



**INNOWACYJNA
GOSPODARKA**
NARODOWA STRATEGIA SPÓJNOŚCI

UNIA EUROPEJSKA
EUROPEJSKI FUNDUSZ
ROZWOJU REGIONALNEGO



Stymulator uwagi słuchowej i wzrokowej



1. Wprowadzenie

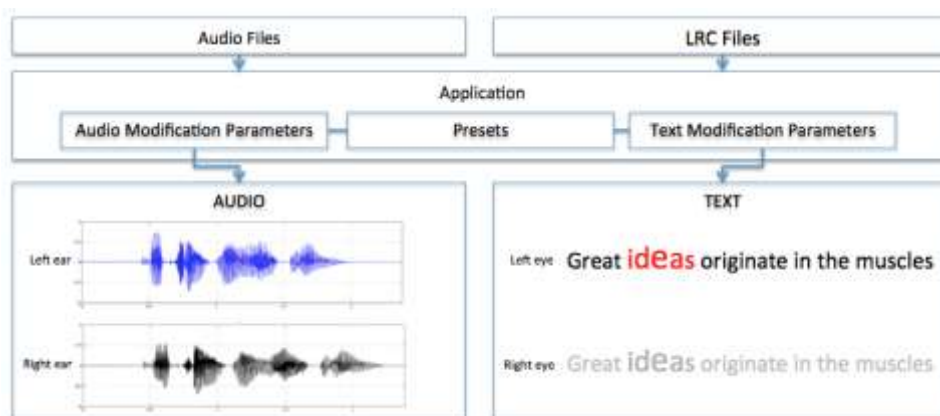
Opisywana metoda polega na zastosowaniu równoczesnej stymulacji wzrokowej i słuchowej, w taki sposób, że tekst wyświetlany na ekranie monitora komputerowego jest modyfikowany synchronicznie z odtwarzanym w słuchawkach głosem lektora. Opracowana metoda pozwala na stymulację wzroku poprzez możliwość modyfikacji koloru, jasności, wielkości i kroju czcionki.

Tempo wypowiedzi może być płynnie zmieniane z użyciem opracowanej metody modyfikacji czasu trwania mowy. Metoda ta jest połączeniem następujących algorytmów: detekcji mowy, detekcji samogłosek, detekcji zająknięć, detekcji transjentów, estymacji tempa wypowiedzi, oraz algorytmu TSM (ang. Time Scale Modification).

Każdy z wykorzystywanych plików dźwiękowych jest w specjalny sposób opisany (z użyciem plików LRC), tak by możliwe było w pełni synchroniczne odtwarzanie dźwięku wraz z modyfikacją tekstu na ekranie. Synchronizacja może odbywać się na poziomie zdań, słów, sylab, czy nawet fonemów (typowo: na poziomie słów).

Użytkownik może sterować współczynnikiem transpozycji czasowej mowy. Po zmianie współczynnika (w trakcie pracy algorytmu) konieczne jest przeprowadzenie procesu adaptacji. Proces adaptacji jest sygnalizowany wizualnie osobie prowadzącej trening. Modyfikacja obrazu jest zsynchronizowana z sygnałem fonicznym.

Ogólny schemat działania systemu został przedstawiony na rys. 1.



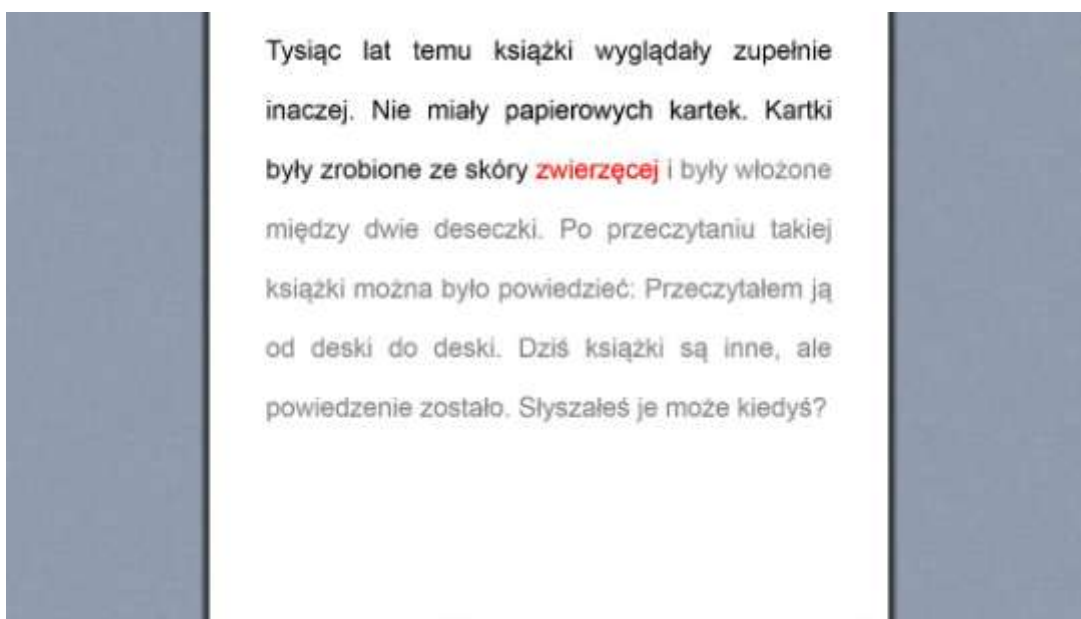
Rys. 1. Schemat działania systemu.

2. Obsługa

Aplikacja stymulatora uwagi słuchowej i wzrokowej składa się z dwóch okien:

- okna głównego, służącego do sterowania,
- okna pomocniczego, służącego do wyświetlania tekstu (rys. 2).

Oprogramowanie jest przystosowane do pracy na dwóch monitorach komputerowych.



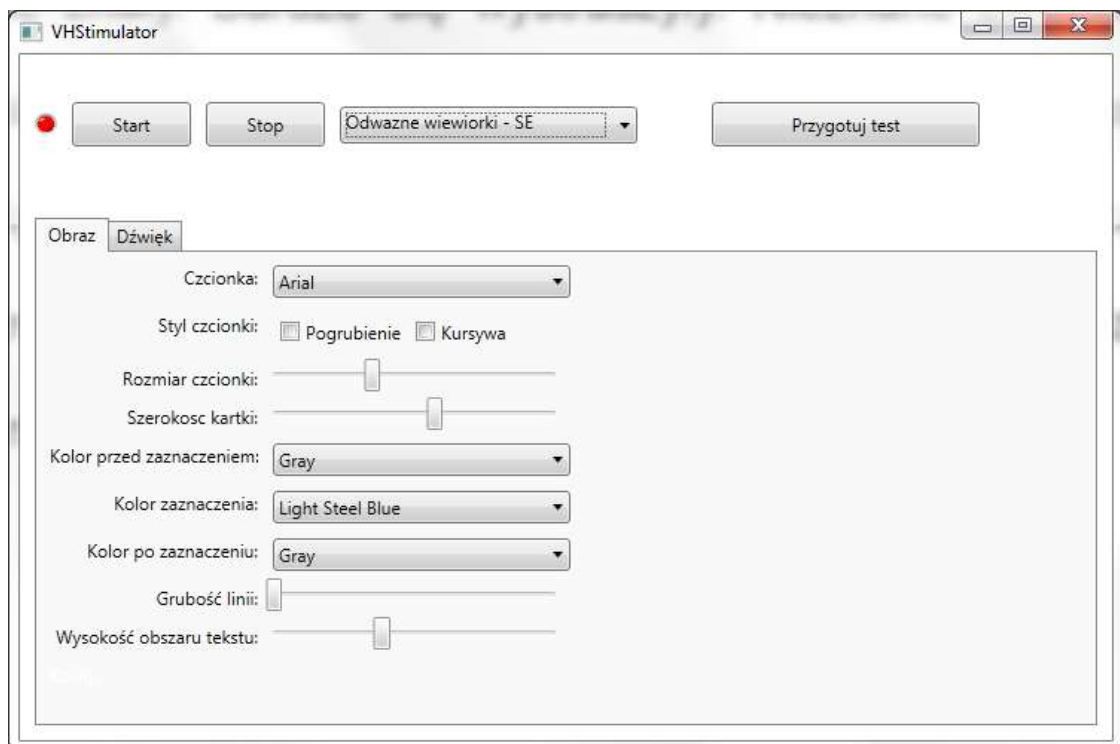
Rys. 2. Okno programu z wyświetlonym tekstem i zaznaczonym słowem (odczytywanym w danej chwili przez lektora)

2.1. Parametry obrazu

Z poziomu okna głównego w zakładce Obraz (rys. 3) można sterować parametrami obrazu (UWAGA: Parametry obrazu mogą być zmieniane podczas odtwarzania dźwięku):

- krój czcionki,
- styl czcionki,
- rozmiar czcionki,
- szerokość kartki (szerokość obszaru na którym jest wyświetlany tekst),

- kolor przed zaznaczeniem (kolor wyświetlanego tekstu, zanim zostanie on przeczytany przez lektora – ustawienie białego koloru spowoduje że tekst nie będzie widoczny przed przeczytaniem),
- kolor zaznaczenia (kolor jaki zostanie użyty do zaznaczenia czytanego tekstu),
- kolor po zaznaczeniu (kolor jaki zostanie użyty dla tekstu po przeczytaniu – ustawienie białego koloru spowoduje że tekst nie będzie widoczny po przeczytaniu).



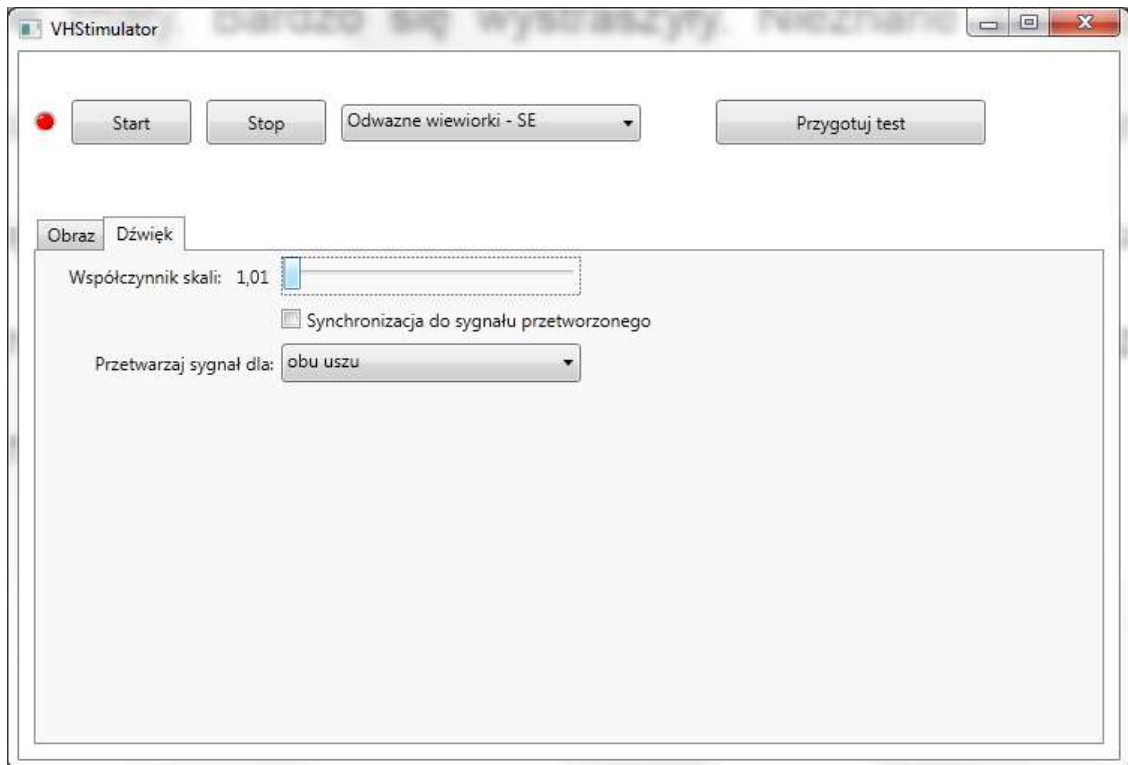
Rys. 3. Okno główne aplikacji – parametry obrazu

2.2. Parametry dźwięku

Z poziomu okna głównego w zakładce Dźwięk (rys. 4) można sterować parametrami dźwięku (UWAGA: Tylko parametr Współczynnik skali może być zmieniany podczas odtwarzania dźwięku, zmiana pozostałych parametrów podczas odtwarzania nie przyniesie efektu ich ustawienie zostanie uwzględnione po zatrzymaniu i ponownym uruchomieniu przetwarzania):

- współczynnik skali,

- włączenie synchronizacji podświetlania tekstu do sygnału przetworzonego lub oryginalnego,
- wybór ucha dla którego będzie przetwarzany dźwięk:
 - lewe,
 - prawe,
 - oba.



Rys. 4. Okno główne aplikacji – parametry dźwięku

2.3. Odtwarzanie dźwięków

W górnym obszarze okna głównego aplikacji (rys. 3, 4) znajduje się lista dostępnych utworów testowych. Przyciski **Start** i **Stop** służą do testowego odtwarzania wybranego tekstu. Uwaga: w tym trybie aplikacja zachowuje pełną funkcjonalność (za wyjątkiem funkcji testera rozumienia mowy) – uwzględniane są wszystkie parametry przetwarzania dźwięku i obrazu, a przetworzony obraz wyświetlany jest w oknie pomocniczym aplikacji.

UWAGA: W trakcie odtwarzania dźwięku nie należy zmieniać utworu.

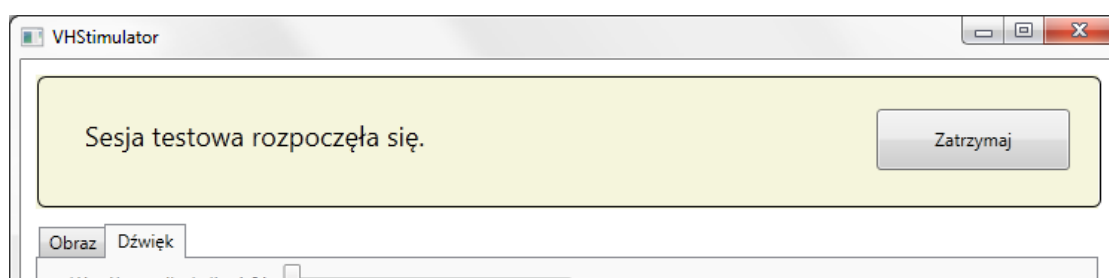
2.4. Przygotowanie i przeprowadzenie testu

Aplikacja stymulatora udostępnia także funkcjonalność prowadzenia prostych testów rozumienia mowy, polegających na zaznaczaniu różnic między tekstem czytany przez lektora, a wyświetlanym na ekranie.

Po wybraniu tekstu (z listy rozwijanej) oraz ustawieniu wszystkich parametrów przetwarzania dźwięku i obrazu można rozpocząć przygotowanie testu poprzez kliknięcie przycisku **Przygotuj test**. Przycisk zlokalizowany jest w górnej części okna głównego aplikacji (rys. 3, 4). W następnym kroku jest dostępna krótka ankieta służąca do zebrania podstawowych informacji o osobie badanej. Po wypełnieniu ankiety należy kliknąć przycisk **Dalej** - pojawi się napis *Sesja testowa rozpoczęła się* (rys. 5). Możliwe jest wcześniejsze zatrzymanie sesji testowej poprzez kliknięcie przycisku **Zatrzymaj** (rys. 6).



Rys. 5. Okno główne aplikacji – ankieta testu



Rys. 6. Okno główne aplikacji – rozpoczęcie testu

Po rozpoczęciu sesji testowej na ekranie pomocniczym (dla osoby badanej) jest wyświetlany komunikat: **Kliknij przycisk SPACJA aby rozpocząć test**. Po kliknięciu przycisku spacja test zostanie uruchomiony (rys. 7). Zadaniem osoby badanej jest wychwycenie różnic pomiędzy tekstem wyświetlonym na

ekranie, a tekstem czytany przez lektora, po znalezieniu różnicy osoba badana ma za zadanie nacisnąć spację. Aplikacja automatycznie zlicza wszystkie kliknięcia użytkownika i po zakończeniu badania zapisuje je w pliku *results.txt* znajdującym się w folderze aplikacji.



Rys. 7. Okno pomocnicze aplikacji – dedykowane dla osoby badanej.

Po zakończeniu badania nie jest wyświetlany żaden wynik. Potrzebna jest dodatkowa analiza zapisanych rezultatów badania.

3. Format plików

W celu synchronizacji odtwarzania plików dźwiękowych z podświetlaniem odpowiednich fragmentów tekstu konieczne jest zaindeksowanie plików zgodnie z formatem:

kolejny_element;czas_początkowy;czas_końcowy;podświetlany_tekst

gdzie:

kolejny_element - kolejny indeks podświetlanego tekstu, począwszy od liczby 1

czas_początkowy - czas rozpoczęcia wypowiedzianego przez lektora danego tekstu; format: *minuty.sekundy.tysięczne_sekundy*

czas_końcowy - czas zakończenia wypowiedzenia tekstu przez lektora; format:
minuty.sekundy.tysiężne_sekundy

podświetlany_tekst - tekst, który ma zostać wyróżniony, jeżeli złożony z więcej niż jednego wyrazu (lub spójnika i wyrazu) konieczne wstawienie znaku o kodzie ASCII 255.

Plik musi mieć rozszerzenie *.sym.

Zawartość przykładowego pliku *.sym przedstawiono poniżej.

```
1;0:00.133;0:00.413;Ala
2;0:00.413;0:00.560;ma
3;0:00.560;0:00.986;małego
4;0:00.986;0:01.480;kotka.
5;0:01.840;0:02.160;Kotek
6;0:02.160;0:02.386;Alę
7;0:02.386;0:02.920;w parku
8;0:02.920;0:03.573;spotkał.
```