

PRZETWARZANIE INFORMACJI JĘZYKOWEJ. PODSTAWY KOGNITYWNE

Zanim przejdziemy do kwestii technicznych, przyjrzyjmy się człowiekowi, a konkretniej jego zdolnościom i możliwościom percepcyjnym oraz lingwistycznym, abyśmy wrażliwi na te zagadnienia, pamiętali o nich i wykorzystywali wiedzę o procesach poznawczych użytkownika do tworzenia przyjaznego dla niego interfejsu.

Ten rozdział składa się z kilku części, w których omówiono wybrane zagadnienia dotyczące: percepcji człowieka (percepcja wzrokowa i słuchowa), uwagi jako procesu kontrolnego oraz, ze względu na fundamentalną rolę języka naturalnego jako narzędzia poznania i komunikacji, przetwarzania informacji językowej. Każdy z tych tematów jest bardzo obszerny i zasługuje na to, by szczegółowo go omówić i rozwinąć. Niestety, ze względu na charakter książki problemy te zostaną potraktowane skrótowo, jednak z uwzględnieniem najistotniejszych potrzebnych informacji, które zdaniem autorów są niezbędne, by pomóc twórcy interfejsu w pracy nad budową przyjaznego dla użytkownika interfejsu.

Zatem zaczniemy od percepcji, bo w ten sposób człowiek, a nim przecież jest użytkownik interfejsu, odbiera bodźce pochodzące ze świata zewnętrznego. Nie będziemy omawiać funkcjonowania wszystkich naszych zmysłów, lecz zajmiemy się krótko wybranymi zagadnieniami z percepcji wzrokowej, która jest głównym narzędziem ludzkiej percepcji, oraz szerzej wybranymi problemami z percepcji słuchowej (w tym percepcji i przetwarzania mowy), która wspomaga percepcję wzrokową, a w przypadku osób z dysfunkcją wzroku musi ją zastąpić.

Pokażemy też, jak procesy percepcyjne są powiązane z funkcjonowaniem uwagi, która odgrywa kluczową rolę w świadomym odbieraniu i przetwarzaniu

informacji. Na koniec skupimy się na języku, który jest skomplikowanym narzędziem komunikacji, a zarazem ważnym narzędziem poznawania świata. Nazywając to, co nas otacza, a więc obiekty, czynności, zjawiska, emocje itd., nazywamy wynik poznania, a zarazem tworzymy i zapamiętujemy znaczenia, które przyjmują formę wyrazów. Jak wiemy, znaczenia są przechowywane w ludzkim umyśle w sposób uporządkowany, są z sobą powiązane (np. tematycznie). Podobieństwo struktur semantycznych przechowywanych w umyśle pozwala ludziom porozumiewać się za pomocą języka. Doświadczenia nad badaniem swobodnych skojarzeń słownych dają możliwość poznania struktury powiązań semantycznych przechowywanych w ludzkim umyśle¹. Wiedza o tych powiązaniach pozwala rekonstruować sposób, w jaki człowiek odbiera i kojarzy informację. Wiedza o naturalnych powiązaniach znaczeń ułatwi także budowę tzw. intuicyjnego interfejsu użytkownika.

1.1. PERCEPCJA WZROKOWA

Percepcja wzrokowa stanowi najważniejsze narzędzie odbioru informacji przez człowieka. Większość informacji dociera do nas tą drogą. Dzięki niej odbieramy świat wieloaspektowo: kolory, wielkość, kształt, odległość, perspektywa, głębia, orientacja obiektów w przestrzeni i pomiędzy sobą, organizacja figura–tło, ruch – to cechy obiektów dostrzegane dzięki zmysłowi wzroku². Ciekawe światło na funkcjonowanie percepcji wzrokowej rzucają tzw. iluzje wzrokowe. Poniekąd demaskują one to, jak przebiega proces percepcji wzrokowej. W historii badań nad ową percepcją nie sposób nie wspomnieć o szkole Gestalt i prawach, które zostały sformułowane na podstawie badań eksperymentalnych³. Warto podkreślić, że narząd wzroku, a więc para oczu, stanowi tylko narzędzie do odbioru wrażeń wzrokowych. Informacja o cechach obiektu zostaje przekazana w formie impulsów nerwowych drogami

¹ Opis autorskiego badania swobodnych skojarzeń słownych znajduje się w końcowej części tego rozdziału.

² Każdy z tych aspektów percepcji wzrokowej wymaga osobnego omówienia, jednak ze względu na specyfikę rozdziału omówienie zostanie pominięte.

³ S. Palmer, *Modern Theories of Gestalt Perception*, w: *Understanding Vision: An Interdisciplinary Perspective Readings in Mind and Language*, red. G.W. Humphreys, Blackwell, Oxford 1992, s. 39–70; I. Rock, S. Palmer, *The Legacy of Gestalt Psychology*, „Scientific American”, December, s. 84–90.

nerwowymi przez wzgórce do mózgu. Tam dopiero następuje proces dekodowania cech bodźca, a następnie scalania i interpretacji postrzeganego obrazu.

Percepcja wzrokowa powoduje, że możemy czytać tekst: artykuły, prasę, książki. Nasz wzrok podąża litera po literze, odczytuje wyrazy i zdania, postrzega znaki graficzne, które pomagają nam zrozumieć tekst: pytajnik, wykrzyknik, pauzę, przecinek, wielokropek czy po prostu kropkę. Czytając nawet bardzo skomplikowany tekst, potrafimy wczytać się w niego, by zrozumieć treść, powrócić wzrokiem do czytanej frazy, przerzucić wzrok na koniec zdania, akapitu bądź strony. I mimo że mamy do czynienia z tekstem językowym, który ze swej istoty jest linearny (elementy występują w nim po kolei), to możemy przerzucać wzrok w różne jego części. Możemy też za pomocą odpowiednich konstrukcji językowych opisać nawet najbardziej skomplikowane zagadnienia, a posługując się językowymi zwrotami określającymi orientację przestrzenną i organizację czasową, wyobrazić sobie daną sytuację wielowymiarowo. Tak więc język ludzki jest linearny jedynie w wypowiedzi: w języku pisanym litera następuje po literze, zaś w wypowiedzi mówionej głoski realizujemy jedna po drugiej.

Gdy przeglądamy strony internetowe, bez problemu możemy przerzucić wzrok do różnych części ekranu, by w poszukiwaniu interesujących nas informacji szybko przejrzeć różne wątki i elementy strony, znaleźć poszukiwane dane, które następnie nasz wzrok nie tyle czyta (jak drukowany tekst), ile skanuje. To nowa metoda, którą rozwinął i posługuje się ludzki system wzrokowy do korzystania z nowego źródła informacji, jaką jest Internet⁴. Jednocześnie ludzki system wzrokowy nie tylko czyta, a właściwie skanuje znajdujący się na ekranie tekst, ale też ogląda grafikę, rysunki, ilustracje i filmy, które powinny uzupełnić i dopełnić informacje językowe.

1.1.1. Kompensacja percepcji wzrokowej u osób z dysfunkcją wzroku

Nieco odmiennie wygląda sytuacja osoby z dysfunkcją wzroku. Dla niej najważniejsza jest percepcja słuchowa i to ona stanowi główne źródło informacji o świecie. Dzięki dźwiękom taka osoba buduje swój świat, ocenia odległość,

⁴ Ch. Baumgärtner, W. Menzel, *Integrating a Model for Visual Attention into a System for Natural Language Parsing*, w: *Natural Language Processing and Cognitive Science*, red. B. Sharp, M. Zock, 9th International Workshop on Natural Language Processing and Cognitive Science, Wrocław 2012, s. 24–33.

przestrzeń, a percepcja mowy jest głównym źródłem informacji językowej, choć nie jedynym – wykorzystywany jest też zmysł dotyku, gdy osoba z dysfunkcją wzroku czyta, posługując się alfabetem Braille’a.

Pozostałe zmysły, a więc dotyk, zapach, smak, dopełniają obrazu świata u osób widzących, natomiast u osób z dysfunkcją wzroku są w miarę możliwości intensywnie rozwijane, by wspomóc percepcję słuchową i w możliwie najpełniejszy sposób pozwolić im poruszać się po otaczającym świecie.

U osób z dysfunkcją wzroku inne zmysły odgrywają istotną rolę, w szczególności zmysł słuchu – dzięki niemu osoby te konstruują swój świat, oraz zmysł dotyku, który pełni ważną funkcję zwłaszcza w początkowym okresie życia, kiedy to wpływa na rozwój emocjonalny, jak i fizyczny. Podobnie dzieje się u osób widzących, gdzie zmysł dotyku dopełnia całości obrazu świata. Zwłaszcza w początkowym okresie życia, gdy dziecko nawiązuje kontakt z matką, zmysł dotyku pomaga nawiązać intymną więź. Także w rozwoju fizycznym dotyk, we współpracy z pozostałymi zmysłami, pozwala rozwinąć się integracji ruchów kończyn, położenia ciała oraz świadomego poruszania się. Wraz z rozwojem osobniczym zmysł dotyku podlega specjalizacji, która polega na umiejętności oceny cech przedmiotów, czyli faktury, kształtu, konsystencji, twardości, temperatury, ciężaru⁵. Jeśli zmysł dotyku jest wcześniej kształtowany i trenowany, to pozwala posługiwać się alfabetem Braille’a podczas czytania i pisania, umożliwia również tworzenie i rozpoznawanie wypukłych rysunków, dając tym samym szansę na rozwój wiedzy, ale też na zabawę.

Innym istotnym dla konstruowania i poruszania się w świecie zmysłem u osób z dysfunkcją wzroku jest zmysł powonienia oraz związany z nim zmysł smaku. Osoby widzące także posługują się tymi zmysłami, jednakże traktują je jako zmysły wspomagające dla percepcji wzrokowej. U niewidomych szczególna rola zmysłów powonienia i smaku wynika z potrzeby konstruowania możliwie najbardziej pełnego obrazu świata z wykorzystaniem informacji pochodzących z dostępnych im zmysłów.

⁵ J. Minogue, M. Jones, *Haptics in Education: Exploring an Untapped Sensory Modality*, „Review of Educational Research” 2006, nr 76 (3), s. 317–348.

1.2. PERCEPCJA SŁUCHOWA

Dla zdrowego człowieka percepcja słuchowa stanowi drugie po percepcji wzrokowej źródło informacji o otaczającym świecie. Dzięki percepcji słuchowej człowiek zyskuje jeszcze inny wymiar poznania rzeczywistości. Mając na myśli percepcję słuchową, myślimy o percepcji dźwięków oraz o percepcji mowy.

Percepcja mowy jest szczególnie rozwinięta ze względu na swoją specyfikę i pełni niezwykle istotną rolę w życiu człowieka. To zagadnienie zostanie jeszcze omówione w dalszej części tekstu. W tym miejscu, analizując dwie najważniejsze percepcje człowieka: wzrokową i słuchową, warto uświadomić sobie, że człowiek odbiera otaczający go świat, korzystając z obydwu (a nawet i większej liczby) zdolności percepcyjnych organizmu, równie cennych i dopełniających obraz postrzeganej rzeczywistości, jak dotyk, zapach czy smak. Nawet nie zauważamy, że korzystając z percepcji słuchowej, dopełniamy obraz percepcją wzrokową – np. podczas bezpośredniej rozmowy zazwyczaj przyglądamy się układowi warg i ruchom języka interlokutora, bo są to cenne informacje wspomagające percepcję słuchową. Kiedy oglądamy program na ekranie telewizora, jednocześnie słuchając, patrzymy na układ ust mówiącego, wcale się nad tym nie zastanawiając. Dopiero gdy ścieżka dźwiękowa nie nakłada się idealnie ze ścieżką obrazu, natychmiast zauważamy takie minimalne nawet przesunięcia. Zjawisko to nazywamy *efektem McGurka*⁶. Ten przykład pokazuje, jak ściśle w umyśle zdrowego człowieka powiązane są obydwie percepcje. Wszystko po to, by dać mu spójną wizję otaczającego go świata, by jak najprecyzyjniej zebrać potrzebne do tej interpretacji informacje.

U zdrowego, czyli normalnie funkcjonującego człowieka percepcja słuchowa wspomaga percepcję wzrokową. Jesteśmy praktycznie cały czas skazani na odbiór informacji dźwiękowych, albowiem nie możemy odciąć się od źródła informacji – w przeciwieństwie do percepcji wzrokowej, podczas której możemy zamknąć oczy, zacisnąć powieki, by przestać na coś patrzeć, a tym samym odciąć się od informacji wzrokowej. Z naszym narządem słuchu jest inaczej: nie jesteśmy w stanie, jak to jest w przypadku oczu, zamknąć

⁶ H. McGurk, J. MacDonald, *Hearing Lips and Seeing Voices*, „Nature” 1976, nr 264, s. 746–748.

uszu. Dlatego ludzki organizm wypracował inne metody radzenia sobie z nadmiarem informacji dźwiękowych – procesy związane z uwagą umożliwiają nam pewną selekcję doboru materiału dźwiękowego. Nie jest to jednak mechanizm doskonały. Poza tym w dużej mierze zależy od właściwości osobniczych oraz od rangi wykonywanego zadania. Potrafimy na przykład odseparować się od dźwięków otaczającej nas rzeczywistości dosłownie – blokując sobie uszy poprzez zatyczki; możemy starać się ograniczyć dopływ informacji dźwiękowych (pracując w ciszy, w nocy), by uniknąć nieustannego rozproszenia uwagi. Niektórzy opanowali w bardzo dobrym stopniu mechanizm związany z uwagą, a mianowicie tak mocno koncentrują się na pracy, że psychicznie wyłączając się, nie słyszą lub prawie nie słyszą odgłosów otoczenia. Niektórzy lubią pracować przy dźwiękach muzyki instrumentalnej bądź słuchając piosenek w języku, którego nie znają. Znajomość języka, w którym śpiewana jest piosenka, powoduje, że uwaga zostaje zaangażowana, co przyczynia się do zmniejszenia efektywności wykonywanego zadania⁷. Jak wiemy, uwaga jest niezwykle ważnym procesem, który organizuje i zarządza wykonywanymi przez człowieka zadaniami, zwłaszcza jeśli nie są to zadania zautomatyzowane. Dlatego uwadze poświęcimy nieco miejsca w dalszej części tego rozdziału.

Procesy uwagowe w kontekście percepcji słuchowej nie ograniczają się tylko do blokowania informacji dźwiękowych. Uwaga pomaga także wyłowić spośród natłoku dźwięków te informacje, które są dla człowieka istotne. Nie bez powodu dźwięki wydawane przez pojazdy uprzywilejowane, gdy jadą na sygnale, czyli karetki pogotowia, wozy strażackie, wozy policyjne, są tak przejmujące, silne i wybijające się spośród odgłosów ruchu ulicy, że pozostali uczestnicy ruchu drogowego ustępują im drogi.

W przypadku percepcji mowy, nawet w gwałtownym otoczeniu, jeżeli człowiek usłyszy swoje imię bądź wyraz znaczący istotną dla niego informację, to potrafi skoncentrować swoją uwagę na tym kanale, z którego pochodzi interesujące go źródło. Naturalnie nazwa tego efektu zwanego *cocktail party* nawiązuje do okoliczności, podczas których został on zaobserwowany, ale oczywiście nie zachodzi on tylko podczas przyjęć, lecz zawsze, gdy człowiek znajduje się w gwałtownym otoczeniu i zostały spełnione opisane wyżej

⁷ Wiemy, że pod kontrolą uwagi możemy wykonać efektywnie jedno zadanie, inne automatycznie nie zmienia jakości i czasu wykonania zadania angażującego uwagę. Jeśli wykonujemy dwa zadania z zaangażowaniem uwagi, wówczas wykonujemy je wolniej bądź mniej dokładnie.

warunki. O tym zjawisku, które zachodzi wówczas, gdy kanał od źródła dźwięku do odbiorcy jest otwarty i żadna przeszkoda go nie przetrnie, co jest związane z funkcjonowaniem uwagi, będzie mowa w dalszej części tego rozdziału poświęconej roli uwagi.

1.2.1. Rola percepcji słuchowej w budowaniu obrazu świata u osób z dysfunkcją wzroku

Szczególną rolę u osób z dysfunkcją wzroku pełni percepcja słuchowa. Dzięki niej osoby niewidome w dużej mierze budują swój świat. Poprzez odgłosy odbierają informacje przestrzenne: gdzie się znajdują, gdzie jest zlokalizowana przeszkoda, którą należy pokonać, gdzie stoi dana osoba i czy otoczenie jest bezpieczne. Wreszcie po usłyszanych dźwiękach rozpoznają postacie: osoby, zwierzęta, maszyny itd. Zmysł słuchu jest wspomagany przez pozostałe zmysły, a więc w szczególności dotyku, a poza tym zapachu i smaku. W tej specyficznej sytuacji percepcja słuchowa jest szczególnie dobrze rozwinięta.

Także informacje językowe, które służą do budowy wyobrażenia świata, w dużej mierze są odbierane za pomocą języka mówionego. W celu tworzenia możliwie pełnego modelu świata wykorzystuje się tzw. werbalizmy, czyli wyrazy i sformułowania odnoszące się do obiektów i zjawisk dostępnych osobom widzącym drogą wzrokową, a osobom z dysfunkcją wzroku drogą słuchową, językową. Zakłada się, że werbalizmy to wyrazy, językowe opisy zjawisk wizualnych, które oczywiście są niedostępne drogą wzrokową dla osób z dysfunkcją wzroku, a które pozwolą takim osobom zbudować w wyobraźni wzrokowy świat. Ta metoda, wykorzystywana często w celu przybliżenia osobom z dysfunkcją wzroku pełni wrażeń, np. estetycznych, jest stosowana podczas spektakli teatralnych bądź niekiedy w trakcie emisji filmu w kinie (osoba z dysfunkcją wzroku otrzymuje wówczas słuchawki, a lektor mówi, co dzieje się na scenie) – są to tzw. audiodeskryptory (zob. rozdział 3.5.1, zasada nr 1). Także opisy obiektów bardzo małych, nietypowych, dużych bądź po prostu niebezpiecznych, których osoba z dysfunkcją wzroku nie może poznać, korzystając z percepcji kinestetycznej (dotyku), pozwalają na pozyskiwanie cennych informacji o otaczającym świecie. Zakłada się przy tym, że dla tych osób wyrażenia językowe są puste, tj. przezroczyste znaczeniowo, czyli stanowią coś w rodzaju dodatkowych informacji umożliwiających pełniejsze zrozumienie rzeczywistości. Pytanie tylko, co to znaczy, że są to opisy przezroczyste

znaczeniowo. Osobną kwestią jest bowiem sposób językowego opisu. Jeśli opis jest rzeczowy, lakoniczny, to wymaga od osoby z dysfunkcją wzroku większego wysiłku, by zbudować własną interpretację. Z kolei jeżeli opis jest językowo barwny, ekspresyjny, to niejako narzuca odbiorcy interpretację, co z pewnością stanowi dla niego ułatwienie, choć nie zawsze musi odbywać się zgodnie z jego wolą. Kolejnym problemem dla piszącego tekst do audio-deskryptora jest wybór złożoności języka, dobór odpowiedniego słownictwa i konstrukcji językowych. Wiadomo, że innych środków językowych używa się w zależności chociażby od grupy wiekowej odbiorców, gdy tekst kierowany jest do dzieci, młodzieży lub dorosłych. Poza tym ważny jest sposób układania zdań, które nie powinny być za długie, wielokrotnie złożone, zaś konstrukcje w nich użyte zbyt skomplikowane, gdyż wszystkie te wymienione elementy utrudniają zrozumienie słyszanego tekstu (chyba że jest on kierowany do odbiorców mających duże doświadczenie w słuchowym, językowym odbiorze świata).

Można mieć wątpliwości, czy jest celowe, by niewidomi korzystali ze sformułowań określających percepcję niedostępną dla nich (np. nazw kolorów). Naszym zdaniem ten sposób przedstawiania świata osobom z dysfunkcją wzroku nie jest pozbawiony sensu, umożliwia im bowiem pełniejsze funkcjonowanie w społeczeństwie, nie odcina ich od całościowej, kompletnej informacji. Poza tym ważny jest moment, kiedy dana osoba straciła wzrok⁸. Jeśli stało się to w takim momencie jej życia, gdy zdążyła poznać np. paletę barw, to wówczas zarówno percepcja słuchowa mowy z wykorzystaniem nazw kolorów, jak i jej posługiwanie się językiem z użyciem owych nazw nie są bez znaczenia i odwołują się do pamięci osoby pozbawionej możliwości wzrokowego poznania. Z kolei osobom, które straciły wzrok wcześniej bądź nawet od urodzenia są niewidome, czyli osobom ociemniałym, kolory mogą być przedstawiane za pomocą temperatury, czyli jako ciepłe (ciepło emitowane przez słońce, rozgrzana powierzchnia) w przypadku koloru żółtego, czerwonego itd., lub zimne (śnieg, zimna woda) w przypadku białego, niebieskiego itp.

Te wszystkie zabiegi, a więc korzystanie głównie z komunikacji językowej oraz pozostałych dostępnych percepcji, czyli dotyku, powonienia, smaku, mają pozwolić osobie z dysfunkcją wzroku możliwie najpełniej poznać otaczający świat i zbudować jego obraz w swoim umyśle. Trzeba pamiętać, że ta

⁸ O różnym podłożu i stopniu uszkodzenia zmysłu wzroku zob. rozdział 3.2.

dysfunkcja nie może osobom niedowidzącym zamykać drogi ich rozwoju intelektualnego, budowania w miarę pełnego obrazu rzeczywistości zgodnie z ich wyobrażeniami.

Internet stanowi obecnie bodaj najpowszechniejsze, dostępne i zawierające bogaty zasób wszelkich informacji źródło wiedzy o świecie. Dlatego tak ważne i cenne jest tworzenie interfejsu przyjaznego użytkownikom, w tym także osobom z dysfunkcją wzroku. W odpowiedzi na potrzeby osób z dysfunkcją wzroku został skonstruowany i zbudowany czytnik stron. Pod koniec tego rozdziału zostaną przedstawione najistotniejsze lingwistyczne wskazówki, którymi warto się kierować, konstruując tekst na stronę internetową, by był on zrozumiały dla użytkownika.

1.3. ROLA UWAGI W PRZETWARZANIU INFORMACJI JĘZYKOWYCH

Wspomnieliśmy już o roli uwagi w procesie przetwarzania informacji. Teraz rozważymy, dlaczego jest ona tak ważna. Uwaga jest procesem cechującym się dużą dynamiką w poszukiwaniu i odnajdywaniu istotnych dla człowieka bodźców. Badania nad złożonością tego procesu umożliwiły ustalenie, że na proces uwagi składają się podsystemy pracujące i mające swoją lokalizację w strukturach mózgowych⁹.

Badania nad uwagą pozwoliły wysunąć wniosek, iż sposób kodowania informacji o różnych bodźcach w różnych układach sensorycznych opiera się na podobnych zasadach ogólnych. Ważną rolę w rejestrowaniu i kodowaniu informacji odgrywa intensywność bodźca. Jest to istotna cecha, albowiem intensywność bodźca wpływa na częstotliwość wyładowań: bodźce silniejsze wywołują większy potencjał receptorowy, co z kolei prowadzi do zwiększonej częstotliwości impulsacji w pojedynczym włóknie. Jednocześnie bodźce silniejsze pobudzają większą liczbę receptorów, czyli włókien (i tak bodźce wzrokowe o większej intensywności odbieramy jako jaśniejsze, bodźce słuchowe zaś jako głośniejsze).

Poza tym bodźce różnią się pod względem jakości. Światło padające na siatkówkę oka może zawierać fale o różnej długości, co odbieramy jako

⁹ M.I. Posner, M.E. Raichle, *Images of Mind*, Scientific American Library, New York 1994.

różną barwę. Dźwięk mogą stanowić fale o różnej częstotliwości, co odbieramy jako różną wysokość dźwięku. Bodźce oddziałujące na skórę mogą mieć charakter termiczny, dotykowy lub bólowy. Owe różnice jakości i intensywności bodźca są odbierane przez wyspecjalizowane receptory i przekazywane niezależnymi kanałami do mózgu.

Ograniczenie roli uwagi do rejestracji, identyfikacji bodźców byłoby jednak zbyt dużym uproszczeniem. Uwaga jest skomplikowanym procesem o wielopoziomowym przebiegu (i angażującym różne struktury mózgowe), który umożliwia człowiekowi jak najlepsze funkcjonowanie w środowisku. A zatem uwaga poszukuje i wybiera istotne dla człowieka informacje, by powiązać je z aktualnymi potrzebami, wykorzystując przy tym zapamiętane już przez człowieka informacje, a także pozwala mu zaplanować i wykonać jego przyszłe działania. Jak widzimy, uwaga jest procesem koordynującym informacje płynące z aktualnego doświadczenia z informacjami zmagazynowanymi w naszej pamięci, a ponadto umożliwia planowanie naszych aktywności (w tym ruchów czy mowy).

Wiemy, że te bodźce otaczającego nas świata, które zostaną zarejestrowane przez uwagę, mają szansę zostać przetworzone, a jeśli są dla nas istotne, mogą być nawet zapamiętane. Dlatego tak ważne jest odpowiednie kierowanie tego naturalnego filtra, jakim jest uwaga, na owe aspekty rzeczywistości, które są dla nas istotne. Pamiętajmy, że uwaga rejestruje tylko pewną porcję informacji. Tylko te, które rejestruje uwaga, są przekazywane do dalszego świadomego przetwarzania i zapamiętywania. Wiemy, że powtórzona informacja zostaje ponownie uchwycona przez uwagę i jeszcze lepiej zapamiętana jako ważna. Dlatego redundancja w języku mówionym, ze względu na specyfikę, czyli selektywność, zmienność uwagi jest konieczna, a w języku piśmianym – pomocna, albowiem powtórzenie treści w konkluzji wywołuje u odbiorcy wrażenie, że tekst jest spójny i zrozumiały, zaś wnioski zostaną przez niego zapamiętane.

Powiedzieliśmy sporo o specyfice uwagi: jej zmienności, poszukiwaniu nowych bodźców oraz znaczeniu dla zapamiętywania ważnych informacji. Człowiek zwraca uwagę na określone słowa: mogą być one poszukiwane przez niego, mogą budzić jego ciekawość, zainteresowanie, chęć bliższego poznania. Niestety, ludzkiej uwagi nie da się określić statycznie. Ze zjawiskiem zmienności ludzkiej uwagi boleśnie zetknęli się zwolennicy statystycznych algorytmów wyszukiwania informacji. Wszelkie bowiem próby statystycznego modelowania potrzeb informacyjnych człowieka nie uwzględniają

natury uwagi i w rezultacie człowiek jest zawsze tylko częściowo zadowolony z wyników wyszukiwania¹⁰.

Aktywność uwagi, która jest powiązana z procesami percepcji, możemy zaobserwować w pewnych jej przejawach. Do takich należy opisany wcześniej efekt *cocktail party*. Jest to zjawisko pojawiające się w sytuacji, gdy człowiek przebywa w gwarным, zatłoczonym miejscu. Ponieważ zostało zaobserwowane na przyjęciu, dlatego uzyskało takie miano, ale oczywiście jego występowanie nie ogranicza się tylko do tego rodzaju sytuacji. Zauważono, że nawet w głośnym miejscu, gdy człowiek usłyszy swoje imię lub interesujący dla niego, toczący się w pewnej odległości dialog, potrafi tak skoncentrować swą uwagę, że może przysłuchiwać się głównie toczącej się w pewnej odległości rozmowie. W tym przypadku selekcja (a więc uwaga) prawdopodobnie jest oparta na kierunkowości dźwięków. Wychwytywane są dźwięki nadchodzące z określonego kierunku, a inne są tłumione. Kłopot pojawia się, gdy zakłócenia (np. intensywny szum wody) docierają z tego samego kierunku lub gdy ocena kierunku jest utrudniona przez zatkanie jednego ucha. W takich sytuacjach proces wyłączenia jest utrudniony.

Nie będziemy pisać o pamięci, gdyż jest to bardzo rozległy przedmiot badań poznawczych, w tym szczególnie psychologicznych. Warto jednak wspomnieć o tzw. pamięci krótkotrwałej. Okazuje się, że w bardzo krótkim czasie trzeba zadbać o to, by ważna informacja zaangażowała pamięć krótkotrwałą, gdzie może zostać przetworzona i przeniesiona do magazynu pamięci długotrwałej. W tym zadaniu nieodzowna jest rola uwagi, która decyduje o tym, jakie informacje będą przetwarzane. Wiemy, choćby z badań George'a Millera¹¹, że pamięć krótkotrwała ma stosunkowo niewielką pojemność (od 5 do 9 elementów, gdzie elementami mogą być zarówno cyfry bądź pojedyncze litery, jak i wyrazy oraz całe dobrze zapamiętane zdania, czyli sensownie zorganizowane całości) i trwa bardzo krótko, bo zaledwie kilka sekund¹².

¹⁰ J. Allan, *HARD Track Overview in TREC 2003. High Accuracy Retrieval from Documents*, w: *Proceedings of the Twelfth Text Retrieval Conference (TREC 2003)*, NIST; C. Buckley, S. Robertson, *Relevance Feedback Track Overview: TREC 2008*, w: *Proceedings of the Seventeenth Text Retrieval Conference (TREC 2008)*, NIST.

¹¹ G.A. Miller, *Human Memory and the Storage of Information*, „IRE Transactions on Information Theory” 1956, 2 (3), s. 129–137.

¹² Wcześniejszy, powstały w latach 60. i 70. XX wieku, model pamięci Richarda Shiffrina i Richarda Atkinsona został skrytykowany; późniejsze badania nad uwagą rzuciły nowe światło na przebieg procesów przetwarzania informacji przez człowieka, niemniej jednak nie mamy nowego, kompletnego i spójnego modelu pamięci.

1.4. PRZETWARZANIE MOWY

Kiedy mówimy o percepcji, mówimy o odbiorze bodźców (informacji). Już wiemy, że uwaga jest tym procesem, swoistym filtrem, który decyduje, jakie informacje spośród ogromu zalewających nas danych wybrać i zacząć je przetwarzać. Teraz, w dużym skrócie, przyjrzymy się, jak przebiega proces przetwarzania informacji językowej.

Ze względu na niezwykle istotną rolę języka w życiu człowieka, ludzki mózg został wyposażony w odpowiednie struktury, które są wyspecjalizowane w zadaniach polegających na zrozumieniu mowy. Naturalnie, odpowiednio inne struktury w ludzkim mózgu odpowiadają za tworzenie mowy. Istnieją badania potwierdzające aktywność lewej półkuli w wyczuciu rytmu i to bez względu na ręczność. Lewa półkula mózgu odpowiada za rozumienie mowy, planowanie i tworzenie wypowiedzi oraz za uruchamianie w układzie nerwowym procesów prowadzących do realizacji głosek. Przyjmuje się, że prawa półkula ludzkiego mózgu pozwala na zrozumienie języka metaforycznego, ironii, humoru, zabarwienia emocjonalnego. Widzimy zatem, że obie półkule mózgowe współpracują w prawidłowym korzystaniu z języka. Taka organizacja ludzkiego mózgu pokazuje, jak ważną rolę pełni język w życiu człowieka. Jak już powiedzieliśmy na wstępie, język naturalny jest narzędziem poznawania i opisu świata, a także służy do komunikowania się z innymi ludźmi: wyrażania siebie, poznawania innych, wymiany poglądów. Przyjrzyjmy się elementom języka, które umożliwiają percepcję i przetwarzanie mowy.

1.4.1. Elementy prozodyczne języka polskiego

Na proces rozumienia mowy składa się warstwa prozodyczna, która poniżej zostanie krótko omówiona, oraz warstwa leksykalna, czyli znaczenia wyrazów, którą omówimy w dalszej części tekstu. Zaczniemy od segmentu, czyli od fragmentu sygnału mowy, który w normalnych warunkach słyszenia daje nierozzerwalne wrażenie dźwiękowe. To minimalny, percepcyjnie jednolity element mowy. Zjawiska segmentalne dotyczą oddzielnych, kolejnych elementów (np. *dom* – 3 segmenty – postać fonetyczna). Bardziej złożonymi elementami mowy są zjawiska suprasegmentalne, które obejmują pewne ciągi elementów. Sylaba (zgłoska) jest ciągiem głosek określonym dla każdego języka przez uniwersalne prawo zmiennej głośności segmentów oraz dodatkowo przez specyficzne dla każdego języka reguły. Każda sylaba składa się z części zwanej szczytem, który stanowi głoska o maksymalnej głośności.

Kolejnym elementem prozodycznym mowy jest rytm. Opisy rytmu w mowie (w przeciwieństwie do poezji) nie są w pełni zgodne ani ściśle. Jako najbardziej ogólną cechę rytmu przyjmuje się zjawisko izochronizmu, polegające na tendencji do wyrównywania czasowego określonych ciągów segmentów. Gdyby nie zachodziła tendencja izochronizmu, wymówiony wyraz *‘spał’* (1 sylaba) byłby dwa razy krótszy niż *‘spalby’* (2 sylaby), trzy razy krótszy niż *‘spalaby’* (3 sylaby) i cztery razy krótszy niż *‘spalibyśmy’* (4 sylaby). W wyniku działania izochronizmu takie relacje nie zachodzą. Co prawda w podanych przykładach wyrazy są coraz dłuższe, ale izochronizm powoduje kompresję iloczynową, w wyniku której jednostka rytmiczna *‘spalibyśmy’* jest około 2–3 razy dłuższa niż *‘spał’*.

Do zjawisk suprasegmentalnych należy też np. intonacja, której jednostki funkcjonalne rozciągają się niejednokrotnie na długie ciągi, nawet kilkadziesiąt segmentów (np. *jaka dzisiaj jest pogoda?* – 1 fraza intonacyjna). Jednostką intonacyjną jest fraza, czyli część wypowiedzi o pewnym określonym przebiegu wysokości tonu, na który składa się część konstytutywna (obowiązkowa) zwana rdzenną oraz część fakultatywna (nieobowiązkowa) zwana wstępną. Bardziej istotne lingwistycznie są intonacje rdzenne. Występują tu przebiegi rosnące, opadające i równe. W języku polskim intonacja ma charakter logiczny i emocjonalny. Jej funkcja polega na wyrażeniu postawy mówiącego do wypowiedzanej treści. Intonacja pełni ważną funkcję w rozumieniu intencji mówiącego. Może być wyrazem subiektywnej postawy mówiącego, jeśli stawia on pytanie: występuje wówczas intonacja rosnąca (wznosząca), czyli tzw. antykadencja. W intonacji rosnącej początek jest wymawiany niżej a koniec wyżej. Taki przebieg intonacji zachęca słuchacza do udzielenia odpowiedzi. Z kolei intonacja opadająca (tzw. kadencja) charakteryzuje zdania, które wyrażają postawę subiektywną wykrzyknikową. Jeśli natomiast nadawca coś twierdzi – wyraża postawę obiektywną, to wówczas najczęściej intonacja jest rosnąco-opadająca. Czasami konstrukcja zdania nie wskazuje na intencję nadawcy: równie dobrze może być pytaniem, co twierdzeniem (np. *dzisiaj jest ładna pogoda*), a wówczas to intonacja będzie wskazywać na intencję nadawcy. Intonacja rosnąca, która pojawia się podczas stawiania pytania, zachęca do udzielenia odpowiedzi, zaś intonacja opadająca jest sposobem podania nacisku. Zmiany w intonacji polegające na ujednoliceniu są diagnozowane jako patologia mowy¹³.

¹³ Np. zmiany w mowie powstałe w wyniku chorób neurologicznych o typie dyzartrii (I. Gatkowska, *Diagnoza dyzartrii u dorosłych w neurologii klinicznej*, Wydawnictwo Uniwersytetu Jagiellońskiego, Kraków 2012).

Kolejnym elementem suprasegmentalnym mowy jest akcent. Akcent wyrazowy polega na wyróżnieniu realizacji sylaby w wyrazie. Akcent wyrazowy w języku polskim manifestuje się poprzez zwiększoną siłę głosu (przycisk) w realizacji akcentowanej sylaby. Jest ruchomy pod względem morfologicznym (np. *młody*, *młodego*) oraz paroksytoniczny, czyli pada najczęściej na drugą od końca sylabę wyrazu. Zdarzają się realizacje, w których akcent pada na trzecią (proparoksytoniczny), a niekiedy czwartą sylabę od końca. Akcent na trzeciej sylabie od końca mają czasowniki w 1. i 2. osobie liczby mnogiej czasu przeszłego oraz w liczbie pojedynczej i w 3. osobie liczby mnogiej trybu przypuszczającego, a także formy liczebników od 400 do 900 oraz te wyrazy zapożyczone, które nie zostały całkowicie przyswojone przez system akcentowy języka polskiego (typu: *fizyka*, *matematyka*, *uniwersytet*). Akcent na czwartej sylabie od końca mają czasowniki w formie 1. i 2. osoby liczby mnogiej trybu przypuszczającego, np. *rozmawialiśmy*.

Akcent zdaniowy w języku polskim polega na zwiększeniu siły głosu na tym wyrazie, który ma szczególne znaczenie ze względu na przekazywane treści bądź emocje (akcent ma charakter logiczny). Na akcent składają się zestroje akcentowy (jednostka rytmiczna) oraz fraza (jednostka intonacyjna). Zestroj akcentowy może składać się z jednej lub więcej sylab, z jedną sylabą wyróżnioną rytmicznie. Może obejmować jeden lub kilka wyrazów. I odwrotnie: jeden dłuższy wyraz może stanowić dwa zestroje akcentowe z dwiema sylabami wyróżnionymi rytmicznie (np. *'nieskazitelny'*). Akcent rytmiczny, objawiający się zestrojami akcentowymi o charakterze rytmicznym, które podlegają izochronizmowi, jest w języku polskim słabo zaznaczony. Znacznie wyraźniejszy jest akcent intonacyjny.

Umiejętne korzystanie z elementów prozodycznych języka, czyli właściwe tempo mowy, akcentowanie, intonacja, znakomicie pozwalają wyrazić intencje mówiącego, podkreślić istotne dla wypowiedzi fragmenty, a w ten sposób ułatwić odbiorcy przyjęcie wypowiedzi zgodnie z intencją mówiącego. Zatem, jak widać, sfera prozodyczna języka jest szczególnie istotna dla rozumienia mowy.

Podsumowując, omówiliśmy krótko elementy prozodyczne języka, czyli segment, tj. najmniejszy dostępny percepcyjnie element mowy (pojedynczy fonem), oraz zjawiska suprasegmentalne (prozodyczne), do których zaliczamy: rytm, tempo, akcent i intonację. To, co człowiekowi ułatwia percepcję (odbior) mowy, to odpowiednie posługiwanie się pauzami przez nadawcę komunikatu.

1.4.2. Rola pauz w percepcji i przetwarzaniu mowy

Pauza stanowi przerwę w ciągu komunikatu słownego i pełni bardzo ważną funkcję: z jednej strony pozwala nadawcy (wygłaszającemu komunikat) zaczerpnąć oddech, przełknąć ślinę, a zatem jest potrzebna ze względów fizjologicznych, a także umożliwia zebranie i sformułowanie myśli, dobranie odpowiedniego słowa, czyli jest potrzebna ze względów intelektualnych, zaś z drugiej strony pozwala na zrozumienie tekstu przez odbiorcę. Jak założył Ernst Pöppel¹⁴, takie 2–3-sekundowe przerwy w sygnale mowy pomagają z jednej strony wypowiedzieć frazę, czyli ciąg wyrazów stanowiących logiczną całość, z drugiej strony można przypuszczać, że istnieją pewne mechanizmy umożliwiające identyfikację sekwencji wydarzeń i integrację informacji. Mechanizm identyfikacji sekwencji wydarzeń zapewnia prawidłową sekwencję głoski, prawdopodobnie trwa kilkadziesiąt milisekund, czyli tyle ile potrzeba na artykulację i rozumienie. Słuchając mowy, słyszymy uporządkowane ciągi myślowe, a nie same tylko dźwięki: głoski czy sylaby. Według tej teorii, mechanizm integracji informacji scala głoski i sylaby w większe jednostki, jakimi są logiczne ciągi myślowe. Pöppel założył, że informacja wypowiedziana (zarówno drogą wzrokową, jak i słuchową) jest pakowana w „większe paczki”, które są oddawane do dyspozycji świadomości. Są to 2–3-sekundowe jednostki, w trakcie których dochodzi do integrowania wydarzeń w całość. Zatem prawdopodobnie pewna część ludzkiej świadomości może przetrwać kilka sekund. Po tym czasie wyczerpuje się nasza zdolność do integracji, a w świadomości musi znaleźć się miejsce dla czegoś nowego. Wydaje się więc, że istnieje wyraźny związek między świadomością i uwagą.

Prawdopodobnie u podstaw rozumienia mowy leży mechanizm opracowania informacji językowej w czasie. Termin ‘zegar wewnętrzny’ wskazuje na specyficzny mechanizm czasowy, który organizuje proces nadawania i odbioru informacji słownych. Taki 2–3-sekundowy rytm jest niezależny od języka. Koncepcja Pöppela była przedmiotem badań dla języków: niemieckiego, angielskiego, polskiego oraz chińskiego i w każdym z nich zauważono, iż mówienie trwa około 3 sekund, potem następuje pauza i kolejna trzysekundowa jednostka, mimo że budowa systemu językowego każdego z wymienionych języków jest całkowicie odmienna. Wiemy też, że przeżywanie czasu

¹⁴ E. Szeląg, E. Pöppel, *Temporal Perception: A Key to Understand Language*, „Behavioral and Brain Sciences” 2000, nr 23 (1); E. Szeląg, *Mózgowe mechanizmy mowy*, w: *Mózg a zachowanie*, red. T. Górską, A. Grabowska, J. Zagrodzka, Wydawnictwo Naukowe PWN, Warszawa 2005, s. 489–524.

angażuje tylko pewne struktury mózgu i to te, które są połączone z określonymi funkcjami języka.

To istotna wskazówka, którą warto wykorzystać, tworząc tekst, by był zrozumiały w odbiorze. Ważne jest odpowiednie, celowe rozmieszczenie pauz, by nie rozdzielały informacji stanowiących konceptualną całość, nie rozdzielały silnych powiązań syntagmatycznych¹⁵.

1.4.3. Modele przetwarzania mowy

Są teorie percepcji mowy zakładające, że ważna jest informacja, którą słyszymy na wejściu. Dźwięki, które do nas docierają, porównujemy ze wzorcem dźwięku istniejącym w naszym umyśle – w ten sposób dochodzi do identyfikacji sygnału ze wzorcem. Zatem, według tego ujęcia, podstawową informacją dla percepcji słuchowej jest sygnał otrzymany na wejściu. Są to teorie typu dół-góra (ang. *bottom-up*), czyli **przetwarzanie wstępujące**. To ujęcie dotyczy różnych percepcji człowieka, także wzrokowej; wśród jego zwolenników można wymienić psychologów ze szkoły Gestalt. Najpierw odbieramy bodziec, zwracamy uwagę na jego charakterystyczne cechy, następnie dopasowujemy do czegoś, co mamy w naszym umyśle (wzorzec, prototyp), poznawczo organizujemy to w konceptualną całość i wiążemy z posiadaną wiedzą. Na tym fundamencie rozwinęły się teorie wzorca, teorie prototypu, teorie cech i teorie opisu strukturalnego.

Na podstawie tego podejścia badawczego dla percepcji słuchowej rozwinęły się tzw. biernie teorie. Teorie typu dół-góra zakładają, że filtrujemy dźwięki ze względu na cechy. Dzieje się to na poziomie sensorycznym (odbiorczym); na tym etapie nie odbywa się przetwarzanie poznawcze, określane jako przetwarzanie wyższego rzędu. To podejście badawcze na pierwszym miejscu stawia dopasowanie wzorca lub detekcję cech na poziomie sensorycznym. Teorie te zakładają, że istnieją wyspecjalizowane neurony, które wykrywają określone cechy sygnału mowy. Co więcej, występują różne etapy przetwarzania neuronalnego: na jednym z nich dźwięki są analizowane ze względu na komponenty, na innym komponenty te są analizowane ze względu na układy, w jakich występują¹⁶. Takie przetwarzanie, w zależności od teorii, może występować równolegle bądź sekwencyjnie. Są też inne koncepcje zakładają-

¹⁵ O relacjach semantycznych, w tym syntagmatycznych, traktuje rozdział 2.

¹⁶ D.W. Massaro, *Speech Perception by Ear and Eye: A Paradigm for Psychological Inquiry*, Erlbaum, Hillsdale, NJ 1987.

ce, że niektóre teorie percepcji mowy nie postulują występowania specjalnych detektorów mowy, które różniłyby się od detektorów wykorzystywanych w słyszeniu innych, niezwiązanych z mową dźwięków¹⁷.

Drugi nurt podejścia badawczego prezentują teorie, które zakładają, że ważna jest nie tyle informacja otrzymana na wejściu, czyli sygnał, który usłyszymy na wejściu, ile kontekst, nasze poznawcze oczekiwania, nasza wiedza, czyli to, co już posiadamy. Są to teorie typu góra-dół (ang. *top-down*), czyli **przetwarzanie zstępujące**. Oczywiście to podejście występuje także w badaniach nad innymi percepcjami człowieka.

Dla percepcji słuchowej teorie góra-dół rozwinęły się w nurcie tzw. teorii aktywnych. Wszystkie teorie aktywne w percepcji mowy zakładają konieczność występowania procesów decyzyjnych. Rozumienie mowy polega nie tylko na percepcji słuchowej, ale uruchamia także procesy poznawcze, w tym pamięć semantyczną. Zatem ważne jest zarówno to, co słyszymy, jak i to, co wnioskujemy na temat wypowiedzi, zamiarów mówiącego. Istotną rolę odgrywa tu kontekst, który stał się przedmiotem badań eksperymentalnych. Szczególne zainteresowanie badawcze budzą uczący się języków, chociażby Japończycy uczący się języka angielskiego, bo jak wiemy, w języku japońskim nie ma opozycji $r : l$, która występuje i jest istotna dla języka angielskiego. Okazuje się, że Japończycy z kontekstu domyślają się, czy w danym wyrazie występuje r czy l .

Jedną z teorii góra-dół jest model kohorty¹⁸ (*the Cohort Model of Word Recognition*), który powstał w wyniku eksperymentów nad rozpoznawaniem przez człowieka słów. Badania te wykazały, że człowiek zwraca największą uwagę na początek wyrazu, mniejszą natomiast na część środkową czy końcową. Z kolei inne badania nad percepcją mowy dowiodły, że dla człowieka ważne jest nie tylko to, co usłyszy, ale dużą rolę w rozpoznawaniu słowa i znaczenia odgrywa też wiedza językowa. Wykorzystując te eksperymentalne

¹⁷ Por. np. P.D. Eimas, J.D. Corbit, *Selective Adaptation of Linguistics Feature Detectors*, „Cognitive Psychology” 1973, nr 4, s. 99–109.

¹⁸ W.D. Marslen-Wilson, *Function and Process in Spoken Word Recognition*, w: *Attention and Performance X: Control of Language Processes*, red. H. Bouma, D. Bouwhuis, Erlbaum, Hillsdale, NJ, s. 125–150; W.D. Marslen-Wilson, *Functional Parallelism in Spoken Word-Recognition*, „Cognition” 1987, nr 25, s. 71–102; W.D. Marslen-Wilson, A. Welsh, *Processing Interactions and Lexical Access during Word Recognition in Continuous Speech*, „Cognitive Psychology” 1978, nr 10, s. 29–63; M.G. Gaskell, W.D. Marslen-Wilson, *Integrating Form and Meaning: A Distributed Model of Speech Perception*, „Language and Cognitive Processes” 1997, nr 12, s. 613–656.

doniesienia, skonstruowano model, który zakłada dwuetapowe rozpoznawanie słowa. W pierwszym etapie przetwarzania słów człowiek słyszy dźwięk, czyli odbiera informację akustyczno-fonetyczną zawartą w pierwszym dźwięku słyszanego słowa, która uruchamia wszystkie znajdujące się w pamięci słowa zaczynające się od tego samego dźwięku (to tzw. wstępna kohorta słów). Następnie kolejne informacje docierające do odbiorcy, ponieważ nie są podobne do rozpoznanego słowa, pozwalają wyeliminować niepożądane słowa bądź zawartość kohorty wstępnej jest zawężana przez kontekst zdania, wypowiedzi. Teoria zakłada, że informacje pochodzące z wyższych poziomów pozwolą wyeliminować słowa, które nie są zgodne z danymi semantycznymi i syntaktycznymi. W rezultacie tych procesów eliminacji pozostanie jeden kandydat, właściwe słowo, które odpowiada wymaganiom, i tym samym zakończy się proces rozpoznawania.

Inną z aktywnych teorii percepcji mowy jest teoria udoskonalenia fonetycznego¹⁹ (*Phonetic Refinement Theory*), która stara się łączyć bierną analizę wrażeń słuchowych z aktywnym przetwarzaniem. Najpierw odbieramy informacje słuchowe, a następnie je aktywnie przetwarzamy, korzystając przy tym z zasobów naszej pamięci. Zatem według tej koncepcji percepcji mowy: najpierw rejestrujemy dźwięki, a po chwili dociera do nas to, co usłyszeliśmy (być może na tym etapie uaktywniają się procesy uwagi i świadomości).

Kolejną ciekawą koncepcją jest model TRACE²⁰, który odwołuje się do działania sieci neuronowych. Stanowi on dość ciekawą propozycję wykorzystania wiedzy dotyczącej percepcji mowy rozumianej jako koneksjonistyczny, interakcyjny model aktywacji do stworzenia komputerowego programu symulującego działanie ludzkiego umysłu/mózgu podczas przetwarzania informacji zawartych na poziomie dźwięków mowy (głosek/fonemów²¹) i słów (wyrazów) w czasie rzeczywistym. W modelu tym założono, że percepcja mowy zaczyna się od trzech poziomów detekcji cech: poziomu cech akustycznych (*Feature Layer*), poziomu fonemów (*Phoneme Layer*) i poziomu słów (*Word Layer*). Model zakłada, że percepcja mowy ma charakter interakcyjny, czyli zarówno niższe poziomy wpływają na wyższe, jak

¹⁹ D.B. Pisoni, H.C. Nusbaum, P.A. Luce, L.M. Slowiaczek, *Speech Perception, Word Recognition and the Structure of the Lexicon*, „Speech Communication” 1985, nr 4, s. 75–95.

²⁰ J.L. McClelland, J.L. Elman, *The TRACE Model of Speech Perception*, „Cognitive Psychology” 1986, nr 18, s. 1–86.

²¹ Fonem to zespół cech dystynktywnych głoski. Głoskę natomiast stanowią wszystkie cechy konkretnego dźwięku realizowanego przez człowieka.

i na odwrót – ten wpływ ma charakter pobudzenia i hamowania poszczególnych węzłów. W tym modelu cechy dystynktywne fonemu i słowa tworzą tzw. węzły reprezentujące poziomy przetwarzania. Każdy węzeł ma poziom spoczynkowy, próg pobudzenia i poziom aktywacji. Kiedy informacja zostanie potwierdzona (czyli dochodząca informacja jest właściwa dla danego węzła), to jego aktywacja wzrasta do momentu osiągnięcia poziomu progowego; gdy brakuje potwierdzania informacji, to aktywacja węzła powoli się zmniejsza aż do poziomu spoczynkowego. System stanowi sieć powiązań między węzłami: jeśli któryś z nich osiągnie progowy poziom aktywacji, to może wpływać na aktywację powiązanych z nim węzłów. Reasumując, węzły, które osiągnęły wartość progową, mogą pobudzać inne węzły z nimi połączone, ale też mogą hamować aktywację powiązanych z nimi węzłów. Ten model zakłada, że węzły odpowiadające fonemom mogą pobudzać węzły odpowiadające słowom i przeciwnie: węzły słów mogą pobudzać węzły fonemów.

Kolejną aktywną teorią percepcji mowy jest model logiki zbiorów rozmytych²² (*Fuzzy Logical Model of Speech Perception*, FLMP). Model ten zakłada, że percepcja (otrzymany na wejściu bodziec wzrokowy bądź słuchowy) zostaje rozpoznana dzięki trzem fazom: ocenie cech, integracji cech, podjęciu decyzji co do zgodności z prototypem. Ta koncepcja opiera się na pojęciu prototypu (wzorca), a właściwie prototypów rozumianych jako sumaryczny opis percepcyjnych jednostek języka powiązanych z różnymi cechami dystynktywnymi. Cechy prototypu odpowiadają wartościom idealnym, jakie powinny cechować egzemplarze należące do danej kategorii. W pierwszej fazie następuje analiza słuchowych (*auditory features* – *af*) i wzrokowych (*visual features* – *vf*) cech. W drugiej fazie informacja cechy jest integrowana, pod warunkiem, że proces decyzyjny w trzeciej fazie może być użyty jako oczywisty dowód zaklasyfikowania dźwięku mowy. Model zakłada, że nieustannie otrzymywane informacje na temat cech są: oceniane, integrowane i porównywane z prototypami zakodowanymi w pamięci. Decyzja identyfikacyjna jest podejmowana na podstawie względnej zgodności w dopasowaniu informacji bodźcowej do odpowiednich prototypów. Poprzez ocenę cechy ustala się stopień dopasowania każdej cechy sylaby do odpowiadającej

²² D.W. Massaro, *Speech Perception by Ear and Eye...*; D.W. Massaro, *Testing between the TRACE Model and Fuzzy Logical Model of Speech Perception*, „Cognitive Psychology” 1989, nr 21, s. 398–421; D.W. Massaro, G.C. Oden G.C., *Speech Perception: A Framework for Research and Theory*, w: *Speech and Language: Advances in Basic Research and Practice*, t. 3, red. N.J. Lass, Academic Press, New York 1980, s. 129–165.

jej cechy prototypowej. W wyniku integracji cechy określany jest stopień, w jakim każdy z prototypów dopasowuje się do przetwarzanej sylaby. Potem następuje proces podejmowania decyzji o cechach: o tym, czy możemy mówić o względnej zgodności dopasowania, decyduje obliczana proporcja przypadków, w których sylaba zostaje zidentyfikowana jako egzemplarz danego prototypu. Model zakłada, że różne cechy odpowiadające danemu kontrastowi fonetycznemu są uzyskiwane z fali głosowej niezależnie, następnie zostają połączone zgodnie z zasadami integracji logicznej. Reguły te opierają się na rozmytych wartościach logicznych, dzięki czemu informacja dotycząca danej cechy może być reprezentowana jako pewien stopień dopasowania, a nie jako forma identyczna z prototypem. Przyjmuje się przy tym, że informacja ma postać ciągłą (a nie dyskretną). Należy zwrócić uwagę, że ten model odnosi się do percepcji sygnałów fonetycznych i jego porównania do cech prototypów bez odniesienia do wyższych poziomów języka, czyli poziomu wyrazów posiadających znaczenie (semantyki), a tym bardziej poziomu zdania (składni).

Warto jeszcze wspomnieć o modelu analizy przez syntezę²³ (*the analysis by synthesis, A x S*), który zakłada ścisły związek pomiędzy percepcją mowy a jej artykulacją. Słuchający odbiera i analizuje mowę, generuje ją wewnętrznie, czyli syntetyzuje na podstawie tego, co usłyszy, a w dalszej kolejności dochodzi do porównania tego, co powstało w wyniku syntezy z bodźcem słuchowym. Zatem proces percepcji zaczyna się od analizy słuchowych cech sygnału mowy. Analiza dostarcza opisu w kategoriach wzorców słuchowych. Następnie budowana jest hipoteza, która odnosi się do reprezentacji wypowiedzi w kategoriach cech dystynktywnych. Jeśli cechy fonetyczne nie są silnie zależne od kontekstu, czyli posiadają pewien niezmiennik, to wzorce słuchowe są dekodowane jako fonemy. Jeśli nie występują niezmienniki identyfikujące cechę fonetyczną, to wówczas uruchamiane jest dodatkowe przetwarzanie. Wtedy konstruowana jest hipoteza odnosząca się do reprezentacji cech dystynktywnych przetwarzanej wypowiedzi (zakłada się, że występuje system dopasowania, który polega na dopasowaniu bodźca do matryc abstrakcyjnych cech dystynktywnych). To punkt wyjścia dla zestawu reguł generatywnych, syntetyzujących potencjalne wzorce, które są porównywane

²³ K.N. Stevens, M. Halle, *Remarks on Analysis by Synthesis and Distinctive Features*, w: *Models for the Perception of Speech and Visual Form*, red. W. Wathen-Dunn, MIT Press, Cambridge, MA 1967, s. 88–102; K.N. Stevens, *Toward a Model for Speech Recognition*, „Journal of the Acoustical Society of America” 1960, nr 32, s. 47–55.

z rzeczywistością wypowiedzą na poziomie neuroakustycznym. Wyniki tego porównania są przekazywane do układu sterującego, który z kolei przekazuje ten opis fonetyczny wyższym poziomom analizy językowej. Późniejsze wersje tej teorii zakładają, że właściwości mowy mogą być jednoznacznie i niezmiennie określone na podstawie samego sygnału akustycznego. Kontynuowano także badania nad zagadnieniem niezmienników i niektóre prace, zwłaszcza dotyczące cech spółgłosek zwartych²⁴, wskazują na ich występowanie. Spotyka się opinie, że jest to model abstrakcyjny oraz że nie znaleziono wystarczająco silnych danych empirycznych, które by go potwierdzały. Wydaje się nam jednak, że takie niezmienniki występują, gdyż pozwalają na identyfikację sygnału mowy nawet przy znacznych zakłóceniach (naturalnie dużym ułatwieniem percepcyjnym jest kontekst, a wówczas już przenosimy się na wyższy niż segment poziom języka). Poza tym możliwa jest identyfikacja sygnału mowy dla osób nieznających danego języka bądź znających go bardzo słabo. Dodatkowo badania nad przetwarzaniem informacji językowej z wykorzystaniem technik neuroobrazowania pozwalają przyjrzeć się sposobom, w jakie ludzki mózg realizuje różne lingwistyczne zadania (w tym percepcję i artykulację). Tak więc być może wraz z większą liczbą doniesień ta koncepcja znajdzie swoje silne potwierdzenie empiryczne, zwłaszcza że percepcja mowy jest specyficzna w stosunku do percepcji słuchowej. Z jednej strony jej specyfika jest widoczna z perspektywy obserwacji: ludzie mają skłonność do słyszenia i nadawania sensu dźwiękom mowy, doszukiwania się treści w sygnale, który odbierają, z drugiej zaś strony badania nad funkcjonowaniem mózgu wskazują na specyficzne struktury mózgowe odpowiedzialne za percepcję, artykulację i przetwarzanie informacji językowej²⁵. Ludzki język ma takie właściwości, które umożliwiają dobrą percepcję mimo nawet sporych zakłóceń. Do takich wniosków skłaniają liczne eksperymenty badające możliwości rozumienia mowy przez człowieka. Zauważono, że w percepcji mowy ważny

²⁴ S.E. Blumstein, K.N. Stevens, *Perceptual Invariance and Onset Spectra for Stop Consonants in Different Vowel Environments*, „Journal of the Acoustical Society of America” 1980, nr 67, s. 648–662; S.E. Blumstein, K.N. Stevens, *Acoustic Invariance in Speech Production: Evidence from Measurements of the Spectral Characteristics of Stop Consonants*, „Journal of the Acoustical Society of America” 1979, nr 66, s. 1001–1017. Współcześnie wykorzystuje się założenie, że mamy w mózgu engramy ruchowe (prawdopodobnie w dodatkowym polu ruchowym), które odpowiadają za właściwe ruchy artykulacyjne. To podejście wykorzystywane w rehabilitacji afazji, czyli zaburzeniu mowy powstałym wskutek uszkodzenia struktur mózgowych odpowiedzialnych za posługiwanie się językiem.

²⁵ Zob. I. Gatkowska, *Jak mózg rozpoznaje języki*, w: *Język polski: nowe wyzwania językoznawcze*, red. J. Dybiec, G. Szpila, Krakowskie Towarzystwo „Tertium”, Kraków 2010, s. 15–29.

jest początek wyrazu. Jeśli mowa jest mniej staranna i poszczególne słowa są niewyraźnie realizowane, to całe usłyszane z dużymi zakłóceniami zdanie może zostać zrozumiane dzięki kontekstowi. Także przyspieszone tempo mowy stanowi pewne utrudnienie percepcyjne, jednakże eksperymenty pokazują, że ludzie lepiej radzą sobie w odbiorze i rozumieją nawet stosunkowo szybką mowę niż w przypadku zwolnionego tempa mowy²⁶. Prawdopodobnie dzieje się tak dlatego, że ludzie są przyzwyczajeni do szybszego tempa mowy (choćby pod wpływem emocji), natomiast zbyt wolne tempo mowy powoduje odpływ uwagi i zaangażowanie jej w procesie przetwarzania innych informacji (np. własnych myśli).

Dotąd nie powstał jeden model, który kompleksowo opisywałby proces percepcji mowy. Wszystkie wybrane przez nas i zaprezentowane powyżej koncepcje zawierają jakiś fragment tego jakże skomplikowanego wielopoziomowego procesu, który jest na tyle elastyczny, że pozwala zrozumieć ludzką mowę nawet przy stosunkowo poważnych zakłóceniach występujących przy jej odbiorze. Wiele z tych teorii powstało wskutek wcześniejszych eksperymentów badających przebieg i specyfikę percepcji jakiegoś aspektu mowy.

Podsumowując, wszystkie teorie aktywne różnią się od teorii biernych tym, że zakładają konieczność występowania, oprócz wykrywania cech i dopasowania wzorca, procesów podejmowania decyzji. Zatem czynniki kontekstowe i poznawcze wpływają na naszą percepcję sygnału odbieranego przez zmysły. W przetwarzaniu mowy ważną rolę pełnią przewidywania i oczekiwania.

1.5. PRZETWARZANIE INFORMACJI TEKSTOWEJ

1.5.1. Rozumienie języka

Prezentując modele przetwarzania mowy, podkreśliliśmy, że każdy z nich stanowi interesującą koncepcję, nie wyczerpuje jednak pełnego modelu przetwarzania mowy (który obejmuje percepcję i rozumienie). W praktyce wydaje nam się bezpieczne założenie, że w procesie interpretacji komunikatów mowy uwzględniamy różne aspekty wiedzy o języku naturalnym, takie jak:

²⁶ Także eksperymenty przeprowadzane przez studentów kierunku Elektroniczne Przetwarzanie Informacji podczas zajęć pt. *Reprezentacja języka w mózgu i umyśle* potwierdzają przedstawione stanowisko.

znajomość dopuszczalnych **sekwencji fonologicznych** w danym języku, znajomość specyfiki **warstwy prozodycznej danego języka**, zrozumienie tematu wypowiedzi (znajomość **semantyki** danego języka), znajomość budowy zdania w danym języku (**struktury syntaktycznej**), która powinna odpowiadać percypowanej sekwencji dźwięków mowy.

1.5.2. Kilka uwag o budowie tekstu zamieszczonego w serwisie

W tej części rozdziału na początku zaprezentujemy kilka uwag dotyczących warstwy językowej, a następnie sformułujemy podstawowe założenia struktury tekstu (w tym tekstu serwisu). Warto podkreślić, że uwagi te mają charakter ogólny, tzn. należy mieć je na względzie, budując każdy interfejs, który w założeniu projektanta ma być przyjazny dla użytkownika.

Znajomość budowy języka jest niezbędna do zrozumienia odbieranych treści. Naturalnie nie chodzi tu o fachową wiedzę językoznawczą, bo gdyby tak było, to porozumiewanie się za pomocą języka stałoby się domeną językoznawców, a przecież tak nie jest. Lingwiści badają system językowy oraz analizują zjawiska językowe, tymczasem za pomocą języka porozumiewają się wszyscy, którzy potrafią zrozumieć, a także zbudować i wyartykułować wypowiedzi w danym języku.

Liczne obserwacje i eksperymenty pozwalają sformułować tezę, że to semantyka języka naturalnego odgrywa najważniejszą rolę w rozumieniu języka. Jednak nie można lekceważyć roli gramatyki, która także pozwala na zrozumienie wypowiedzi, a więc zgodność i poprawną formę użytych wyrazów wraz z prawidłowym łączeniem wyrazów w bardziej złożone jednostki języka, czyli zdania. Nie bez znaczenia jest fakt, jak konstruowane są zdania: proste, złożone, wielokrotnie złożone. Im bardziej skomplikowana, zawiła jest struktura zdania, tym większą trudność sprawia zrozumienie go przez odbiorcę. W przypadku języka pisanego czytelnik może powrócić do jakiegoś fragmentu zdania, myślowo ustalić logiczne powiązania pomiędzy jego poszczególnymi składowymi, natomiast w przypadku języka mówionego takie operacje umysłowe są znacznie trudniejsze. Dlatego dla zrozumienia mowy szczególnie ważne jest, by zdania nie były nadmiernie rozbudowane, gdyż w najlepszym przypadku utrudnia to, zaś w najgorszym uniemożliwia zrozumienie tekstu. Zdania powinny być krótkie, proste, a wówczas będą łatwiejsze w percepcji i przetwarzaniu.

Ważny jest też dobór słownictwa. Należy dbać o to, by wykorzystywane wyrazy nie nastroczały odbiorcy problemów z odnalezieniem odpowiedniego znaczenia, zatem powinny to być powszechnie używane słowa. Jeśli są to wyrazy wieloznaczne²⁷, to tematyka serwisu, kontekst zdania, nagłówki będą rozstrzygać i pomagać odbiorcy w znalezieniu właściwego znaczenia. Ponadto trzeba mieć na uwadze, by odbiorca mógł bez wysiłku wybrać właściwe znaczenie wyrazu. Jeśli konieczne jest korzystanie z wyrazów specjalistycznych, to należy pamiętać o tym, by wyjaśnić znaczenia terminów. Powinno się przy tym dążyć do tego, by tekst był jak najbardziej zrozumiały, logiczny i precyzyjny, jasny dla przeciętnego odbiorcy. Zbyt zawile teksty będą traktowane przez odbiorcę jako niezrozumiałe, a w konsekwencji odrzucane. Należy unikać skrótów i skrótowców, posługując się raczej nazwami pełnymi, a jeśli koniecznie chcemy korzystać ze skrótowców, to musimy od razu podać pełną nazwę i to nawet w sytuacji, gdy jesteśmy pewni, że skrót jest powszechnie znany.

W dążeniu do zrozumiałości tekstu zgodnie z intencją nadawcy należy unikać form językowych, które mogą stanowić pewne utrudnienie dla odbiorcy, a więc żargonu, zbędnych przenośni, ironii. Ironia może być zrozumiana dosłownie, a wtedy odbiór rozminie się z intencją nadawcy. Słowa w umyśle użytkownika są z sobą powiązane, czyli kojarzą się tematycznie. To ważna wskazówka: musimy zdawać sobie sprawę z tego, że tekst, który często zawiera wyrazy wieloznaczne, powinien być osadzony w jasnym kontekście, użytkownik zaś będzie miał swoje oczekiwania, nastawienia co do poszukiwanych informacji. Zatem dobieramy słowa w ten sposób, by nie nastroczały wątpliwości percepcyjnych, lecz harmonijnie łączyły się z sobą. Jeśli chcemy zwrócić uwagę odbiorcy na jakąś informację, to wówczas stosujemy interesujące, zaskakujące połączenie. Jednak dla tekstu znajdującego się w serwisie (w przeciwieństwie do chociażby tekstu literackiego, poetyckiego) taka sytuacja musi być wyjątkowa, gdyż tylko wtedy spełni intencję nadawcy; w przeciwnym wypadku, gdy w tekście będzie wiele zaskakujących zestawień, spowoduje to chaos percepcyjny i w konsekwencji tekst zostanie odrzucony przez odbiorcę jako niezrozumiały.

Choć istotnie semantyka pełni zasadniczą rolę w zrozumieniu tekstu, należy dbać o poprawność form gramatycznych, gdyż także inne aspekty języka

²⁷ Często wyrazy mogą mieć bardzo różne znaczenia, np. *zamek*: *budowla*, *zamek w drzwiach*, *sposób zapięcia odzieży*.

naturalnego ułatwiają zrozumienie tekstu, a więc prawidłowe formy wyrazów, przestrzeganie zasad ortografii oraz interpunkcji. Gdy tekst jest czytany przez syntezytor mowy, to znaki interpunkcyjne przełożą się na przerwy oraz sposób odczytania tekstu²⁸.

Ważna jest właściwa **struktura tekstu**. Zacznijmy od tytułu. Sugeruje on, czego będzie dotyczył tekst. W przypadku serwisów internetowych bardzo ważne jest, by tytuł współgrał z treścią strony serwisu. Jest to szczególnie istotne dla osób z dysfunkcją wzroku, które korzystając z czytnika ekranu, zaraz po wejściu na stronę usłyszą najpierw tytuł strony. Zrozumiały tytuł, który w sposób zwięzły opisuje bądź nazywa treść strony, będzie znakomitym ułatwieniem dla użytkownika z dysfunkcją wzroku. Naturalnie, zrozumiały tytuł ma niebagatelne znaczenie także dla użytkownika widzącego, niemniej jednak osoba widząca jest w stanie szybko przerzucić wzrokiem treść strony i w ten sposób zorientować się, czy na stronie znajdują się poszukiwane przez nią informacje. Takiej możliwości nie ma osoba z dysfunkcją wzroku. Dlatego tytuł jest jednym z najważniejszych elementów stron internetowych. Dobrze, jasno skonstruowany tytuł pozwala użytkownikom na szybkie i wygodne korzystanie z serwisów internetowych.

Analizując dobrze skomponowany tekst, należy zacząć od podziału tekstu na mniejsze, czytelne fragmenty. Podstawową jednostką podziału tekstu jest akapit. Logiczny podział tekstu znajdującego się na stronie za pomocą akapitów może znacznie ułatwić jego zrozumienie. Akapit powinien zawierać logiczną całość myślową. Nie może być zbyt rozbudowany, należy raczej dążyć do zwięzłości przy zachowaniu wszelkich logicznie powiązanych informacji stanowiących wątek. Przyjmuje się, że pierwsze słowa powinny wskazywać, o czym dany akapit traktuje, albowiem to na nie odbiorca zwraca swą uwagę.

Kolejnym ważnym elementem struktury tekstu są nagłówki (śródtytuły). Zadaniem nagłówków stron internetowych jest zatytułowanie logicznie wyróżnionych akapitów (sekcji) krótkim, zwięzłym tekstem. Nagłówki służą do zwiększenia czytelności całej strony. Osoby zdrowe zwykle przerzucają tekst strony, skanują ją wzrokiem, wybierając nagłówki, oceniają, czy informacje zawarte na stronie są dla nich interesujące. Nagłówki są także bardzo ważną wskazówką dla osób z dysfunkcją wzroku. Dzięki nim mogą one szybko dowiedzieć się (usłyszeć), jakie treści zawiera dany fragment serwisu (strona). Nagłówki powinny być stosowane konsekwentnie na każdej stronie,

²⁸ Por. rozdziały 1.4.1 i 1.4.2.

gdyż to przyczynia się do uporządkowania treści serwisu. Poza tym nagłówki na stronie oddzielają odrębne treści. Dobrze jest stosować hierarchię nagłówków: ma to szczególne znaczenie dla osoby z dysfunkcją wzroku, albowiem umożliwia jej zbudowanie wyobrażenia układu strony z podziałem na ważne i mniej ważne treści. Poza tym czytnik ekranu pozwala na odczytywanie samych nagłówków i przechodzenie od jednego do drugiego z wykorzystaniem skrótu klawiszowego. Nagłówki stanowią też znakomity punkt orientacyjny w tekście.

Omówiliśmy warstwę prozodyczną języka i to, w jakim stopniu właściwe posługiwanie się elementami prozodycznymi umożliwia odbiorcy zrozumienie tekstu zgodnie z intencją nadawcy. Jest to szczególnie ważne w przypadku języka mówionego (w tym także czytanego na głos). Przedstawione informacje nie wyczerpują problemu, bowiem z punktu widzenia projektanta istotne jest pytanie, jak skonstruować interfejs, by umożliwić osobie z dysfunkcją wzroku słuchowy odbiór interfejsu tekstowego. Innymi słowy, jak skonstruować interfejs, który ułatwi użytkownikowi percepcję informacji przekazywanej za pomocą dźwięków mowy. Na tak postawione pytanie można odpowiedzieć, że **kluczową rolę w rozumieniu informacji językowych pełni semantyka (czyli znaczenie)**.

1.6. JAK BADAĆ NATURALNE POWIĄZANIA SEMANTYCZNE?

Zakładamy, że przetwarzanie semantyczne jest niezależne od tego, czy człowiek potrafi percypować wzrokowo, czy też nie. Naszą tezę opieramy na następujących przesłankach: osoby z dysfunkcją wzroku potrafią płynnie porozumiewać się za pomocą języka naturalnego z osobami widzącymi. Poza tym osoby z dysfunkcją wzroku rozumieją tekst, zatem zakładamy, że przetwarzanie semantyczne jest porównywalne z tym, jakie występuje u osób widzących, bowiem osoby widzące i niewidzące posługują się tym samym językiem.

Naszym zdaniem niewątpliwie tym, co różni osoby widzące od osób z dysfunkcją wzroku, jest proces materializacji znaczeń, tj. proces przyporządkowania znaczeniu konkretnego elementu rzeczywistości (zob. rozdział 2.1). Różnice te wynikają oczywiście z różnych możliwości percepcyjnych i różnych

indywidualnych obrazów świata, nie są to jednak różnice w zakresie przetwarzania semantycznego.

Współcześnie rozwijające się nauki poznawcze pozwalają zgłębiać naturalne uwarunkowanie języka. Dzięki wykorzystaniu metod pozyskiwania materiału językowego, jakie daje nam psycholingwistyka, możemy zbudować i zweryfikować na drodze eksperymentalnej modele lingwistyczne. Możemy przy tym korzystać z osiągnięć współczesnej techniki, a w szczególności informatyki, która okazuje się bardzo pomocna podczas gromadzenia i porządkowania dużej liczby danych. W ten sposób otwierają się nowe możliwości badań językoznawczych koniecznych do zgłębiania struktury języka naturalnego, co pozwala stworzyć propozycję modelu kognitywnego zbliżonego do funkcjonowania języka w ludzkim umyśle.

Naszym zdaniem dobrą metodą badawczą, dostarczającą materiał lingwistyczny umożliwiający najpełniejszy opis przetwarzania semantycznego w ludzkim umyśle, jest eksperyment polegający na badaniu swobodnych skojarzeń słownych. W historię naukową badań swobodnych skojarzeń słownych wpisali się Grace H. Kent i Aaron J. Rosanoff, którzy po raz pierwszy przeprowadzili takie badania w 1910 roku w Stanach Zjednoczonych. Przebadali oni 1000 osób o różnym wykształceniu i wykonujących różne zawody, sprawdzając powszechność skojarzeń dla 100 wyrazów (głównie rzeczowników i przymiotników). Lista Kent-Rosanoffa została przetłumaczona na kilka języków i w ten sposób badano skojarzenia w innych krajach. W Polsce także przeprowadzono badania swobodnych skojarzeń słownych – dokonała tego w latach 1964–1965 Ida Kurcz²⁹. W swym eksperymencie badała ona częstość odpowiedzi na słowo-hasło z podziałem na kobiety i mężczyzn. Była to lista skojarzeniowa dla 100 haseł z listy Kent-Rosanoffa, przy czym zebrano odpowiedzi od 1000 studentów Uniwersytetu Warszawskiego i Politechniki Warszawskiej. Kompletne wyniki oraz krótki artykuł wprowadzający opublikowano w VII tomie „Studiów Psychologicznych” z 1967 roku. Prawdopodobnie ze względów technicznych nie brano pod uwagę indywidualnego czasu reakcji³⁰. Słowa-hasła były podawane wszystkim badanym w tym samym czasie.

²⁹ I. Kurcz, *Polskie normy powszechności skojarzeń swobodnych na 100 słów z listy Kent-Rosanoffa*, „Studia Psychologiczne” 1967, t. VIII, s. 122–255.

³⁰ Czas reakcji jest ważnym kryterium uwzględnianym w psycholingwistyce; przyjmuje się, że im krótszy czas reakcji, tym w umyśle człowieka powiązania między słowami są silniejsze.

Później w Stanach Zjednoczonych przeprowadzano badania porównawcze na grupach liczących 500 osób, uznano bowiem, że taka liczba zebranych odpowiedzi stanowi już wystarczającą podstawę do solidnych analiz lingwistycznych³¹. Aktualnie badania swobodnych skojarzeń słownych przeprowadziliśmy wśród 560 studentów Wydziału Zarządzania i Komunikacji Społecznej UJ. Skorzystaliśmy z listy skojarzeniowej Kent-Rosanoffa w polskiej wersji językowej, którą posłużyła się I. Kurcz. Po wstępnej analizie lingwistycznej zdecydowaliśmy, by do opisanego eksperymentu przeznaczyć te wyrazy z listy Kurcz, które pod względem gramatycznym są rzeczownikami, a dodatkowo wybraliśmy pięć najczęstszych skojarzeń do tych rzeczowników. Jeśli dane skojarzenie pojawiało się dla różnych haseł, np. *biały* dla *doktor*, *ser*, *owca*, to wyraz skojarzony jako hasło w naszym eksperymencie wystąpił tylko raz. W rezultacie materiał obejmował 60 rzeczowników z listy Kent-Rosanoffa oraz 260 wyrazów reprezentujących najczęstsze skojarzenia dla haseł z tej listy w polskiej wersji językowej. Zatem nie jest to dokładne powtórzenie eksperymentu sprzed 45 lat: grupa badanych pozostaje taka sama jak wówczas, czyli stanowią ją studenci, zmieniły się natomiast warunki przeprowadzonego eksperymentu, a także metoda analizy wyników.

Opisując w największym skrócie metodę naszego eksperymentu, powiemy, że był on prowadzony za pomocą systemu komputerowego, który wyświetlał na monitorze każdego uczestnika eksperymentu słowa-hasła i zapisywał wprowadzane za pomocą klawiatury skojarzenie. Przy takiej organizacji eksperymentu musieliśmy empirycznie ustalić parametry czasowe, które pozwalają badanemu przeczytać wyraz i udzielić odpowiedzi³². Ważnym elementem eksperymentu było też takie sformułowanie instrukcji dla badanych, by była

³¹ D.S. Palermo, J.J. Jenkins, *Word Associations Norms: Grade School through College*, University of Minnesota Press, Minneapolis 1964.

³² Ponieważ niektórzy badani szybciej wypełniają pole swoim skojarzeniem i w rezultacie szybciej kończą swój test, zastanawialiśmy się nad parametrami czasowymi, czyli jak długo można pozwolić badanemu wpisywać skojarzenie. Nie może to być czas zbyt krótki, by badany się nie zdenerwował, zniecierpliwił i zniechęcił do wypełnienia testu. Ale też nie może być za długi, bo wówczas nie będzie to pierwsze skojarzenie, tylko kolejne. Czas ten ustalono empirycznie, analizując zachowanie i opinie pierwszej grupy badanych w badaniu pilotażowym, którą stanowili studenci studiów informatyczno-humanistycznych (Elektroniczne Przetwarzanie Informacji w Katedrze Lingwistyki Komputerowej WZiKS UJ), gdyż przyjęto, że badani mają podobny trening lingwistyczny i podobną biegłość w korzystaniu z klawiatury komputera. Aktualnie przeprowadzone badania studentów różnych kierunków na Wydziale Zarządzania i Komunikacji Społecznej UJ potwierdziły, że czas ustalony na przeczytanie i udzielenie odpowiedzi przez badanego został określony prawidłowo i zdaniem badanych był wystarczający.

jasna, zwięzła i zawierała najistotniejsze informacje³³ (w tym np. dotyczące czcionki – wielkości liter, użycia polskich liter).

W wyniku eksperymentu uzyskaliśmy zbiór par: wyraz hasłowy – wyraz skojarzony, gdzie skojarzenie jest charakteryzowane przez liczbę odpowiedzi i parametry czasowe związane z udzielaniem odpowiedzi (czas reakcji uczestnika). Ze względu na specyfikę pozyskania materiału nie mamy do dyspozycji zdań lub tekstów, a jedynie pary: wyraz hasłowy – wyraz skojarzony. Interpretację wyników musiały poprzedzić założenia. Przyjmujemy, że liczba skojarzeń pozwala wnioskować, iż konkretne powiązanie jest stosunkowo mocne (bardzo częste) lub stosunkowo słabe (rzadkie), jednak chcąc określić jakość powiązania, a więc relację zachodzącą pomiędzy wyrazem hasłowym a wyrazem skojarzonym, możemy odwołać się do słownika umysłowego (np. *lekarz* : *fartuch*), kontekstu językowego lub kulturowego, który motywuje skojarzenie (np. *pies ogrodnika* dla pary *pies* : *ogrodnik*, czy *Chata wuja Toma* dla pary *chata* : *Tom*). Jeśli jednak nie znajdziemy żadnego z wymienionych kontekstów dla konkretnej pary, powiedzmy *stół* : *ciocia*, wówczas przyjmujemy, że skojarzenie jest umotywowane pragmatycznie (pozajęzykowo).

Poniżej przedstawiamy wyniki skojarzeń uzyskanych w naszym eksperymencie dla wyrazu hasłowego „dom”.

Układ według częstości skojarzeń do wyrazu-hasła „dom” (zachowano oryginalną pisownię)

Skojarzenie	Ilość	Średni czas [s]	Min. czas [s]	Max. czas [s]
rodzinny	101	3.79	2.04	10.11
mieszkanie	87	3.58	2.09	8.50
rodzina	69	3.93	2.17	10.91
spokój	18	3.97	2.45	8.43
ciepło	18	3.75	2.27	5.54
ogród	16	4.14	2.51	6.13
bezpieczeństwo	13	5.91	3.47	11.00
mój	12	3.29	1.86	5.12
mama	11	3.18	1.43	5.93
pokój	9	3.91	2.60	6.61
dach	9	3.83	2.63	6.19
duży	9	3.20	2.05	7.34
budynek	7	4.42	2.57	7.35
pusty	6	3.07	1.82	4.14

³³ Z badań psychologicznych oraz psycholingwistycznych wiadomo, że instrukcja może decydować o uzyskanych w eksperymencie wynikach.

Skojarzenie	Ilość	Średni czas [s]	Min. czas [s]	Max. czas [s]
zły	5	2.36	2.17	2.64
własny	4	3.83	2.39	5.52
ognisko	4	3.43	2.99	3.87
wielki	3	2.98	2.17	4.39
miłość	3	3.85	2.97	4.58
rodzice	3	3.95	3.37	4.54
mieszkać	3	3.37	2.25	5.13
chata	3	4.37	2.93	6.85
wieś	3	3.69	2.97	4.94
mieszkanie	3	2.94	2.29	3.91
stół	3	4.04	3.64	4.69
dom	3	1.91	1.17	2.58
komin	3	2.94	2.25	3.37
matka	3	3.00	1.76	4.51
własny	2	3.18	1.88	4.49
jednorodzinny	2	4.48	4.03	4.93
dziecka	2	2.36	1.91	2.82
twierdza	2	4.07	3.91	4.23
daleko	2	3.48	3.45	3.51
ciepło	2	5.00	3.12	6.89
schronienie	2	4.66	4.38	4.95
Tom	2	2.74	2.20	3.27
cegła	2	3.27	2.92	3.61
domek	2	4.00	3.63	4.36
szczęście	2	7.19	7.12	7.27
ciepły	2	3.23	2.97	3.49
azyl	2	5.19	4.89	5.48
dobro	1	5.00	5.00	5.00
Tatry	1	3.67	3.67	3.67
z cegły	1	7.23	7.23	7.23
przytulny	1	4.08	4.08	4.08
podwórko	1	3.62	3.62	3.62
życie	1	3.87	3.87	3.87
cegły	1	2.95	2.95	2.95
Ryki	1	8.90	8.90	8.90
ja	1	1.92	1.92	1.92
ściana	1	3.27	3.27	3.27
własne miejsce	1	8.94	8.94	8.94
mieszkalny	1	3.00	3.00	3.00
wspólny	1	7.12	7.12	7.12
spanie	1	7.60	7.60	7.60
złoto	1	2.56	2.56	2.56

Skojarzenie	Ilość	Średni czas [s]	Min. czas [s]	Max. czas [s]
ogromny	1	4.70	4.70	4.70
opieki	1	4.22	4.22	4.22
ostoja	1	4.80	4.80	4.80
na wsi	1	3.61	3.61	3.61
drzwi	1	3.27	3.27	3.27
wadowice	1	3.82	3.82	3.82
stan	1	2.43	2.43	2.43
CHAŁUPA	1	3.45	3.45	3.45
okno	1	5.31	5.31	5.31
ikea	1	11.06	11.06	11.06
rmieszkać	1	5.86	5.86	5.86
cztery ściany	1	3.65	3.65	3.65
na piasku	1	3.81	3.81	3.81
mieszaknie	1	3.17	3.17	3.17
przystań	1	4.14	4.14	4.14
garaż	1	4.39	4.39	4.39
mam	1	1.98	1.98	1.98
pięny	1	2.61	2.61	2.61
pokoje	1	4.44	4.44	4.44
spokoójny	1	3.92	3.92	3.92
piękny	1	2.35	2.35	2.35
otwarty	1	3.84	3.84	3.84
gran torino	1	5.36	5.36	5.36
ściany	1	8.14	8.14	8.14
spokojny	1	4.45	4.45	4.45
kompost	1	5.93	5.93	5.93
pies	1	3.64	3.64	3.64
wartość	1	7.36	7.36	7.36
zamieszkany	1	2.96	2.96	2.96
obiad	1	2.29	2.29	2.29
pełny	1	4.33	4.33	4.33
cisza	1	3.33	3.33	3.33
soplicowo	1	5.64	5.64	5.64
Przemyśl	1	4.71	4.71	4.71
posiadłość	1	7.59	7.59	7.59
przytulny	1	5.09	5.09	5.09
ochrona	1	3.97	3.97	3.97
wież	1	6.29	6.29	6.29
święta	1	3.87	3.87	3.87
chatka	1	2.54	2.54	2.54
na brzegu morza	1	7.39	7.39	7.39
młot	1	3.31	3.31	3.31

Skojarzenie	Ilość	Średni czas [s]	Min. czas [s]	Max. czas [s]
ludzie	1	4.05	4.05	4.05
mieszkanje	1	2.57	2.57	2.57
wspomnienia	1	7.12	7.12	7.12
sad	1	5.76	5.76	5.76
drewniany	1	5.73	5.73	5.73
cichy	1	3.50	3.50	3.50
bezpieczeństwo	1	6.40	6.40	6.40
ładny	1	4.25	4.25	4.25
blok	1	2.52	2.52	2.52
bezpieczny	1	6.53	6.53	6.53
gość	1	7.16	7.16	7.16
ojczyzna	1	4.10	4.10	4.10
rodzinny	1	2.24	2.24	2.24
tęsknić	1	5.16	5.16	5.16
publiczny	1	2.99	2.99	2.99
kasa	1	2.28	2.28	2.28
niewiadoma	1	6.05	6.05	6.05
przuytulny	1	4.10	4.10	4.10
jedyny	1	2.77	2.77	2.77
mama tata	1	3.16	3.16	3.16

Jak widzimy, skojarzenia do wyrazu hasłowego „dom” układają się w dwa nurty znaczeniowe: dom jako obiekt fizyczny, ‘miejsce do mieszkania’ z typowymi dla niego atrybutami, oraz dom jako ‘miejsce rodziców’ widziane oczami dziecka, czyli dom rodzinny wraz z właściwymi dla niego osobami, więzami międzyludzkimi, emocjami. Przeważają skojarzenia tematyczne (głównie semantyczne). Co ciekawe, uzyskaliśmy porównywalną liczbę skojarzeń do obydwu znaczeń: *rodzinny* – 101 (najwięcej), *mieszkanie* – 90 (87 + 3 napisane z błędem literowym). Kolejne skojarzenia to *rodzina* (69), *spokój* (18) i *ciepło* (18), które możemy potraktować jako przymioty domu rodzinnego, czyli są to stany. Następny *ogród* (16) możemy powiązać ze znaczeniem domu jako miejsca do mieszkania. Kolejne skojarzenie *bezpieczeństwo* (13) wydaje się, że można powiązać z ‘miejszem jakim jest dom’, który daje poczucie bezpieczeństwa (gdyby chodziło o ‘miejsce do mieszkania’, pojawiłaby się forma *bezpieczny*). I dalej: *mój* (12) wskazuje na własność, lecz nie mamy pewności, o który dom chodzi, zaś *mama* (11) to osoba, która wyraźnie kojarzy się z domem rodzinnym. Ale już następne skojarzenia: *pokój* (9), *dach* (9), *duży* (9), *budynek* (7), *pusty* (6), *własny* (4) odsyłają nas do znaczenia domu

jako ‘miejsca do mieszkania’. W dalszej kolejności przeplatają się obydwa znaczenia: *ognisko* (4) wydaje się, że można interpretować jako ‘ognisko domowe’, czyli ‘miejsce rodziców’, *wielki* (3) może odnosić się do wielkości ‘miejsca do mieszkania’, *miłość* (3) może dotyczyć miejsca, w którym żyje para odczuwająca uczucie miłości, ale może też oznaczać ‘miejsce rodziców’, zaś skojarzenie *rodzice* (3) wskazuje jednoznacznie na znaczenie ‘miejsca rodziców’. Następne skojarzenia: *mieszkać* (3), *chata* (3), *wieś* (3), *stół* (3), *dom* (3), *komin* (3) odnoszą się do ‘miejsca do mieszkania’. Dalej pojawia się *matka* (3) jako osoba tworząca dom, ‘miejsce rodziców’, dalsze: *własny* (2), *jednorodzinny* (2) to skojarzenia do znaczenia ‘miejsce do mieszkania’. Występuje też skojarzenie *dziecka* (2), które może być skojarzeniem do ‘miejsce rodziców’, ale może też być członem nazwy ‘dom dziecka’. Następnie pojawiają się odpowiedzi *twierdza* (2), *azyl* (2), *szczęście* (2), *ciepło* (2), *daleko* (2), które można kojarzyć z pojęciem domu miejsca, gdzie wytwarzają się emocje, więzi, czyli ‘dom rodziców’. Wreszcie pojawia się materiał, z którego ‘miejsce do mieszkania’ zbudowano – *cegła* (2), i cecha takiego domu – *ciepły* (2), a także skojarzenie gramatyczne: *domek* (2), *Tom* (2)³⁴.

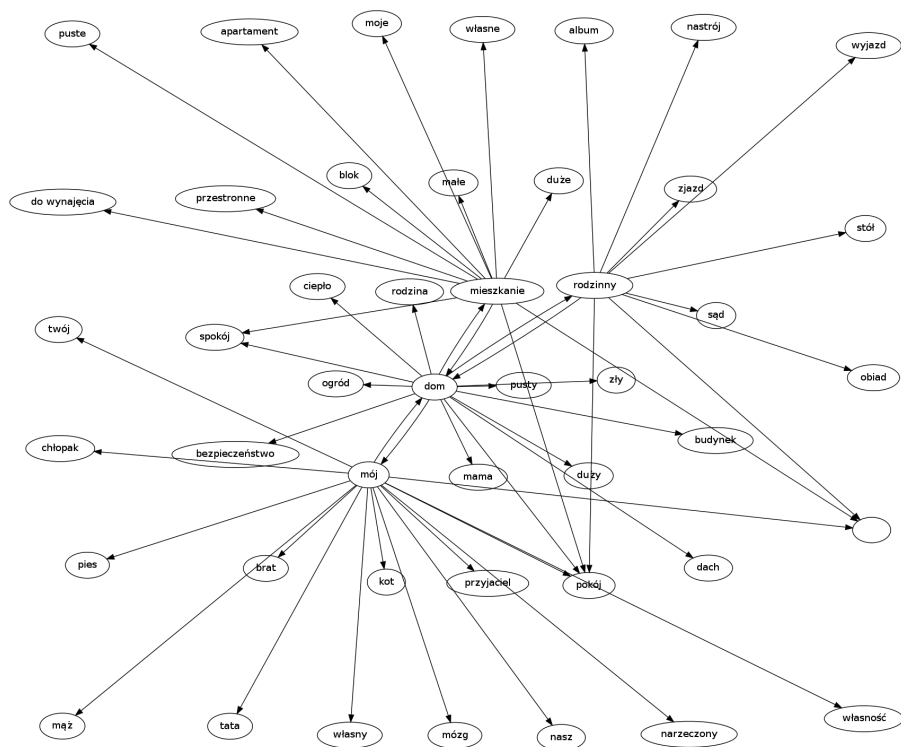
Podsumowując, widzimy, że obydwa znaczenia uzyskały porównywalną liczbę skojarzeń: ‘miejsce rodziców’ 247, ‘miejsce do mieszkania’ 191³⁵. Jednak nawet to zestawienie ilościowe niepoparte analizą lingwistyczną pokazuje, co dzieje się w sytuacji, gdy człowiekowi (odpowiadającemu skojarzeniem) nie podano kontekstu (zdania, tematu), dając tym samym możliwość, by odpowiedział pierwszym skojarzeniem, które przyszło mu na myśl. Dzięki odpowiedziom widzimy, że znaczenia w umyśle człowieka są silnie powiązane nawet bez jasno przedstawionego kontekstu. W takiej sytuacji zazwyczaj dominuje znaczenie najczęściej używane. W omawianym przez nas przykładzie widać, że żadne ze znaczeń wyrazu „dom” nie uzyskało wyraźnej, zdecydowanej przewagi. A zatem można sformułować tezę, że nadanie kontekstu będzie powodować aktywizowanie konkretnych znaczeń w umyśle użytkownika³⁶.

³⁴ Uzyskana w wyniku eksperymentu lista skojarzeń wykazuje dużą zbieżność z listą dla wyrazu ‘house’ w EAT: Edinburgh Associative Thesaurus, <http://www.eat.rl.ac.uk/>.

³⁵ Na uwagę zasługuje stabilność znaczeń „dom”, bowiem np. w utworach Jana Kochanowskiego najczęstsze są także oba wspomniane znaczenia. Zob. *Słownik polszczyzny Jana Kochanowskiego*, t. 1, red. M. Kucała, Instytut Języka Polskiego PAN, Kraków 1994.

³⁶ Ta teza wiąże się z częścią dotyczącą struktury tekstu. W psycholingwistyce znane jest zjawisko torowania, które mówi o tym, że kontekst decyduje o wyborze znaczenia wyrazu.

Inną kwestią są typy skojarzeń. Wspomniano już, że przeważają skojarzenia tematyczne, czyli semantyczne i pragmatyczne. Znacznie rzadziej pojawiają się skojarzenia gramatyczne, a jeszcze rzadziej dźwiękowe. Gdyby teraz rozszerzyć eksperyment i przebadać skojarzenia do wyrazów skojarzonych, uzyskamy kolejne listy, które zintegrowane pokażą, że skojarzenia odwzorowują sieć semantyczną przechowywaną w ludzkim umyśle. Fragment sieci odnoszącej się do wyrazu „dom” pokazuje rysunek, na którym widać wyrazy skojarzone bezpośrednio z wyrazem-bodźcem oraz wykryte w eksperymencie skojarzenia do wyrazów kojarzących się bezpośrednio z wyrazem-bodźcem. Strzałka → prowadząca od wyrazu-bodźca do wyrazu skojarzonego oznacza, że wyraz skojarzony został uzyskany w reakcji na wyraz-bodziec, np. *dom* → *mieszkanie*, *wielki* → *dom*. Skojarzenie zwrotne, polegające na tym, że dwa wyrazy wymiennie występują w roli bodźca i skojarzenia, np. *dom* : *mieszkanie* jest reprezentowane przez dwie strzałki. Dla zachowania czytelności rysunek przedstawia tylko wybrane, najczęstsze skojarzenia.



Rysunek 1. Fragment sieci semantycznej

Patrząc na rysunek, możemy wyciągać wnioski odnośnie do organizacji leksykonu w ludzkim umyśle. Widzimy, że słowa (symbole języka naturalnego) są z sobą powiązane. Najczęściej są to powiązania tematyczne, jednak jak pokazuje rysunek, powiązania te nie muszą być bezpośrednie. Może się też zdarzyć, że w ciągu skojarzeń brakuje jakiegoś ogniwa, gdyż odnosi się ono do kontekstu, którego nie mamy w słowniku. Jest to kontekst kulturowy, poznawczy, właściwy dla danego wieku czy środowiska. Niektóre z takich skojarzeń są stałe i mimo upływu lat oraz zmiany pokoleniowej pozostają niezmienione. Inne ulegają modyfikacji³⁷.

Nie ulega zatem wątpliwości, że pełna interpretacja wyniku eksperymentu wymagałaby jakościowej interpretacji powiązań. Niestety lista skojarzeń uzyskana w wyniku eksperymentu dostarcza danych ilościowych. Dane ilościowe pozwalają psychologom rozróżniać powiązania stałe (semantyczne) i chwilowe (pragmatyczne). Jednakże taka analiza nie jest wystarczająca, by opisać struktury ludzkiej pamięci semantycznej, nie pokazuje też rzeczywistych lingwistycznych zależności pomiędzy znaczeniami, których poznanie jest konieczne, by skonstruować model, z którego można wyprowadzić naturalne reguły konstrukcji interfejsu użytkownika. Chcąc tego dokonać, potrzebujemy językoznawczej klasyfikacji powiązań semantycznych. O tym traktuje rozdział 2.

BIBLIOGRAFIA

- Allan J., *HARD Track Overview in TREC 2003. High Accuracy Retrieval from Documents*, w: *Proceedings of the Twelfth Text Retrieval Conference (TREC 2003)*, NIST.
- Baumgärtner Ch., Menzel W., *Integrating a Model for Visual Attention into a System for Natural Language Parsing*, w: *Natural Language Processing and Cognitive Science*, red. B. Sharp, M. Zock, 9th International Workshop on Natural Language Processing and Cognitive Science, Wrocław 2012, s. 24–33.

³⁷ Autorka niniejszego rozdziału przeprowadziła w 2012 roku eksperyment lingwistyczny, badając swobodne skojarzenia słowne na grupie studentów Wydziału Zarządzania i Komunikacji Społecznej UJ. Dysponowała też wynikami badań przeprowadzonych przez I. Kurcz na grupie studentów Uniwersytetu Warszawskiego i Politechniki Warszawskiej w latach 1964–1965, dzięki czemu takie obserwacje i porównania są możliwe.

- Blumstein S.E., Stevens K.N., *Acoustic Invariance in Speech Production: Evidence from Measurements of the Spectral Characteristics of Stop Consonants*, „Journal of the Acoustical Society of America” 1979, nr 66, s. 1001–1017.
- Blumstein S.E., Stevens K.N., *Perceptual Invariance and Onset Spectra for Stop Consonants in Different Vowel Environments*, „Journal of the Acoustical Society of America” 1980, nr 67, s. 648–662.
- Buckley C., Robertson S., *Relevance Feedback Track Overview: TREC 2008*, w: *Proceedings of the Seventeenth Text Retrieval Conference (TREC 2008)*, NIST.
- EAT: Edinburgh Associative Thesaurus, <http://www.eat.rl.ac.uk/>.
- Eimas P.D., Corbit J.D., *Selective Adaptation of Linguistics Feature Detectors*, „Cognitive Psychology” 1973, nr 4, s. 99–109.
- Gaskell M.G., Marslen-Wilson W.D., *Integrating Form and Meaning: A Distributed Model of Speech Perception*, „Language and Cognitive Processes” 1997, nr 12, s. 613–656.
- Gatkowska I., *Diagnoza dyzartrii u dorosłych w neurologii klinicznej*, Wydawnictwo Uniwersytetu Jagiellońskiego, Kraków 2012.
- Gatkowska I., *Jak mózg rozpoznaje języki*, w: *Język polski: nowe wyzwania językoznawcze*, red. J. Dybiec, G. Szpila, Krakowskie Towarzystwo „Tertium”, Kraków 2010, s. 15–29.
- Gatkowska I., *Jak słowa łączą się z sobą w umyśle użytkowników*, referat na konferencji Tertium, Kraków 2012.
- Jassem W., *Biolingwistyka: wybrane zagadnienia*, w: *Foniatrya kliniczna*, red. A. Pruszevicz, Państwowy Zakład Wydawnictw Lekarskich, Warszawa 1992, s. 73–86.
- Kurcz I., *Polskie normy powszechności skojarzeń swobodnych na 100 słów z listy Kent-Rosanoffa*, „Studia Psychologiczne” 1967, t. VIII, s. 122–255.
- Marslen-Wilson W.D., *Function and Process in Spoken Word Recognition*, w: *Attention and Performance X: Control of Language Processes*, red. H. Bouma, D. Bouwhuis, Erlbaum, Hillsdale, NJ, s. 125–150.
- Marslen-Wilson W.D., *Functional Parallelism in Spoken Word Recognition*, „Cognition” 1987, nr 25, s. 71–102.
- Marslen-Wilson W.D., Welsh A., *Processing Interactions and Lexical Access during Word Recognition in Continuous Speech*, „Cognitive Psychology” 1978, nr 10, s. 29–63.
- Massaro D.W., *Speech Perception by Ear and Eye: A Paradigm for Psychological Inquiry*, Erlbaum, Hillsdale, NJ.
- Massaro D.W., *Testing between the TRACE Model and Fuzzy Logical Model of Speech Perception*, „Cognitive Psychology” 1989, nr 21, s. 398–421.
- Massaro D.W., Oden G.C., *Speech Perception: A Framework for Research and Theory*, w: *Speech and Language: Advances in Basic Research and Practice*, red. N.J. Lass, t. 3, Academic Press, New York 1980, s. 129–165.
- McGurk H., MacDonald J., *Hearing Lips and Seeing Voices*, „Nature” 1976, nr 264, s. 746–748.

- McClelland J.L., Elman J.L., *The TRACE Model of Speech Perception*, „Cognitive Psychology” 1986, nr 18, s. 1–86.
- Miller G.A., *Human Memory and the Storage of Information*, „IRE Transactions on Information Theory” 1956, 2 (3), s. 129–137.
- Miller G.A., Johnson-Laird P.N., *Language and Perception*, Belknap Press of Harvard University Press, Cambridge 1976.
- Minogue J., Jones M., *Haptics in Education: Exploring an Untapped Sensory Modality*, „Review of Educational Research” 2006, nr 76 (3), s. 317–348.
- Palermo D.S., Jenkins J.J., *Word Associations Norms: Grade School through College*, University of Minnesota Press, Minneapolis 1964.
- Palmer S., *Modern Theories of Gestalt Perception*, w: *Understanding Vision: An Interdisciplinary Perspective Readings in Mind and Language*, red. G.W. Humphreys, Blackwell, Oxford 1992, s. 39–70.
- Pisoni D.B., Nusbaum H.C., Luce P.A., Slowiaczek L.M., *Speech Perception, Word Recognition and the Structure of the Lexicon*, „Speech Communication” 1985, nr 4, s. 75–95.
- Posner M.I., Raichle M.E., *Images of Mind*, Scientific American Library, New York 1994.
- Rock I., Palmer S., *The Legacy of Gestalt Psychology*, „Scientific American”, December, s. 84–90.
- Słownik polszczyzny Jana Kochanowskiego*, t. 1, red. M. Kucała, Instytut Języka Polskiego PAN, Kraków 1994.
- Stevens K.N., *Toward a Model for Speech Recognition*, „Journal of the Acoustical Society of America” 1960, nr 32, s. 47–55.
- Stevens K.N., Halle M., *Remarks on Analysis by Synthesis and Distinctive Features*, w: *Models for the Perception of Speech and Visual Form*, red. W. Wathen-Dunn, MIT Press, Cambridge, MA 1967, s. 88–102.
- Szeląg E., *Mózgowe mechanizmy mowy*, w: *Mózg a zachowanie*, red. T. Górską, A. Grabowska, J. Zagrodzka, Wydawnictwo Naukowe PWN, Warszawa 2005, s. 489–524.
- Szeląg E., Pöppel E., *Temporal Perception: A Key to Understand Language*, „Behavioral and Brain Sciences” 2000, nr 23 (1), s. 53.