

Daniel Śledziński

Etapy oraz koncepcja automatyzacji badań nad percepcją mowy

1. Wstęp

W niniejszym artykule omówiono etapy badań nad percepcją mowy oraz przedstawiono koncepcję automatyzacji takich badań. Koncepcja ta jest zrealizowana w rozwiązaniu technicznym obejmującym trzy programy – program przeznaczony do projektowania testów percepcyjnych, program do przeprowadzania testów oraz program do analizy wyników badań.

Idea oprogramowania ułatwiającego i przyspieszającego badania percepcji mowy powstała podczas wcześniejszych badań prowadzonych przez autora tekstu. Te doświadczenia pozwoliły na wypracowanie procesu obejmującego kilka etapów. W artykule przeanalizowano rolę wymienionych programów oraz problemy napotymane na poszczególnych etapach badań.

Przedstawione w artykule rozwiązania nie są ostateczne – przewidywany jest dalszy ich rozwój – głównie przez wprowadzenie nowych funkcji w wymienionych programach oraz usprawnianie ich działania. W czasie przygotowywania artykułu istniały prototypowe wersje omawianych programów. Planowane jest nieodpłatne udostępnienie pakietu programów w tworzonym serwisie internetowym poświęconym tego typu oprogramowaniu (serwis będzie uruchomiony pod adresem: speechtools.org).

2. Badania wcześniejsze i literatura

Idea oprogramowania ułatwiającego i przyspieszającego badania percepcji mowy powstała podczas wcześniejszych eksperymentów

wykonywanych przez autora artykułu. Jedno z badań polegało na wygenerowaniu syntetycznych połączeń samogłosek oraz spółgłosek (miejsce łączenia pokrywało się z granicą ulokowaną między ośrodkiem sylaby otwartej a nagłosem następnej sylaby). W teście percepcyjnym oceniana była jakość połączeń syntetycznych oraz połączeń niezmienionych. Inne wykonane przez autora badanie polegało na identyfikacji sylab typu CV (ang. Consonant–Vowel), które zostały pobrane z wyrazów nagranych w izolacji.

Istnieje wiele publikacji poświęconych percepcji mowy. Usystematyzowany przekrój aktualnych badań z tej dziedziny można odnaleźć w *The Handbook of Speech Perception* (D.B. Pisoni, R.E. Remez 2008). Jest to obszerna publikacja, która ujmuje tę tematykę w sposób bardzo wnikliwy – liczba omawianych badań i przytaczanych autorów jest imponująca.

Inną wartą polecenia publikacją jest podręcznik *Psycholingwistyka* (J.B. Gleason, N.B. Rathner 2009), w którym rozdział trzeci poświęcono percepcji mowy. Przedstawiono tam wiele istotnych zagadnień, takich jak: percepcja samogłosek i spółgłosek, czas rozpoczęcia dźwięczności (ang. Voice Onset Time – VOT), wpływ tempa mowy na percepcję czy modele percepcji mowy.

Niezmiernie wartościową pozycją jest monografia pod redakcją Andrzeja Obrębowskiego *Wybrane zagadnienia z audiometrii mowy* wydana w 2011 roku. W publikacji tej przedstawiono między innymi problemy związane z metodologią badań, strukturą testów czy możliwościami wykorzystania audiometrii mowy w orzecznictwie lekarskim.

3. Etapy badań nad percepcją mowy

Badania percepcji mowy obejmują kilka etapów. Nakład pracy związany z poszczególnymi etapami jest uzależniony od konkretnego zadania badawczego. Również zastosowane rozwiązania techniczne mają wpływ na organizację pracy. Poniżej przedstawiono propozycję etapów badań uwzględniających użycie opisywanego w artykule oprogramowania:

- planowanie badań,
- przygotowania materiału badawczego,
- przygotowanie testu percepcyjnego,
- przeprowadzenie testu,
- analiza wyników oraz wnioski.

Etapy te – ich rolę w procesie badawczym, napotykanne problemy oraz sugestie praktycznych rozwiązań – omówiono w rozdziale czwartym wspomnianego podręcznika. Przedstawione rozwiązania zostały w dużej mierze oparte na wcześniejszym doświadczeniu autora – zatem nie mają charakteru ścisłych reguł czy norm – mogą natomiast podlegać dyskusji oraz dalszym modyfikacjom.

W rozdziale drugim omawianej monografii wymieniono wytyczne dotyczące metodologii badania słuchu. Poniższa lista pochodzi z tej publikacji i określa czynności, jakie należy wykonać, przystępując do badania:

- określić wymagania, jakie powinien spełniać nagrany materiał słowny,
- określić poziom nagrania (sygnału mowy) i poziom maskowania (jeśli jest to konieczne),
- określić poziomy ciśnienia akustycznego hałasu tła w pomieszczeniu pomiarowym, tak przy badaniu przez słuchawki, jak i przez głośnik,
- przygotować i udzielić instrukcji osobie badanej,
- ustalić sposób odpowiadania osoby badanej,
- wyznaczyć progowe poziomy detekcji mowy (próg wykrywania i rozpoznawania mowy) bez i z dźwiękiem zakłócającym (jeśli jest to konieczne),
- ustalić poziom maskowania w uchu lepiej słyszającym przy badaniu ucha gorszego,
- wyznaczyć i wykreślić postać audiogramu słownego.

Przytoczone punkty oparte są na normach europejskich oraz normie opracowanej przez Polski Komitet Normalizacyjny PN-EN ISO 88253-3 z grudnia 2005 roku pt. *Akustyka. Metodyka pomiarów audiometrycznych, część 3, Audiometria słowna*.

3.1. Planowanie badań

Podjęcie właściwych decyzji na tym etapie ma zasadniczy wpływ na powodzenie eksperymentu. Z kolei decyzje błędne mogą sprawić, że eksperyment będzie nieudany, a uzyskane wyniki nie będą miały istotnej wartości naukowej. Planując badania, należy przeanalizować aktualny stan wiedzy dotyczący podejmowanego zagadnienia badawczego. Istotna jest odpowiedź na pytanie, czy istnieją przesłanki uzasadniające podjęcie danego zadania badawczego lub podjęcie tego zadania dla konkretnego języka.

Kolejnym problem, który trzeba uwzględnić już na etapie planowania badań, jest związany z możliwościami technicznymi, jakimi dysponuje autor badań lub zespół badawczy. W niniejszym artykule omówiono rozwiązanie, którego celem jest ułatwienie badań percepcji audytywnej w szczególności osobom, które nie dysponują zaawansowanymi możliwościami i umiejętnościami technicznymi (na przykład umiejętnością tworzenia programów komputerowych).

Istotny problem techniczny w obszarze badań percepcji audytywnej dotyczy sposobu przeprowadzania testu percepcyjnego – rozwiązania, które pozwala na odtwarzanie kolejnych nagrań oraz rejestrowane odpowiedzi. Najprostsze podejście nie wymaga angażowania zaawansowanych rozwiązań – wystarczy dysponować zbiorem plików dźwiękowych zapisanych w folderze i ułożonych w kolejności zgodnej z wykazem, do którego wgląd ma osoba badana. Takie rozwiązanie ma jednak istotną wadę – nie jest ono wygodne dla osoby badanej, zatem mogą wystąpić trudności z koncentracją na treści testu. Poza tym podejście to narzuca pewne ograniczenia dotyczące konstrukcji testu oraz analizy wyników. Inne rozwiązanie polega na utworzeniu specjalnej aplikacji przeznaczonej do realizacji konkretnej koncepcji badawczej. Takie podejście umożliwia utworzenie wygodnego interfejsu dostosowanego do specyfiki określonego zadania badawczego, jednak wymaga zatrudnienia programisty. Celem omawianych w artykule narzędzi jest połączenie zalet dwóch wymienionych rozwiązań – łatwej dostępności oraz wysokiej jakości badań.

Na etapie projektowania badań trzeba też zastanowić się, czy realizacja danego zadania badawczego jest w ogóle możliwa. Problem ten w szczególności dotyczy badań nad percepcją mowy, ponieważ dostęp

do procesów zachodzących w ludzkim mózgu jest bardzo ograniczony. Uzyskane wyniki można porównywać z dostępnymi teoriami dotyczącymi percepcji mowy¹.

Mówiąc o możliwościach realizacji zadań badawczych związanych z percepcją mowy, trzeba pamiętać o dużej liczbie niektórych jednostek lub trudnościach związanych z ich wyznaczeniem. Jako przykład można podać sylaby – badacz napotyka problem dużej liczby tych jednostek (są ich tysiące) oraz problem z wyznaczeniem ich granic – trudność ta w szczególności dotyczy języka polskiego ze względu na występujące powszechnie wieloelementowe grupy spółgłosek o niespotykanej w innych językach strukturze². Wymienione powody sprawiają, że niektóre problemy badawcze mogą być rozpatrywane tylko fragmentarycznie.

3.2. Przygotowanie materiału badawczego

Po sformułowaniu problemu badawczego można przystąpić do przygotowania materiału badawczego. Badania percepcji audytywnej polegają na prezentowaniu kolejnych bodźców dźwiękowych oraz rejestrowaniu reakcji osoby badanej. Materiał badawczy obejmuje zatem:

- zapis tekstowy bodźców wraz z danymi towarzyszącymi,
- zapis dźwiękowy bodźców (zbiór plików dźwiękowych).

Opracowanie listy bodźców (zawierającej treści, które będą odtwarzane w teście percepcyjnym) nie jest zadaniem łatwym. Przygotowany materiał badawczy powinien być adekwatny do przyjętych założeń badawczych. Jeżeli potencjalna liczba bodźców jest zbyt duża, to badacz zmuszony jest do ograniczenia zakresu badań do rozmiaru moż-

¹ Zob. J. Elman, J.L. McClelland, *The interactive activation model of speech perception. Language and speech*, New York 1984; D. Massaro, *Testing between TRACE model and fuzzy logical model of speech perception*, „Cognitive Psychology” 1989, 21 (3), s. 398–421.

² Zob. J. Szpyra-Kozłowska, *Wprowadzenie do współczesnej fonologii*, Lublin 2002, s. 147–155; H. Rogers, *The Sounds of Language. An Introduction to Phonetics*, London 2000, s. 89; P. Ladefoged, *A course in phonetics*, London 1975, s. 217.

liwego do zrealizowania. Dla każdego bodźca trzeba opracować zestaw danych niezbędnych do realizacji testu – na przykład informacje, które będą pojawiały się na ekranie komputera w czasie odtwarzania plików dźwiękowych lub etykiety przycisków pojawiających się po odtworzeniu bodźca.

Dysponując gotową listą bodźców, można przystąpić do nagrywania mowy. Czasami generuje się mowę syntetyczną dla badań percepcyjnych (zagadnienia związane z syntezą mowy nie są poruszane w tym artykule)³. Oto pytania, na które należy odpowiedzieć przed rozpoczęciem nagrań mowy:

- Czy nagrywana będzie mowa ciągła (czytany tekst), czy wyrazy izolowane lub mniejsze jednostki?
- Ile osób będzie uczestniczyło w nagraniach?
- Jaka będzie płeć osób nagrywanych?
- Jaki będzie wiek osób nagrywanych?
- Czy istnieją dodatkowe wymagania dotyczące osób nagrywanych?
- Czy istnieją dodatkowe wymagania dotyczące warunków, w których odbędą się nagrania?

Często nagrywane są fragmenty, które obejmują wiele wyrazów, dlatego wymagają dalszej obróbki – muszą być pocięte (rozdzielone na pojedyncze bodźce) i umieszczone w oddzielnych plikach dźwiękowych. Umieszczenie bodźców w oddzielnych plikach daje możliwość między innymi randomizacji bodźców. Przed rozdzieleniem na mniejsze pliki często wykonuje się inne czynności – na przykład normalizację natężenia dźwięku lub konwersję podstawowych parametrów dźwięku cyfrowego – liczby kanałów, częstotliwości próbkowania lub rozmiaru próbek dźwięku⁴.

³ R.S. Cooper [et al.], *Some experiments on the perception of synthetic speech sound*, „Journal of the Acoustical Society of America” 1952, 24 (6), s. 597–606.

⁴ Zob. W. Burty, *Dźwięk cyfrowy, systemy wielokanałowe*, Warszawa 2001, s. 9–43; R. Tadeusiewicz, *Sygnal mowy*, Warszawa 1988, s. 99–193.

3.3. Przygotowanie testu percepcyjnego

Dysponując zbiorem plików dźwiękowych oraz niezbędnymi danymi, można przystąpić do tworzenia testu. Omawiany w artykule pakiet programów obejmuje program, który umożliwia realizację tego zadania. Wynikiem działania tego programu jest plik zawierający wszystkie potrzebne dane dotyczące przebiegu testu. Oto lista najważniejszych funkcji programu:

- wczytywanie zbiorów plików dźwiękowych,
- ustalanie struktury bodźca – liczby powtórzeń nagrania oraz przerwy między kolejnymi odtworzeniami nagrania (liczonej w milisekundach),
- ustalanie struktury odpowiedzi – wprowadzono podział na odpowiedzi jawne (wyświetlane w przyciskach) oraz odpowiedzi niejawne (wpisywane do pól tekstowych przy użyciu klawiatury),
- ustalenie dodatkowych pól z danymi – istnieje możliwość, aby do każdego bodźca dołączyć dane, które nie są wykorzystywane w teście, jednak mogą być użyte na etapie analizy wyników,
- randomizacja bodźców,
- zapisywanie do pliku oraz odczytywanie z pliku wersji roboczej testu.

3.4. Przeprowadzenie testu percepcyjnego

Kolejny etap to przeprowadzenie właściwego testu. Liczba osób badanych jest uzależniona od przyjętych założeń. Zaangażowanie zbyt małej liczby osób może sprawić, że wyniki będą niewiarygodne. Inny problem związany jest z tym, że pojedynczy test percepcyjny nie powinien być zbyt obszerny (nie powinien zawierać zbyt wielu bodźców), ponieważ może to spowodować znużenie lub zmęczenie słuchu u osób badanych oraz wpłynąć negatywnie na jakość wyników. Istotne jest również miejsce badania, które musi spełniać określone wa-

runki akustyczne – poziom hałasu nie powinien przekraczać 40 dB, a czas pogłosu powinien być mniejszy niż 0,5 s⁵.

Omawiany w artykule pakiet programów obejmuje program przeznaczony do przeprowadzania testów percepcyjnych – jest to jedyny program, z którym styczność mają osoby biorące udział w eksperymencie. Program przeprowadza test zgodny z informacjami zawartymi w pliku wygenerowanym przez program opisany w podrozdziale 3.3. Przed rozpoczęciem testu trzeba wskazać folder przechowujący wszystkie niezbędne pliki dźwiękowe. Po zakończeniu testu odpowiedzi zapisywane są do pliku tekstowego w postaci tabelarycznej (jako separator kolumn używany jest znak tabulacji). Jeżeli zaprojektowany test obejmuje dodatkowe pola z danymi, to te dane również zostają zapisane do pliku z wynikami.

3.5. Analiza wyników badań

Analiza wyników to ostatni etap badań. Rozwijana jest koncepcja programu ułatwiającego i przyspieszającego analizę wyników uzyskanych z testów percepcyjnych. Poniżej przedstawiono główne funkcje dotyczące programu:

- jednoczesne odczytywanie wielu wyników (pochodzących z badań przeprowadzanych na różnych osobach),
- grupowanie odpowiedzi na podstawie dodatkowych informacji zapisanych w polach z danymi,
- tworzenie rozkładów pojedynczych odpowiedzi lub rozkładów dla grup odpowiedzi,
- zapisywanie wyników do pliku.

4. Podsumowanie

W artykule omówiono najważniejsze zadania związane z badaniami percepcji mowy oraz przedstawiono koncepcję automatyzacji takich badań. Przedstawione informacje w dużej mierze oparto na doświad-

⁵ Zob. A. Obrębowski, *Wybrane zagadnienia z audiometrii mowy*.

czeniu autora tekstu, zatem nie mają one charakteru reguł czy norm – zaprezentowane koncepcje mogą podlegać dyskusji oraz dalszym modyfikacjom. Wydaje się jednak, że już na obecnym etapie udało się zrealizować najważniejsze założenia dotyczące technicznego systemu przeznaczonego do wspomagania badań percepcji audytywnej. Powstałe programy powinny uprościć i przyspieszyć procedurę badawczą.

W artykule omówiono również szereg potencjalnych problemów występujących na kolejnych etapach badań. Wiedza ta może być pomocna przy projektowaniu badań, a także ułatwić uniknięcia błędów wpływających negatywnie na jakość wyników.

W czasie pisania artykułu istniały prototypowe wersje omawianych w nim programów. Następne prace będą związane z testowaniem, usprawnianiem, a także wprowadzaniem nowych funkcji, na przykład możliwości regulowania natężenia odtwarzanych dźwięków, wykrywania poziomu progowego detekcji, czy też tworzenia audiogramów. Planowane jest nieodpłatne udostępnienie pakietu w aktualnie tworzonej stronie internetowej speechtools.org.

Stages and automatisisation of speech perception investigations

SUMMARY

In the article a process of investigations of the speech perception was described. A conception of automatisisation of the process was introduced. That conception was realized in a package of software developed by the author. The package includes three programs. The first program is purposed for perceptual tests building. The second program is aimed to execute tests. The third program is intended for data analysis but it is not ready at the moment. The ideas showed in the article arose during previous investigations realized by the author. The presented solutions are not finished and they are intended to be developed in the future. All programs will be published as freeware available at the currently developed Internet service: speechtools.org.

O Autorze

Daniel Śledziński – doktor nauk humanistycznych,
Instytut Językoznawstwa, Uniwersytet im. Adama
Mickiewicza w Poznaniu. Zainteresowania: przetwa-
rzanie i analiza tekstów oraz sygnału mowy, percep-
cja mowy, fonetyka akustyczna, fonologia, języki
programowania, tworzenie aplikacji, bazy danych,
sztuczne sieci neuronowe, statystyka.
E-mail: danielsl@poczta.onet.pl