hey | Coma | Myslovitz | Muchy | Piersi | PidZama  
[Porno];  
$zespolbiernik = Hey | Heya | ComE | Myslovitz | Muchy | Piersi | PidZame  
[Porno];  
$zespoldopelniacz = Hey | Heya | Comy | Myslovitz | Much | Piersi |  
(PidZamy|PidZama) [Porno];  
  
$naco = ($zespolbiernik) | (koncert $zespoldopelniacz) | (koncert_zespoLu  
$zespolmianownik);  
  
( SENT-START (  
[$zyczenie] ([jeden] bilet|jeden [bilet]) na $naco |  
[$zyczenie] $ile [bilety] na $naco |  
[$zyczenie] $ile2 [biletOw] na $naco |  
([jeden] bilet|jeden [bilet]) na $naco poproszE |  
$ile [bilety] na $naco poproszE |  
$ile2 [biletOw] na $naco poproszE  
) SENT-END )
```

Schemat blokowy zaproponowanej gramatyki:



gdzie:



2. Opis nagrań

Do nagrań zarówno zbioru treningowego, jak i zbioru testowego, wykorzystany został mikrofon Behringer ECM8000 podłączony do interfejsu M-Audio Fast Track Pro. Nagrania zostały przeprowadzone w cichym pokoju mieszkalnym, dzięki czemu były w dużym stopniu pozbawione wpływu innych dźwięków zakłócających. Użycie mikrofonu o w przybliżeniu płaskiej charakterystyce częstotliwościowej pozwoliło na uzyskanie brzmienia głosu lektora podobnego do słyszalnego w rzeczywistości.

Naszym celem było sprawdzenie, jak opracowywany system będzie radził sobie z naturalną mową ciągłą. Przyjęcie takiego założenia spowodowało, że również do nagrań wykorzystano mowę ciągłą, a tym samym w zdaniach wypowiedzianych przez lektora brak jest przerw między poszczególnymi słowami. W konsekwencji, anotacja słów w naszych nagraniach nie była łatwym zadaniem – często trudno było określić granicę między ostatnią literą słowa poprzedzającego, a pierwszą literą słowa następnego. Łącznie nasz zbiór treningowy składał się z 100 zdań, które łącznie stanowiło 3 minuty 56 sekund nagrania. Zbiór testowy nagrany został w podobnych warunkach z wykorzystaniem głosu tego samego lektora i składał się on z 50 zdań, które łącznie stanowiło 1 minutę 56 sekund nagrania.

3. Warunki testowania i wyniki testów z HResults

Nasz zbiór treningowy składał się z 100 nagrań poszczególnych zdań mowy ciągłej, natomiast zbiór testowy składał się z 50 nagrań poszczególnych zdań wymawianych w ten sam sposób przez tego samego lektora.

Wyniki uzyskane przez nas dla 15-krotnej estymacji przy 50-ciu próbkach testowych zostały przedstawione poniżej:

```
===== HTK Results Analysis =====  
  
Date: Mon May 28 17:56:06 2012  
  
Ref : testref_norm50.mlf  
  
Rec : recout_50.mlf  
  
----- Overall Results -----  
  
SENT: %Correct=0.00 [H=0, S=50, N=50]  
  
WORD: %Corr=36.10, Acc=28.22 [H=87, D=49, S=105, I=19, N=241]  
  
=====
```

Jednak przetestowanie różnej liczby reestymacji mogłoby wpłynąć na poprawę wyników, dlatego wyniki dla kolejnych estymacji, od 0 do 15, przedstawiono w poniższej tabeli i poniższym wykresie.

Wyniki uzyskane przez nas dla kolejnych estymacji, od 0 do 15, przedstawiają się następująco:

Numer kolejnej estymacji	Uzyskane wyniki
0	<pre> ===== HTK Results Analysis ===== Date: Wed May 30 13:39:36 2012 Ref : testref_norm50.mlf Rec : recout/recout0.mlf ----- Overall Results ----- SENT: %Correct=0.00 [H=0, S=50, N=50] WORD: %Corr=26.56, Acc=26.56 [H=64, D=91, S=86, I=0, N=241] ===== </pre>
1	<pre> ===== HTK Results Analysis ===== Date: Wed May 30 13:39:46 2012 Ref : testref_norm50.mlf Rec : recout/recout1.mlf ----- Overall Results ----- SENT: %Correct=0.00 [H=0, S=50, N=50] WORD: %Corr=31.95, Acc=28.22 [H=77, D=59, S=105, I=9, N=241] ===== </pre>
2	<pre> ===== HTK Results Analysis ===== Date: Wed May 30 13:39:56 2012 Ref : testref_norm50.mlf Rec : recout/recout2.mlf ----- Overall Results ----- SENT: %Correct=4.00 [H=2, S=48, N=50] WORD: %Corr=40.25, Acc=31.95 [H=97, D=35, S=109, I=20, N=241] ===== </pre>

=====
HTK Results Analysis
=====
Date: Wed May 30 13:40:06 2012
Ref : testref_norm50.mlf
3 Rec : recout/recout3.mlf
----- Overall Results -----
SENT: %Correct=0.00 [H=0, S=50, N=50]
WORD: %Corr=35.68, Acc=26.97 [H=86, D=42, S=113, I=21, N=241]
=====

=====
HTK Results Analysis
=====
Date: Wed May 30 13:40:16 2012
Ref : testref_norm50.mlf
4 Rec : recout/recout4.mlf
----- Overall Results -----
SENT: %Correct=0.00 [H=0, S=50, N=50]
WORD: %Corr=36.51, Acc=28.22 [H=88, D=43, S=110, I=20, N=241]
=====

=====
HTK Results Analysis
=====
Date: Wed May 30 13:40:26 2012
Ref : testref_norm50.mlf
5 Rec : recout/recout5.mlf
----- Overall Results -----
SENT: %Correct=0.00 [H=0, S=50, N=50]
WORD: %Corr=38.17, Acc=30.29 [H=92, D=43, S=106, I=19, N=241]
=====

=====
HTK Results Analysis
=====
Date: Wed May 30 13:40:37 2012
Ref : testref_norm50.mlf
6 Rec : recout/recout6.mlf

----- Overall Results -----

SENT: %Correct=0.00 [H=0, S=50, N=50]

WORD: %Corr=36.51, Acc=28.22 [H=88, D=47, S=106, I=20, N=241]

=====

===== HTK Results Analysis =====

Date: Wed May 30 13:40:48 2012

Ref : testref_norm50.mlf

7 Rec : recout/recout7.mlf

----- Overall Results -----

SENT: %Correct=0.00 [H=0, S=50, N=50]

WORD: %Corr=36.51, Acc=28.22 [H=88, D=48, S=105, I=20, N=241]

=====

===== HTK Results Analysis =====

Date: Wed May 30 13:40:59 2012

Ref : testref_norm50.mlf

8 Rec : recout/recout8.mlf

----- Overall Results -----

SENT: %Correct=0.00 [H=0, S=50, N=50]

WORD: %Corr=34.44, Acc=25.73 [H=83, D=50, S=108, I=21, N=241]

=====

===== HTK Results Analysis =====

Date: Wed May 30 13:41:10 2012

Ref : testref_norm50.mlf

9 Rec : recout/recout9.mlf

----- Overall Results -----

SENT: %Correct=0.00 [H=0, S=50, N=50]

WORD: %Corr=33.20, Acc=24.48 [H=80, D=52, S=109, I=21, N=241]

=====

=====
HTK Results Analysis
=====
Date: Wed May 30 13:41:19 2012
Ref : testref_norm50.mlf
10 Rec : recout/recout10.mlf
----- Overall Results -----
SENT: %Correct=0.00 [H=0, S=50, N=50]
WORD: %Corr=34.85, Acc=26.14 [H=84, D=52, S=105, I=21, N=241]
=====

=====
HTK Results Analysis
=====
Date: Wed May 30 13:41:30 2012
Ref : testref_norm50.mlf
11 Rec : recout/recout11.mlf
----- Overall Results -----
SENT: %Correct=0.00 [H=0, S=50, N=50]
WORD: %Corr=34.02, Acc=25.73 [H=82, D=53, S=106, I=20, N=241]
=====

=====
HTK Results Analysis
=====
Date: Wed May 30 13:41:39 2012
Ref : testref_norm50.mlf
12 Rec : recout/recout12.mlf
----- Overall Results -----
SENT: %Correct=0.00 [H=0, S=50, N=50]
WORD: %Corr=33.20, Acc=24.90 [H=80, D=54, S=107, I=20, N=241]
=====

=====
HTK Results Analysis
=====
Date: Wed May 30 13:41:50 2012
Ref : testref_norm50.mlf
13 Rec : recout/recout13.mlf

----- Overall Results -----

SENT: %Correct=0.00 [H=0, S=50, N=50]

WORD: %Corr=36.51, Acc=27.80 [H=88, D=50, S=103, I=21, N=241]

===== HTK Results Analysis =====

Date: Wed May 30 13:42:01 2012

Ref : testref_norm50.mlf

14 Rec : recout/recout14.mlf

----- Overall Results -----

SENT: %Correct=0.00 [H=0, S=50, N=50]

WORD: %Corr=35.68, Acc=26.97 [H=86, D=52, S=103, I=21, N=241]

===== HTK Results Analysis =====

Date: Wed May 30 13:42:18 2012

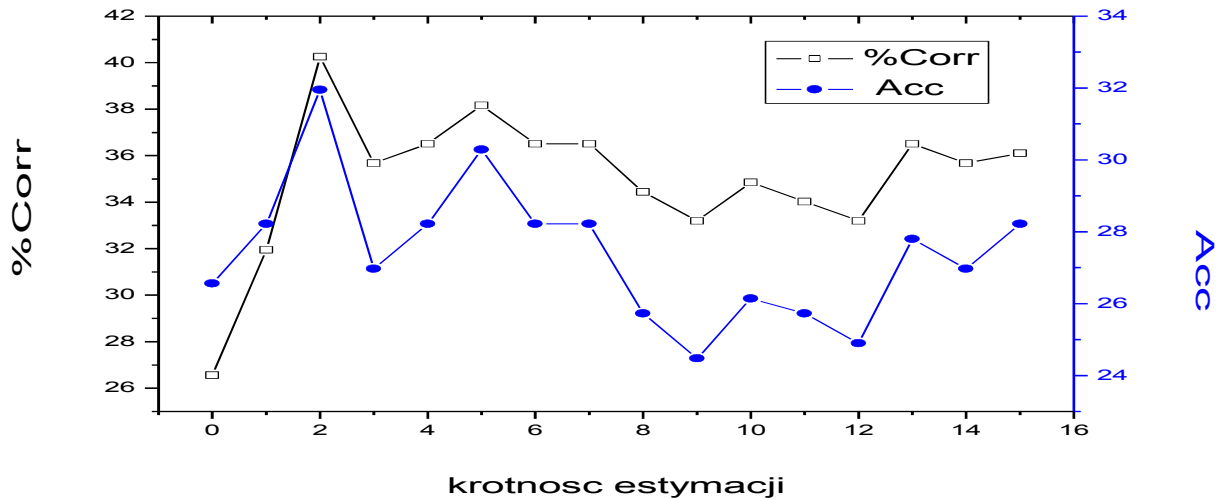
Ref : testref_norm50.mlf

15 Rec : recout/recout15.mlf

----- Overall Results -----

SENT: %Correct=0.00 [H=0, S=50, N=50]

WORD: %Corr=36.10, Acc=28.22 [H=87, D=49, S=105, I=19, N=241]



Analizując powyższe wyniki, można zauważyć, że przy estymacji z numerem 0 uzyskane wyniki rozpoznawalności mowy przyjmują najniższe wartości [WORD: %Corr=26.56, Acc=26.56 [H=64, D=91, S=86, I=0, N=241]]. Wraz z kolejnymi pięcioma estymacjami wyniki poprawności wzrastają [do wartości: WORD: %Corr=38.17, Acc=30.29 [H=92, D=43, S=106, I=19, N=241] przy 5 estymacji], jednak to przy 2 estymacji wyniki rozpoznawalności mowy przyjmują najwyższe wartości [WORD: %Corr=40.25, Acc=31.95 [H=97, D=35, S=109, I=20, N=241]], co odpowiada poprawnemu rozpoznaniu 4 sekwencji. Przy kolejnych czterech estymacjach rozpoznawalność mowy spada [do wyników: WORD: %Corr=33.20, Acc=24.48 [H=80, D=52, S=109, I=21, N=241] przy 9 estymacji]. Zjawisko takie mogłoby sugerować przetrenowanie systemu, jednak w kolejnych przeprowadzonych estymacjach można zauważyć zarówno wzrost, jak i spadek wartości opisujących rozpoznawalność mowy. Przyjmując wyniki 15 estymacji [WORD: %Corr=36.10, Acc=28.22 [H=87, D=49, S=105, I=19, N=241]] za średnią rozpoznawalność tego systemu, można uznać te wyniki rozpoznawalności mowy za stosunkowo dobre, jednak dalekie od ideału.

4. Analiza błędów rozpoznania

System miał problemy z rozróżnianiem słów „jeden” i „siedem”, ze względu na ich podobne brzmienie. System miał również problemy z rozpoznawaniem poszczególnych odmian nazw zespołów. Bez względu jednak na rozpoznanie, następująca po liczbie biletów odmiana słowa „bilet”, zależna od wcześniejszej ilości wybranych biletów, była poprawna z punktu widzenia reguł języka polskiego. System też ani razu nie wytworzył zdania niepoprawnego, bądź brzmiącego nienaturalnie, co sugeruje poprawność konstrukcji gramatyki systemu, natomiast powstałe błędy mogą wynikać z samego rozpoznawania oraz ewentualnych popełnionych błędów w wyborze do nagrań mowy ciągłej zamiast mowy izolowanej, a w konsekwencji trudnościach w anotacji poszczególnych słów w wypowiedzianych przez lektora zdaniach.

5. Analiza różnych rozwiązań

Niskie wyniki procentowych wartości opisujących rozpoznawalność mowy mogły być spowodowane wykorzystaniem w nagraniach mowy ciągłej, ze względu na przyjęte założenie o jak najwierniejszym odzwierciedleniu w nagraniach naturalnej mowy ludzkiej, a także konsekwencjami w trudnościach w anotacji poszczególnych słów w zdaniach w nagraniach ze zbioru treningowego. Zapewne wykorzystanie w nagraniach mowy izolowanej poskutkowałoby większą rozpoznawalnością poszczególnych słów przez system.

Kolejnym stopniem trudności dla systemu mogłaby być również ocena rozpoznawalności mowy, kiedy sygnał mowy jest zakłócany, symulując tym samym bardziej rzeczywiste warunki zakupu. Mogłoby to zostać uzyskane na przykład poprzez wykorzystanie jako zbioru testowego nagrań sygnału mowy z dodanym szumem, reprezentującym rzeczywiste tło akustyczne, a więc takim, jak na przykład występująca w tle dyspozycji klienta muzyka, odgłosy ulicy ruchliwego miasta czy pomieszczenia, w którym w niewielkiej odległości prowadzone są inne rozmowy, bądź też występuje ogólny hałas biurowy.

Przeprowadzenie tego typu testów z pewnością wpłynęłoby na większą wiarygodność wyników, jak również mogłoby być pomocne przy eliminowaniu ewentualnych błędów rozpoznawania i kolejnych krokach doskonalenia systemu.

6. Analiza wpływu tła akustycznego

Postanowiliśmy tutaj sprawdzić, jak system zachowa się w sytuacji, gdy sygnał mowy zostanie minimalnie zaszumiony, symulując rzeczywiste tło akustyczne i rozmowy prowadzone w niewielkiej odległości w stosunku do klienta wydającego dyspozycję zakupu biletów w naszym systemie.

Jako zbiór testowy użyto tych samych 50 nagrań, co poprzednio, z dodanym na niewielkim poziomie szumem rozmów prowadzonych w tle. Wyniki uzyskane przez nas dla 15-krotnej estymacji przy opisanych 50-ciu próbkach testowych zostały przedstawione poniżej:

```
===== HTK Results Analysis =====  
  
Date: Sun Jul 08 14:29:44 2012  
  
Ref : testref_norm50.mlf  
  
Rec : recout/recout15.mlf  
  
----- Overall Results -----  
  
SENT: %Correct=0.00 [H=0, S=50, N=50]  
  
WORD: %Corr=36.10, Acc=28.22 [H=87, D=51, S=103, I=19, N=241]  
  
=====
```

Niewiele więc różni się od wyników dla testowanego wcześniej niezaszumionego sygnału mowy.

Jednak przetestowanie różnej liczby reestymacji mogłoby wpłynąć na poprawę wyników, dlatego wyniki dla kolejnych estymacji, od 0 do 15, przedstawiono w poniższej tabeli i poniższym wykresie.

Wyniki uzyskane przez nas dla kolejnych estymacji, od 0 do 15, przedstawiają się następująco:

Numer kolejnej estymacji	Uzyskane wyniki
0	<pre> ===== HTK Results Analysis ===== Date: Sun Jul 08 14:27:30 2012 Ref : testref_norm50.mlf Rec : recout/recout0.mlf ----- Overall Results ----- SENT: %Correct=0.00 [H=0, S=50, N=50] WORD: %Corr=26.56, Acc=26.56 [H=64, D=91, S=86, I=0, N=241] ===== </pre>
1	<pre> ===== HTK Results Analysis ===== Date: Sun Jul 08 14:27:41 2012 Ref : testref_norm50.mlf Rec : recout/recout1.mlf ----- Overall Results ----- SENT: %Correct=0.00 [H=0, S=50, N=50] WORD: %Corr=26.97, Acc=26.97 [H=65, D=91, S=85, I=0, N=241] ===== </pre>
2	<pre> ===== HTK Results Analysis ===== Date: Sun Jul 08 14:27:52 2012 Ref : testref_norm50.mlf Rec : recout/recout2.mlf ----- Overall Results ----- SENT: %Correct=4.00 [H=2, S=48, N=50] WORD: %Corr=43.57, Acc=39.00 [H=105, D=34, S=102, I=11, N=241] ===== </pre>

=====
HTK Results Analysis
=====
Date: Sun Jul 08 14:28:02 2012
Ref : testref_norm50.mlf
3 Rec : recout/recout3.mlf
----- Overall Results -----
SENT: %Correct=0.00 [H=0, S=50, N=50]
WORD: %Corr=36.10, Acc=25.31 [H=87, D=36, S=118, I=26, N=241]
=====

=====
HTK Results Analysis
=====
Date: Sun Jul 08 14:28:15 2012
Ref : testref_norm50.mlf
4 Rec : recout/recout4.mlf
----- Overall Results -----
SENT: %Correct=2.00 [H=1, S=49, N=50]
WORD: %Corr=38.17, Acc=33.61 [H=92, D=44, S=105, I=11, N=241]
=====

=====
HTK Results Analysis
=====
Date: Sun Jul 08 14:28:23 2012
Ref : testref_norm50.mlf
5 Rec : recout/recout5.mlf
----- Overall Results -----
SENT: %Correct=4.00 [H=2, S=48, N=50]
WORD: %Corr=40.25, Acc=36.93 [H=97, D=43, S=101, I=8, N=241]
=====

=====
HTK Results Analysis
=====
Date: Sun Jul 08 14:28:31 2012
Ref : testref_norm50.mlf
6 Rec : recout/recout6.mlf

----- Overall Results -----

SENT: %Correct=0.00 [H=0, S=50, N=50]

WORD: %Corr=37.76, Acc=32.78 [H=91, D=49, S=101, I=12, N=241]

=====

===== HTK Results Analysis =====

Date: Sun Jul 08 14:28:39 2012

Ref : testref_norm50.mlf

7 Rec : recout/recout7.mlf

----- Overall Results -----

SENT: %Correct=0.00 [H=0, S=50, N=50]

WORD: %Corr=41.08, Acc=36.51 [H=99, D=47, S=95, I=11, N=241]

=====

===== HTK Results Analysis =====

Date: Sun Jul 08 14:28:48 2012

Ref : testref_norm50.mlf

8 Rec : recout/recout8.mlf

----- Overall Results -----

SENT: %Correct=0.00 [H=0, S=50, N=50]

WORD: %Corr=39.00, Acc=32.37 [H=94, D=49, S=98, I=16, N=241]

=====

===== HTK Results Analysis =====

Date: Sun Jul 08 14:28:56 2012

Ref : testref_norm50.mlf

9 Rec : recout/recout9.mlf

----- Overall Results -----

SENT: %Correct=0.00 [H=0, S=50, N=50]

WORD: %Corr=37.34, Acc=31.12 [H=90, D=49, S=102, I=15, N=241]

=====

=====
HTK Results Analysis
=====
Date: Sun Jul 08 14:29:05 2012
Ref : testref_norm50.mlf
10 Rec : recout/recout10.mlf
----- Overall Results -----
SENT: %Correct=0.00 [H=0, S=50, N=50]
WORD: %Corr=38.17, Acc=31.54 [H=92, D=50, S=99, I=16, N=241]
=====

=====
HTK Results Analysis
=====
Date: Sun Jul 08 14:29:13 2012
Ref : testref_norm50.mlf
11 Rec : recout/recout11.mlf
----- Overall Results -----
SENT: %Correct=0.00 [H=0, S=50, N=50]
WORD: %Corr=37.34, Acc=30.29 [H=90, D=51, S=100, I=17, N=241]
=====

=====
HTK Results Analysis
=====
Date: Sun Jul 08 14:29:20 2012
Ref : testref_norm50.mlf
12 Rec : recout/recout12.mlf
----- Overall Results -----
SENT: %Correct=0.00 [H=0, S=50, N=50]
WORD: %Corr=37.76, Acc=30.71 [H=91, D=51, S=99, I=17, N=241]
=====

=====
HTK Results Analysis
=====
Date: Sun Jul 08 14:29:30 2012
Ref : testref_norm50.mlf
13 Rec : recout/recout13.mlf

----- Overall Results -----

SENT: %Correct=0.00 [H=0, S=50, N=50]

WORD: %Corr=37.34, Acc=29.46 [H=90, D=52, S=99, I=19, N=241]

===== HTK Results Analysis =====

Date: Sun Jul 08 14:29:37 2012

Ref : testref_norm50.mlf

14

Rec : recout/recout14.mlf

----- Overall Results -----

SENT: %Correct=0.00 [H=0, S=50, N=50]

WORD: %Corr=37.76, Acc=29.88 [H=91, D=51, S=99, I=19, N=241]

===== HTK Results Analysis =====

Date: Sun Jul 08 14:29:44 2012

Ref : testref_norm50.mlf

15

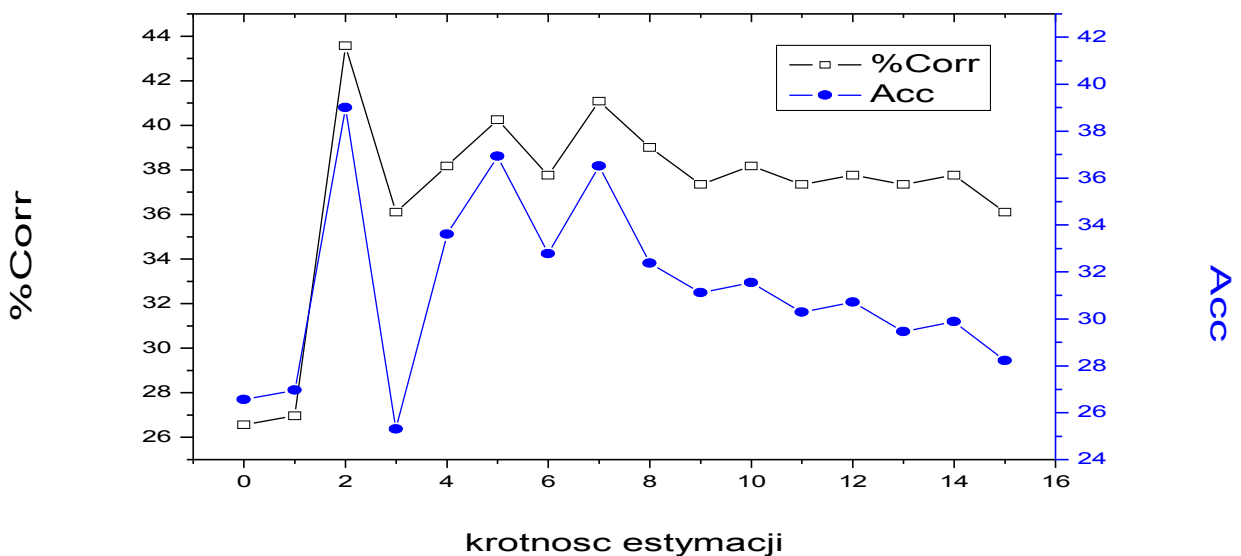
Rec : recout/recout15.mlf

----- Overall Results -----

SENT: %Correct=0.00 [H=0, S=50, N=50]

WORD: %Corr=36.10, Acc=28.22 [H=87, D=51, S=103, I=19, N=241]

===== HTK Results Analysis =====



Analizując powyższe wyniki, można zauważyć, że przy estymacji z numerem 0 i 15 uzyskane wyniki rozpoznawalności mowy dla zaszumionego sygnału niewiele różnią się od wyników dla testowanego wcześniej czystego, niezaszumionego sygnału mowy. W przypadku estymacji z numerem 1 dla sygnału zaszumionego uzyskano nieco niższe wartości procentowe parametrów opisujących rozpoznawalność mowy, a w przypadku 3 estymacji uzyskano wyższe wartości %Corr, lecz niższe Acc niż poprzednio dla czystego, niezaszumionego sygnału mowy. Tendencja spadkowa zauważalna na wykresie już od 9 estymacji może sugerować natomiast przetrenowanie systemu.

Co ciekawe, dla estymacji z numerem 2 uzyskano nawet wyższe wyniki rozpoznawalności mowy dla sygnału zaszumionego [WORD: %Corr=43.57, Acc=39.00 [H=105, D=34, S=102, I=11, N=241]] niż poprzednio wyniki dla czystego, niezaszumionego sygnału mowy [WORD: %Corr=40.25, Acc=31.95 [H=97, D=35, S=109, I=20, N=241]], co również miało miejsce w przypadku estymacji z numerem 4, gdzie uzyskano wyniki [WORD: %Corr=38.17, Acc=33.61 [H=92, D=44, S=105, I=11, N=241]], odpowiadające 2 poprawnie rozpoznany sekwencjom, a więc wyższe w stosunku do poprzednich wyników [WORD: %Corr=36.51, Acc=28.22 [H=88, D=43, S=110, I=20, N=241]] i braku rozpoznania sekwencji w przypadku czystego, niezaszumionego sygnału mowy. Sytuacja powtórzyła się również dla kolejnej 5-tej estymacji, gdzie dla zaszumionego sygnału mowy uzyskano wyniki [WORD: %Corr=40.25, Acc=36.93 [H=97, D=43, S=101, I=8, N=241]], odpowiadające ponownie poprawnemu rozpoznaniu 4 sekwencji, a więc wyższe w stosunku do wyników dla czystego, niezaszumionego sygnału mowy [WORD: %Corr=38.17, Acc=30.29 [H=92, D=43, S=106, I=19, N=241]], gdzie sekwencje nie zostały rozpoznane. Podobnie dla kolejnych estymacji, zarówno wyniki %Corr, jak i Acc, były wyższe dla zaszumionego sygnału mowy niż poprzednio dla czystego, niezaszumionego sygnału mowy. Na tej podstawie można więc wyciągnąć wniosek, że nasz system wciąż uczy się wraz z każdą kolejną przeprowadzaną estymacją.

Porównując wyniki uzyskane dla zaszumionego sygnału mowy, symulującego wydawanie dyspozycji zakupu biletów w rzeczywistym środowisku akustycznym, w którym trudno ustrzec się wpływu innych sygnałów akustycznych, z wynikami uzyskanymi uprzednio dla czystego niezaszumionego sygnału mowy, można stwierdzić, że nasz system radzi sobie równie dobrze z rozpoznawaniem zaszumionego na niewielkim poziomie sygnału mowy, co w przypadku testowanych uprzednio czystych, niezakłóconych sygnałów mowy. Kolejnym krokiem w doskonaleniu systemu mogłoby być przeprowadzenie analogicznych testów dla większego poziomu zaszumienia. Jednak uzyskane przez nas wyniki wydają się być obiecujące.

Wyrażamy zgodę na dołączenie naszych nagrań do korpusu mowy AGH. Nagrania mogą być odtwarzane, ale wyłącznie bez podawania tożsamości mówców (np. w celu prezentacji jakości, rodzaju nagrań itd.).