

BOŻYDAR L. J. KACZMAREK

*Instytut Psychologii, Uniwersytet Marii Curie-Skłodowskiej
Plac Litewski 5, 20 080 Lublin.
e-mail: Bkaczmar@sokrates.umcs.lublin.pl*

SŁUCH FONEMATYCZNY A PERCEPCJA WYPOWIEDZI SŁOWNYCH

SŁUCH FONEMATYCZNY

Współczesne definicje słuchu fonematycznego koncentrują się jedynie na wąskim wycinku tego zjawiska. Przede wszystkim podkreśla się jego znaczenie w percepcji wypowiedzi pomijając istotną rolę, jaką odgrywa ta umiejętność w prawidłowym wypowiedzianiu się. Warto więc przypomnieć poglądy pioniera fonologii, Jana Baudouina de Courtenay, wybitnego polskiego językoznawcy, którego twórcza działalność przypada na przełom XIX i XX wieku. Podał on bardzo przystępną definicję fonemu określając go jako wyobrażenie głoski w mózgu użytkownika danego języka (BAUDOIN DE COURTENAY 1984).

Jest to jednak definicja psychologiczna, dlatego też jego uczniowie i następcy, do których należał między innymi znany lingwista, Roman Jakobson, podjęli ambitną próbę sformułowania definicji uwzględniającej czysto językoznawcze kryteria. W tym bowiem okresie językoznawcy zaczęli unikać „psychologizowania”, a psychologowie przestali zajmować się językiem. Jest to typowa tendencja dla młodych dyscyplin naukowych, które nadmiernie dbają o „czystość” podejmowanych badań. Charakterystyczne jest także to, że w miarę dojrzewania owych dziedzin rośnie wiadomość o ograniczeń wynikających ze zbyt wąskiego ujmowania danego zagadnienia, co w rezultacie prowadzi do powstawania dyscyplin pogranicznych. Wystarczy wymienić psycholingwistykę, neurolingwistykę czy socjolingwistykę, których burzliwy rozwój obserwujemy w ostatnim okresie. Przyczynia się do tego fakt, iż wielu zjawisk językowych, psychologicznych czy neurologicznych, nie da się wyjaśnić z jednej tylko perspektywy.

Przekonuje nas o tym wspomniana wyżej próba sformułowania czysto językoznawczej definicji fonemu. Z reguły podaje się, że jest to

wiązka cech dystynktywnych, dodając przy tym, iż jest to twór abstrakcyjny w odróżnieniu od poszczególnych głosek (por. JAKOBSON i HALLE 1964). W rezultacie studenci często odtwarzają tę definicję w trakcie egzaminu, a sam fonem pozostaje dla nich niezrozumiałą abstrakcją. Dlatego też bardziej przydatne okazuje się określenie fonemu jako modelu głoski. Jak wiadomo, każdy model jest abstrakcją, która odzwierciedla rzeczywiste zjawisko w schematyczny i uproszczony sposób. W modelu tym uwzględnia się też jedynie te właściwości odpowiadającej mu głoski, które pozwalają odróżnić ją od innych głosek i to są właśnie owe cechy dystynktywne, czyli mówiąc bardziej prostym językiem, cechy różnicujące poszczególne głoski. W języku polskim są to na przykład różnice między *s* oraz *ś* (czyli twardą i miękką) lub *s* i *z* (czyli dźwięczną i bezdźwięczną).

Każdy model jest też swego rodzaju abstrakcją, co odpowiada klasycznej definicji językoznawczej. Owe abstrakcyjne (mentalne) modele realizowane są w postaci fizycznych znaków, czyli głosek w przypadku mowy oraz liter w piśmie. W trakcie słownej realizacji fonemów poszczególne głoski obejmują nie tylko cechy fonemiczne (dystynktywne), lecz także fonetyczne. Odzwierciedlają bowiem brzmienie i wysokość głosu mówiącego, ewentualną chrypkę, i tym podobne. Dlatego też językoznawcy podkreślają, iż cechy fonetyczne obejmują szerszy zakres dźwięków od cech fonemicznych, co ilustruje rycina 1.

Powyższy schemat stanowi jednak duże uproszczenie, gdyż w trakcie mówienia do głosek dodajemy nie tylko cechy fonetyczne (nieistotne dla komunikacji), lecz co gorsza obcinamy szereg cech składających się na cechy dystynktywne. Zjawisko to przedstawia rycina 2, w której trójkąt oznacza fonem a okrąg głoskę.

grupowania wyrazów w kategorii gramatyczne. Przejawia się to także w silnej skłonności dzieci do różnego typu rymowanek, w których rym bywa często ważniejszy od treści. Za przykład służyć może znana wszystkim wyliczanka: „ene due ryke fake...”.

3. Obszaru (w przypadku percepcji słuchowej znacznie ma długość lub czas trwania danego dźwięku) — im krócej trwa dane zjawisko akustyczne, tym wyraźniej wyróżnia się jako figura. W ten właśnie sposób wyróżniają się w naszej świadomości okrzyki, rozkazy i wiele emocjonalnie naładowanych wypowiedzi.

4. Zamykania — tendencja do wypełniania brakujących (niedosłyszanych elementów) wypowiedzi. Wystarczy tu wymienić tak zwany „phonetic restoration effect”. Zaobserwono bowiem, że jeśli badanym prezentuje się nagrane zdania, w których jedną z głosek zastąpiono tak zwanym czystym tonem, wówczas twierdzą oni, iż nie brakuje tu żadnego dźwięku.

5. Symetrii — odzwierciedla się w powtórzeniach i w różnego typu konstrukcjach porównawczych. Na przykład w przysłowiach: „Jaki Pan, taki kram”, „Jaka mać, taka nać”, „Jak sobie pościelesz, tak się wyśpisz”, i tym podobne.

6. Dobrej struktury — czyli inaczej zasady pregnancji — polegającej na tym, że bardziej znany „dobry” kształt wyodrębnia się łatwiej niż mniej znany lub nieregularny. Znajduje to swe odzwierciedlenie w takich stwierdzeniach, jak: „okrągła twarz” lub „kwadratowa szczeka”, chociaż ani twarz ani szczeka nie miewają ściśle geometrycznych kształtów.

W przypadku odbioru słuchowego zasada pregnancji wiąże się z selektywnością słuchania, która ma pierwszorzędne znaczenie dla porozumiewania się. Z reguły bowiem słuchamy wypowiedzi innych osób w sytuacjach dalekich od idealnych, gdyż zazwyczaj odbywa się to na tle różnego typu hałasów. Ujawnia się to szczególnie jaskrawo w różnego typu dysfunkcjach mózgowych, w następstwie których pojawia się często tzw. „coctail party phenomenon” (które po

polsku możnaby określić mianem zjawiska imienin). Utrudnia to zwłaszcza rehabilitację chorych z zaburzeniami mowy spowodowanymi uszkodzeniami mózgu. Otóż chory taki, jeżeli nawet nauczy się porozumiewać z terapeutą, gubi się, gdy mówi jednocześnie kilka osób. Przede wszystkim zaś nie jest w stanie zrozumieć, co mówi do niego znajdujący się przed nim rozmówca.

Wiąże się to z podkreślaną jeszcze przez Pawłowa selektywnością działania układu nerwowego. Zazwyczaj zdrowy człowiek z łatwością koncentruje się na wybranym rozmówcy nawet wówczas, gdy mówi wiele osób naraz. Z reguły też nie zapamiętujemy treści owych prowadzonych w tle rozmów. Nie jesteśmy nawet w stanie określić, czy były formowane w naszym ojczystym czy też obcym języku. Prowadzone badania wykazały ponadto, iż szczególnie silne oddziaływanie maskujące na odbiór głosek mają tony niskie oraz dźwięki, którym nie przypisujemy żadnego znaczenia. Przy czym taki sam negatywny wpływ miały tony czyste, jak i hałas (por. MILLER 1963).

Jednakże każdy z nas może się z łatwością przekonać, że w hałaśliwym otoczeniu łatwo zrozumieć nam konwencjonalne słowa powitania, dużo trudniej zaś jest uchwycić brzmienie nazwiska przedstawianej nam osoby. Dzieje się tak dlatego, że w konwencjonalnych wyrażeniach z łatwością uzupełniamy niedosłyszane elementy, trudno zaś to uczynić w przypadku nazwisk, gdyż ich właściwej formy nie możemy być pewni.

Warto również dodać, że takie czynniki jak wysokość i brzmienie głosu, jego intensywność oraz tempo mówienia pomagają nam wyodrębnić głos mówiącego. Wskazuje to na znaczącą rolę prawej półkuli mózgu w odbiorze wypowiedzi słownych. Rola tej półkuli wiąże się również z jej funkcjonalnym znaczeniem dla całościowego odbierania określonego materiału, co ma przecież decydujące znaczenie dla zrozumienia przekazywanych nam treści.

ROZUMIENIE WYPOWIEDZI SŁOWNYCH

Podsumowując dotychczasowe rozważania należy stwierdzić, że nie sprawdził się model zakładający, iż odbieranie wypowiedzi polega na stopniowym przechodzeniu od analizy poszczególnych dźwięków do ich syntezy, która prowadzi do zrozumienia poszczególnych słów, a następnie zdań. Znana angielska psycholingwistka, ARCHISON (1991), stwierdza wręcz, iż ów model biernej sekretarki nie znajduje uzasadnienia z następujących powodów:

1. Szybkość mówienia uniemożliwia „zapisywanie” i „dopasowywanie” poszczególnych dźwięków do siebie.

2. Nie istnieją stałe odpowiedniki akustyczne fonemów.

3. W trakcie mówienia jedna głoska stopniowo przechodzi w drugą. Oznacza to, że odcyfrowywanie dźwięków mowy musi być procesem czynnym a nie biernym. Zgodnie ze współczesnymi danymi przyjmuje się więc model zstępu-

jący (top-down) zakładający przechodzenie „od ogółu do szczegółu” w przeciwieństwie do tradycyjnego modelu wstępującego (bottom-up), w którym zakładano stopniowe odcyfrowywanie znaczenia danej wypowiedzi dzięki przechodzeniu „od szczegółu do ogółu” (por. KRASOWICZ 1997). Zakłada się tutaj, iż słuchacz nie ogranicza się do stopniowego dekodowania odbieranych wypowiedzi, lecz często stawia hipotezy dotyczące odbieranych treści, wyprzedzając niejako ich rzeczywiste zrozumienie. Tylko w ten sposób wytłumaczyć można niezwykłą szybkość odbioru przekazywanych informacji.

Jak bardzo powolny i mało efektywny jest czysto linearny proces odcyfrowywania znaczenia tekstów przekonuje nas obserwacja sposobu czytania przez małe dzieci, które mozolnie składają z liter poszczególne wyrazy, by przy końcu zdania zapomnieć o wcześniej odkodowanych wyrazach. Zaś wprawny czytelnik wyprzedza wzrokiem czytany fragment tekstu z łatwością ujmując całą jego treść. Jednakże ten sposób czytania wymaga nie tylko umiejętności językowych, lecz także pewnej wiedzy o zawartych w danym tekście treściach. Dlatego też łatwiej jest nam czytać literaturę piękną niż teksty naukowe, które wymagają powrotu do stopniowego dekodowania znaczenia czytanych paragrafów.

Podobna sytuacja zachodzi przy słuchowym odbiorze wypowiedzi. Efektywne słuchanie wymaga nie tylko znajomości reguł językowych,

lecz także wiedzy o otaczającej nas rzeczywistości oraz doświadczenia, które nabywamy w różnego typu interakcjach z innymi ludźmi. W związku z powyższym wprowadzono pojęcie kompetencji komunikacyjnej, które uwypukla znaczenie czynników społecznych w porozumiewaniu się.

Znajduje to odzwierciedlenie także w strukturze wypowiedzi, które uznajemy za bardziej naturalne od innych. Przede wszystkim za takie uważamy zazwyczaj raczej czynną a nie bierną formę zdań. Bardziej typowe jest bowiem oddziaływanie istoty żywej na martwy przedmiot. Każdy z nas uzna więc za zręczniejszą konstrukcję „Marek podziwia tę scenerię” od „Ta sceneria jest podziwiana przez Marka”, mimo że obydwie te zdania wyrażają tę samą myśl.

Zawsze też „umieszczamy” mniejszy przedmiot w odniesieniu do większego. Stąd za prawidłowe uznawane jest raczej zdanie pierwsze z przytoczonych niżej par: „Piłka toczy się po trawie”, a nie „Po trawie toczona jest piłka”, a także „Miska znajduje się blisko ściany” zamiast „Ściana znajduje się blisko miski”. Powiemy też raczej „Krzesło stoi z lewej strony stołu”, a nie „Stół stoi z prawej strony krzesła”. Uznajemy również, iż to mniejszy przedmiot znajduje się nad, lub pod większym. Z tego też powodu mówimy: „Obraz znajduje się nad kanapą”, a nie „Kanapa znajduje się pod obrazem” lub „Mucha jest pod obrazem”, nie zaś „Obraz jest nad muchą”.

INTERPRETACJA WYPOWIEDZI

Istotną, choć z reguły niedocenianą rolę, odgrywa też słuch fonematyczny w segmentacji wypowiedzi, czyli wyodrębnianiu mniejszych elementów (syłab i wyrazów) w ciągu słownym. Wielu ludzi nie uświadamia sobie bowiem, że nie wypowiadamy poszczególnych wyrazów, lecz nieprzerwany ciąg dźwięków, które dzielimy lub łączymy w określone całości. Umożliwia to odmienną od zamierzonej interpretację intencji nadawcy, co prowadzi do wielu często nieuświadomianych sobie przez obie strony nieporozumień. Ilustruje to następujący przykład podany przez znanego angielskiego matematyka i pisarza, LEWISA CARROLA (1994): „— O, goni smutny i długi mnie los, smutny i długi jak ta opowieść! — powiedziała Myszka, zwracając się do Alicji i wzdychając. — Ogon i smutny, i długi? — powiedziała Alicja spoglądając ze zdumieniem na jej ogon.” (s. 28–29). Przykład ten uwypukla również fakt, że czynnikiem decydującym o tak zwanej segmentacji tekstów (czyli wyodrębnianiu poszczególnych wyrazów) jest

znaczenie nadawane danej wypowiedzi, w mniejszym zaś stopniu jej cechy formalne. Tłumaczy to trudności, z jakimi borykali się językoznawcy tradycyjnie skupiający swą uwagę na budowie gramatycznej wypowiedzi. Dotyczy to zwłaszcza tak zwanej strukturalizmu językoznawczego, w którym koncentrowano się głównie na stosunkach zachodzących między elementami języka rozumianego jako system, czyli hierarchicznie uporządkowana całość.

Tego typu nieporozumienia występują również znacznie częściej w przebiegu naszych codziennych interakcji, niż to się zakłada. W przeciwnym razie zbędni byłiby psychoterapeuci, którzy często wyjaśniają zwaśnionym współmałżonkom rzeczywiste intencje ich partnera. Błędne odczytywanie intencji drugiej osoby prowadzi do wielu nieporozumień, zazwyczaj zapominamy bowiem, że słuchacz interpretuje odbierane przez niego wypowiedzi przez pryzmat swych upodobań i uprzedzeń. Stanowi to swoisty filtr emocjonalny, który decyduje często na-

dawaniu swoistego znaczenia nawet niewinnym stwierdzeniom. Nawiasem mówiąc, płat skroniowy odgrywający podstawową rolę w percepcji słuchowej ma ściśle anatomiczne powiązanie z układem limbicznym, który wiąże się właśnie z regulacją stanów emocjonalnych.

Zresztą nie może być inaczej, gdyż proste dekodowanie ciągu wymawianych dźwięków mowy (głosek) jest po prostu niemożliwe, chociażby ze względu na mechanizmy artykulacji. Fonetycy już dawno stwierdzili, że w toku wypowiedzania się poszczególne głoski ulegają znacznym przekształceniom z uwagi na konieczność przystosowywania się aparatu mowy do artykułowania ciągu dźwięków. Mówimy wówczas o zjawisku koartykulacji, czyli wzajemnym oddziaływaniu głosek na siebie. Najszerzej opisywano zjawisko tak zwanej antycypacji artykulacyjnej, czyli przygotowywania się narządów mowy do wymówienia następnej głoski. Wybitny polski językoznawca Tytus Benii zwracał uwagę, iż może to dotyczyć nawet głosek oddzielonych od poprzedzającej kilkoma innymi. Wystarczy porównać w lustrze wymowę wyrazów: *wstrętny*, w *strumieniu*, w *Stryju* oraz w *Stradomiu*, by przekonać się, że do samogłoski *ę*, *u*, *y* oraz *a* upodabniają się wszystkie cztery poprzedzające je spółgłoski — *fstr* (patrz KACZMAREK 1986).

Prowadzi to nie tylko do upodobnień ciągu wypowiedzianych głosek, lecz wręcz uproszczeń wypowiedzianych słów. Stąd na przykład wyraz *sześćset* z reguły wymawiany jest jako *sześset*, *szejset* lub wręcz *szeset*. Wiele z tych uproszczeń znajduje z czasem znajdując swe odbicie w piśmie. Za przykład służyć może wyraz *wszystko*, bardzo często wymawiany — *wszysko* oraz pokrewny mu wyraz *wszyscy*, w którym głoska *t* uległa zanikowi.

Jeżeli dodamy do tego upodobnienie pod względem dźwięczności, na przykład *stóg siana* — wymawiany jako *stuksiana* oraz *jak dobrze* artykułowane jako *jagdobre*, a także upodobnienie pod względem artykulacji, w wyniku którego wyraz *pański* przekształca się w *pajski* (z nosowym *j*) dziwnym staje się fakt, że w ogóle jesteśmy w stanie się porozumiewać. Również z akustycznego punktu widzenia poszczególne

tony nakładają się na siebie i tylko dzięki selekcyjnemu działaniu analizatora słuchowego odbieramy poszczególne głoski, jako oddzielne dźwięki.

Warto również podkreślić, że wielu z powyższych zasad użytkownicy danego języka w ogóle sobie nie uświadamiają, chociaż wykorzystują je w codziennych rozmowach. Dlatego też większość Polaków uważa, iż w słowie *chleb* końcowa głoska brzmi *b* a nie *p*. Wielu też twierdzi, że nie zna zasad gramatyki ojczystego języka, mimo że wykorzystują je codziennie formułując własne wypowiedzi i słuchając wypowiedzi innych osób.

Dla uzmysłowienia czytelnikowi faktu, iż potrafi on bez trudu dokonać oceny, czy dana wypowiedź jest zgodna z regułami jego ojczystego języka, przytoczę wierszyk znany każdemu polskiemu dziecku: „Idzie rak nieborak, jak uszczypnie, będzie znak”. Jeżeli w wierszu tym w każdym wyrazie początkowe spółgłoski zastąpimy literą *k*, uzyskamy wówczas następujące zdanie:

1. Ikie kar kieborak, kak ukczyknie, kędzie knak.

Można też dokonać głębszych transformacji wymieniając wszystkie spółgłoski literą *m*:

2. Imie mam miemomam, mam ummymmie, męmmie mmam.

Mimo, iż powstałe zdanie nie jest zrozumiałe, każdy Polak zgodzi się, że brzmi ono po polsku. Podobnie jak nieco nonsensowna konstrukcja powstała w wyniku zamiany wyrazów:

3. Idzie rak niecnota, jak ugryzie, będzie słota.

Inaczej wszakże przedstawia się sprawa w przypadku, gdy poprzedzamy poszczególne litery i spróbujemy odczytać omawiany wierszyk wspak:

4. Eidzi kar karobein, kaj einpyczszu, eidżeb kanz.

Tego typu „zdanie” sprawia wrażenie struktury obcojęzycznej, bowiem złamane zostały zasady łączenia elementów języka w większe całości. Potwierdza to ponadto założenie, że zasady te znane są każdemu mówiącemu po polsku, lecz z reguły nie są w pełni uświadamiane.

UWAGI KOŃCOWE

Na zakończenie warto więc przytoczyć słowa wybitnego etologa KONRADA LORENZA (1977), który wskazuje na istnienie fizjologicznych mechanizmów stałości. Właśnie owe mechanizmy, czy też wyróżnione przeze mnie modele, pozwalają nam zachować orientację w złożonym i

zmiennym świecie otaczających nas zjawisk. Równie istotne jest zwrócenie przez niego uwagi, że daremne jest szukanie w osobnych, niezależnie funkcjonujących podsystemach tych właściwości systemowych, które zaistnieją dopiero na wyższym poziomie integracji. Sytuacja

taka zachodzi właśnie w przypadku percepcji słuchowej. Stąd ograniczenie się jedynie do analizy systemu fonologicznego (dotyczącego dźwięków mowy) niewiele mówi nam o procesie percepcji słowej. W procesie tym bowiem, decydującą rolę odgrywa ogólna wiedza słuchacza, która umożliwi mu właściwe odcyfrowywanie treści przekazywanych przez mówiącego.

Warto też uwypuklić niedocenianą często, lecz jakże istotną, funkcję słuchu fonematycznego umożliwiającą segmentację słyszanych wypowiedzi. Jak już wspominałem, wielu ludzi nie uświadamia sobie, iż nie wypowiadamy poszczególnych słów lecz całe ciągi dźwięków i jedynie słuchacz znający dany kod językowy

potrafi w tym ciągu wyodrębnić poszczególne wyrazy. Wystarczy posłuchać dowolnego (nieznanego sobie) języka prze radio, by się o tym przekonać. Decydującym czynnikiem pozostaje jednak treść danej wypowiedzi. Od jej znaczenia zależy bowiem, czy uznamy dane wypowiedzenie za jeden czy dwa wyrazy. Na przykład tak samo brzmiący szereg dźwięków *dowody* uznamy za jeden wyraz w zdaniu „Czy są na to dowody?“, a za dwa w wypowiedzi „Janek skończył do wody”. Ma to ścisły związek z „zegarem mózgowym”, który stanowi swego rodzaju mechanizm umożliwiający odpowiednie integrowanie odbieranych bodźców słuchowych (por. SZELAĞ 1995).

PHONEMIC HEARING AND PERCEPTION OF VERBAL UTTERANCES

Summary

The role of phonemic hearing in perceiving verbal utterances is discussed. First of all, the characteristics of the phonemic hearing itself is given. It is pointed out that the speaker not only adds irrelevant phonetic features to the pronounced sounds to be identified at the same time but cuts off the features that are essential for a given phoneme. Most important, however, is the often neglected fact that the phonemic hearing enables both distinguishing of individual speech sounds and segmentation of utterances into smaller linguistic units. The listener does it in accordance with the linguistic rules that are stored in his/her mind, and uses these rules in an automatic way without being

aware of them. On the other hand, the appropriate segmentation is also provided for due to the understanding of the meaning of a given message by the listener, in which his general knowledge plays an important role.

Auditory perception is processed in accordance with the rules of human cognition which were best described by Gestalt psychology. Most significant are the rules of constancy and grouping. Grouping of separate elements on the basis of the laws of proximity, similarity, area, closure, symmetry, and good structure enables us to reconstruct speech sounds which are blurred due to the process of phonetic co-articulation as well as by surrounding noises.

LITERATURA

- ARCHISON J., 1991. *Ssak, który mówi*. PWN, Warszawa.
- BAUDOUIN DE COURTENAY J. N., 1984. *O języku polskim*. PWN, Warszawa, str. 453.
- CARROL L., 1994. *Przygody Alicji w krainie czarów*. Wrocław, Wydawnictwo Jasiończyk.
- JAKOBSEN R., HALLE M., 1964. *Podstawy języka*. Ossolineum, Wrocław.
- KACZMAREK B. L. J., 1986. *Wzrokowa percepcja wypowiedzi słownych*. UMCS, Lublin.
- KACZMAREK B. L. J., 1995. *Mózgowa organizacja mowy*. Polskie Towarzystwo Neurolingwistyczne. AGH-AD, Lublin.
- KACZMAREK B. L. J., 1998. *Mózg, język, zachowanie*. UMCS, Lublin.
- KRASOWICZ G., 1997. *Język, czytanie, dysleksja*. Polskie Towarzystwo Neurolingwistyczne. AGH-AD, Lublin.
- KWIATKOWSKA A., 1997. *The Visuo-spatial Determinants of Natural Language*. Wydawnictwo Uniwersytetu Łódzkiego, Łódź.
- MILLER G. A., 1963. *Language and Communication*. McGraw-Hill, New York.
- LORENZ K., 1977. *Odwrotna strona zwierciadła*. PIW, Warszawa.
- SPELRLING A., 1995. *Psychologia*. Zysk i S-ka, Poznań.
- SZELAĞ E., 1995. *Neurobiologiczne podłoże mowy człowieka*. Kosmos 45, 179–200.