

SYLWESTER ORZECZOWSKI, SŁAWOMIR WACEWICZ,  
PRZEMYSŁAW ŻYWICZYŃSKI

## PROBLEM 'ZMIANY MODALNOŚCI' W HIPOTEZACH PIERWSZEŃSTWA GESTÓW W EWOLUCJI JĘZYKA: W STRONĘ HIPOTEZ MULTIMODALNYCH<sup>1</sup>

### 1. WPROWADZENIE

#### 1.1. TŁO BADAWCZE – EWOLUCJA JĘZYKA

Mimo istniejącego jeszcze pod koniec XX wieku przekonania, że pochodzenie języka nie może być problemem rozpatrywanym naukowo (Campbell, 1998; Fisiak, 1985), obecnie *ewolucja języka*, czyli badania nad filogenetycznymi źródłami i rozwojem zdolności językowej, jest dobrze ugruntowaną, rozpoznawalną i dynamicznie rozwijającą się dziedziną nauki. Według danych z *ISI Web of Science* liczba publikacji na temat ewolucji języka w latach 1990–1999 zwiększyła się dziesięciokrotnie, natomiast w latach 2000–2009 – trzykrotnie w stosunku do poprzedniej dekady. Dostępne są już opracowania podręcznikowe (np. Johansson, 2005; Hurford, 2007) i encyklopedyczne (Tallerman i Gibson, 2011) tej dziedziny, a prowadzone w jej zakresie debaty mają centralny charakter dla językoznawstwa jako dyscypliny (Hauser, Chomsky i Fitch, 2002; Pinker i Jackendoff, 2005). W dużej mierze stało się tak za sprawą znaczących postępów, jakie w ostatnich dekadach poczyniono w dziedzinach empirycznych – w szczególności badaniach komunikacji i zdolności poznawczych zwierząt, neuronauce, genetyce, a także modelowaniu komputerowym. Dokładniejszego przeglądu dziedziny dostarczają m.in. Fitch (2010) oraz Tallerman i Gibson (red. 2011); przeglądamy

---

<sup>1</sup> Fragmenty artykułu wykorzystano za zgodą wszystkich jego autorów oraz Redaktora *Studiów Semiotycznych*, w monografii „Ewolucja języka. W stronę hipotez gesturalnych” (Żywiczyński i Wacewicz 2015).

w języku polskim zawierają nasze wcześniejsze prace (m.in. Wacewicz, 2008, 2013, Żywicznyński i Wacewicz, 2015).

## 1.2. HIPOTEZY PIERWSZEŃSTWA GESTÓW W EWOLUCJI JĘZYKA

Według hipotez *pierwszeństwa gestów*, względnie *gesturalnych*<sup>2</sup>, język filogenetycznie wywodzi się od systemu komunikacji mającej formę gestów – przy czym przyjmujemy tu możliwie szeroką definicję gestów, a najważniejsze kryterium stanowi kanał wizualny (zob. 1.3). Są to zatem wyjaśnienia pochodzenia języka *konkurencyjne wobec hipotez głosowych*, tj. odnoszących się do wokalizacji (np. Burling, 2005; Dunbar, 1998; Mithen, 2005; MacNeilage, 2008), które nadal pozostają w przewadze. Hipotezy głosowe zwykle bazują na intuicyjnym założeniu, że rozwój komunikacji ludzkiej od form pierwotnych do języka w jego obecnym kształcie odbywał się całkowicie w modalności głosowej. Założenie to jest bardzo naturalne i przez to często, acz niesłusznie, przyjmowane *implicite* jako oczywiste. Zwłaszcza w popularnych rozważaniach na temat pochodzenia języka, prowadzonych bez dogłębnej znajomości literatury przedmiotu, możliwości inne niż rozwój języka wyłącznie w modalności głosowej często nie są w ogóle dostrzegane (np. Kenneally, 2007). Pokrewnym problemem jest brak świadomości statusu języka migowego jako pełnoprawnego języka ludzkiego (zob. 3.2.); za wymowny przykład może posłużyć tu szeroko znana typologia cech definicyjnych języka Charlesa Hocketta (np. 1960), która w pierwotnej postaci odnosi się wyłącznie do mowy i tym samym wyklucza migi jako rodzaj komunikacji językowej.

Scenariusze gesturalnego pochodzenia języka rozważane były przez wielu autorów: zarówno historycznych w kontekście filozoficzno-spekulatywnym, jak i współczesnych w kontekście naukowym. Hipotezy pierwszeństwa gestów występują w kilku odmianach różniących się ujmowaniem relacji pomiędzy komunikacją gesturalną a głosową oraz postulowanym stopniem organizacji systemu gesturalnego. Wyróżnić można:

- hipotezy języka gestów poprzedzającego mowę
- hipotezy gesturalnego protojęzyka

Do spektrum hipotez gesturalnych można zaliczyć także:

- hipotezy „gestów i mowy razem”.

Hipotezy języka gestów poprzedzającego mowę (np. Corballis, 2002; Stokoe, 2001) zakładają, że przed pojawieniem się mowy funkcjonował rozwinięty system języka gestowego. Choć autorzy ci nie wykluczają udziału wokalizacji, podkreślają, że jej rola ograniczała się do przekazywania informacji o charakterze niewerbalnym, np. emocjonalnym.

<sup>2</sup> W literaturze przedmiotu spotyka się nazwy *gestural primacy hypothesis* (GPH); *gestural hypotheses*; *gesture-first hypotheses*; „*from hand to mouth*”; „*language from gesture*”; za wariant hipotez gesturalnych uważamy także „*gesture together with speech*”.

Hipotezy gesturalnego protojęzyka przyjmują z kolei, że na etapie poprzedzającym wyłonienie się mowy komunikacja opierała się głównie na gestach funkcjonujących jako pojedyncze znaki pozbawione składni. Hipotezy gesturalnego protojęzyka występują w dwóch odmianach:

- *syntetyczny* protojęzyk gesturalny – gest odpowiada pojedynczemu słowu, reprezentując określone desygnaty (obiekty lub czynności), możliwe jest jednak budowanie z nich krótkich ciągów charakteryzujących się kompozycjonalnością treści (znaczenie całości wynika ze znaczenia składowych), lecz brakuje im organizacji składniowej czy morfologicznej (np. Hewes, 1973);
- *holistyczny* protojęzyk gesturalny – gest tworzy całą wypowiedź, reprezentując złożone myśli lub sytuacje np. „jestem głodny” (Arbib, 2005).

Hipotezy „gestów i mowy razem” (*gesture together with speech*), bazując na ścisłych związkach pomiędzy mową i gestykulacją występujących podczas ekspresji językowej, zakładają, że rozwój języka przebiegał od początku przy współdziałaniu obu modalności, głosowej i wizualnej (np. Goldin-Meadow, 2011; McNeill, 2012; Kendon, 1991). Do tego rozwiązania powracamy w sekcji 5.

Mimo dużych różnic między tymi podejściami na tym etapie traktujemy je zbiorczo; kluczowy jest tu dla nas element gesturalny, dzięki któremu wszystkie ww. podejścia lokują się w opozycji do popularnego intuicyjnego założenia o ewolucji języka wyłącznie w kanale głosowym. Nasz artykuł nie dokonuje obszernego omówienia obecnych w literaturze argumentów na rzecz hipotez gesturalnych (w tym celu zob. np. Corballis 2002; Armstrong i Wilcox, 2007; Fitch 2010), ograniczając się do krótkiego ich przeglądu w sekcjach 2.2. i 2.3. Zasadniczym celem pozostaje dla nas propozycja rozwiązania centralnego problemu większości hipotez gesturalnych, czyli problemu zmiany modalności opisanego w sekcji 3: *jeżeli język powstał jako system w przeważającej mierze gesturalny, jak wyjaśnić jego przejście do formy obecnej, czyli w przeważającej mierze głosowej?*

### 1.3. GESTY – DEFINICJA

Jednorodna i neutralna teoretycznie definicja *gestu* następuje pewnych problemów ze względu na wielość użyc tego terminu – tak potocznych, jak i specjalistycznych. W szerokim znaczeniu do zbioru gestów zaliczyć można każdy ruch ekspresyjny, tzn. wyrażający emocję lub myśl, wykonany dowolną częścią ciała (*Oxford English Dictionary*, za: Kendon 2004), włączając w to także twarz i oczy. Kendon (2004) proponuje natomiast, by mianem gestów określać tylko te działania, które cechuje wyraźnie zamierzona ekspresja, a odbiorca interpretuje je jako wolicjonalne i raczej ekspresyjne niż służące innym celom (np. praktycznym). Według najbardziej chyba restrykcyjnego

podejścia termin *gesty* zarezerwowany jest przede wszystkim dla idiosynkrycznych i spontanicznych ruchów rąk, zsynchronizowanych z mową (McNeill, 1992). Niektórzy autorzy określają tym mianem również czynności instrumentalne (zorientowane na działanie fizyczne), takie jak np. chwytywanie za pomocą rąk (zob. Fogassi i Ferrari, 2004). Co interesujące, również ruchy artykulacyjne narządów mowy mogą być rozpatrywane jako rodzaj gestów (szerzej podejście to zostanie omówione w sekcji 4.5).

Gesty nie są wyłączną domeną komunikacji międzyludzkiej. Co szczególnie istotne w kontekście ewolucji języka, posługują się nimi małpy człekokształtne (DeWaal i Pollick, 2011; Pika, Liebal, Call i Tomasello, 2005; Pollick i DeWaal, 2007; Tomasello, 2008), a także dalej spokrewnione z ludźmi małpy ogoniaste (Maestriperi, 2007; Meguerditchian, Cochet i Vauclair, 2011). Poniżej omówione zostaną dla porównania dwie różne perspektywy w definiowaniu gestów: perspektywa komunikacji międzyludzkiej i perspektywa prymatologiczna.

### 1.3.1. GESTY W KOMUNIKACJI MIĘDZYLUDZKIEJ

Gesty są integralną częścią ludzkiego systemu porozumiewania się. Ze względu na wielość form i funkcji pełnionych w procesie komunikacji (Goldin-Meadow, 2003) niełatwo poddają się klarownemu uporządkowaniu. Jedną z najbardziej kompleksowych propozycji usystematyzowania zachowań zaliczanych do gestów, jest zaproponowana przez McNeilla koncepcja kontinuum gestów (1992; 2005; 2012). Zgodnie z nią poszczególne rodzaje gestów można rozmieścić na kontinuum według następującego porządku:

gestykulacja — gesty osadzone w języku — pantomima — emblematy/deiksy — języki migowe

Porządek ten wyznaczają trzy kryteria: 1) stopień wymaganej obecności mowy maleje (od lewej do prawej strony kontinuum); 2) stopień posiadania własności językowych wzrasta; 3) stopień konwencjonalizacji znaku gestowego również wzrasta.

Termin *gestykulacja* odnosi się do ruchów rąk występujących jedynie podczas mówienia (Kendon, 2004). Jednakże mimo że ruchy te są ściśle powiązane z tokiem mowy, nie wykazują językowego usystematyzowania. Ich forma wyłania się z bieżącego sprzęgania ze znaczeniami wyrażanymi w słowach i ma charakter spontaniczny. Gestykulacja nie stanowi jednolitej kategorii i mimo że różni autorzy nieco inaczej dzielą gesty wchodzące w jej skład (zob. Ekman i Friesen, 1969; Krauss, Chen i Gottesman, 2000), sprowadzić je można do następujących rodzajów (McNeill, 1992):

– *gesty ikonizacyjne* – przejawiają wyobrazeniowe podobieństwo do semantycznej zawartości wypowiedzi, reprezentując konkretne obiekty lub działania;

- *gesty metaforyczne* – podobne do ikonicznych, z tą różnicą, że reprezentują bardziej abstrakcyjne pojęcia lub idee;
- *gesty batutowe* („uderzenia”, ang. *beats*) – forma tych gestów pozostaje jednakowa (zwykle rytmiczne ruchy ramienia lub dłoni w górę i w dół lub w przód i w tył) bez semantycznych odniesień do towarzyszącej im wypowiedzi; natomiast są one zgrane z jej rytmem;
- *gesty deiktyczne* (deiksy, wskazujące) – ich podstawową funkcją jest wskazywanie na obiekty lub czynności, także te nieobecne fizycznie, w kontekście zachodzącej właśnie komunikacji. Ze względu na swoją strukturę i funkcję mogą tworzyć oddzielną kategorię, jednakże wraz z emblematami mogą być absorbowane przez gestyfikację.

*Gesty osadzone w języku* (*language-slotted gestures*) wykazują podobieństwo do gestyfikacji, jednak wyróżnia je syntagmatyczny związek, w jaki wchodzi z słowami. Za przykład niech posłuży następujące zdanie: „pogoda była dobra, lecz jedzenie [gest machnięcia ręką]”. Gest ten kończy zdanie, wypełniając gramatyczną lukę, jaka powstała w wyniku pominięcia predykatu. Odrębny charakter ma *pantomima*, która nie wymaga obecności mowy. Obiekty i zdarzenia reprezentowane są tutaj za pomocą gestów składanych w dłuższe sekwencje. Przykładem pantomimy może być następujący ciąg gestów: przywołanie kogoś ruchem ręki, przyłożenie palca do ust oznaczające „zachowaj ciszę” i następnie wskazanie palcem kierunku, w którym zachodzi wydarzenie mające być przedmiotem wspólnej uwagi. Mimo że występuje tu pewna sekwencyjna organizacja znaków, nie wykazuje ona właściwości charakteryzujących składnię. Natomiast główną funkcją *emblematów* jest zastępowanie pojedynczych słów. Każda kultura czy społeczność ma w swoim repertuarze komunikacyjnym określony zestaw emblematów, które cechuje arbitralność, normatywność – czyli precyzyjne wymogi dotyczące sposobu wykonania (analogicznie do reguł fonologicznych obowiązujących w danym języku), intencjonalność<sup>3</sup> i kulturowa transmisja. Przykładem emblematu funkcjonującego w naszym kręgu kulturowym jest gest „OK” wykonywany poprzez podniesienie przedramienia i zetknięcie czubków palca wskazującego i kciuka. Emblematy posiadają wyraźne cechy znaku językowego, jednak funkcjonują one zwykle jako pojedyncze sygnały przekazujące krótką wiadomość. *Migi* z języków migowych zamykają kontinuum gestów. Zastępując mowę, tworzą wobec niej równoważny sposób wymiany informacji językowej. Posiadają one wszystkie cechy systemu językowego, a najważniejsza różnica w porównaniu z mową polega na operowaniu inną modalnością – głosową w przypadku mowy i wzrokową w przypadku języka migowego (zob.

<sup>3</sup> Tu i dalej w tekście używamy terminu *intencjonalny* w sensie psychologicznym, tj. ‘celowy’, ‘wynikający z intencji podmiotu’ – nie zaś w sensie filozoficznym tego terminu, tj. ‘mający treść intencjonalną’, ‘będący o czymś’.

sekcja 3.2.). Jeszcze raz należy z całą mocą podkreślić, że wbrew potocznym intuicjom miגי stanowią system całkowicie różny od gestykulacji, nie tylko pod względem ich własności formalnych i komunikacyjnych, lecz także implementacji mózgowej.

Przedstawiona wyżej charakterystyka gestów osadzona jest w kontekście komunikacji międzyludzkiej. Odzwierciedla więc specyficznie ludzkie zdolności poznawcze, a także warunki ekologiczne, strukturę społeczną oraz obecność systemu językowego realizowanego głównie za pomocą mowy. Definiowanie gestów występujących w innym kontekście, np. komunikacji małp, wymaga uwzględnienia odmiennych warunków. Niżej przedstawiamy więc wybrane definicje i typologie gestów sformułowane na gruncie badań prymatologicznych.

### 1.3.2. GESTY W KOMUNIKACJI INNYCH NACZELNYCH

Mimo że naczelne posługują się wieloma rodzajami komunikacji, to jednak szczególne zainteresowanie prymatologów przyciągnęły wykorzystywane przez nie gesty. Od pozostałych środków porozumiewania odróżnia je (de Waal i Pollick, 2011; Pollick i de Waal, 2007; Tomasello, 2008):

- nabywanie na drodze indywidualnego uczenia się,
- intencjonalne użycie i plastyczność,
- względna niezależność od procesów emocjonalnych,
- ukierunkowanie na określonego odbiorcę.

Dla porównania, wokalizacje małp są sztywnymi i działającymi nioselektywnie zachowaniami emocjonalnymi w dużej mierze uwarunkowanymi genetycznie. Komunikacja gestowa oparta na zrytualizowanych zachowaniach, występuje głównie u małp człekokształtnych, co sugeruje filogenetycznie późne wyłonienie się tej formy komunikacji (de Waal i Pollick, 2011).

Badania porównawcze nad komunikacją gestową małp człekokształtnych (bonobo, *Pan paniscus*; szympanсів *Pan troglodytes*; goryli, *Gorilla gorilla*; orangutanów, *Pongo pygmaeus*) ujawniły, że stosowane przez nie gesty przejawiają pewne podobieństwa (mimo istotnych różnic) do gestów obserwowanych u niemowląt i małych dzieci, dopiero zaczynających posługiwać się mową (Pika i Liebal, 2006). Podobieństwa dotyczą zróżnicowania repertuaru gestów oraz ich intencjonalnego użycia w obu porównywanych grupach. Różnice natomiast związane są z naturą gestów i sposobem ich wykorzystywania: u małp większość gestów ma charakter diadyczny – nadawca, kierując na siebie uwagę odbiorcy, wyraża prośbę lub życzenie (imperatywne użycie gestu); dzieci natomiast potrafią używać gestów triadycznie – kierują uwagę odbiorcy na zewnętrzne zdarzenie lub obiekt celem dzielenia uwagi lub skomentowania przedmiotu wspólnej uwagi (deklaratywne użycie gestu). Zaobserwowane różnice wynikają m.in. z odmienności podstaw ko-

gnitywnych („infrastruktury poznawczo-społecznej”, zob. Tomasello, 2008), jakie posiadają ludzie i małpy człekosształtne. Dostarczają ponadto cennych wskazówek dotyczących gesturalnych źródeł języka, a także samej natury komunikacji międzyludzkiej.

Prymatolodzy studiujący zachowania komunikacyjne małp wypracowali kilka definicji gestów. DeWaal i Pollick (2011 oraz Pollick i deWaal, 2007) proponują by zastosowanie terminu *gesty* zawęzić do ruchów manualnych, ponieważ wykazują one cechy wyraźnie odrębne w stosunku do pozostałych rodzajów ruchów komunikacyjnych (zob. charakterystykę gestów opisaną na początku niniejszej sekcji). Tomasello (2008) charakteryzuje gesty szerzej, tj. jako intencjonalne zachowania komunikacyjne działające w modalności wizualnej, obejmujące głównie postawę, ekspresję twarzy i gesty manualne – wykazujące plastyczność oraz nabyte na drodze rytualizacji ontogenetycznej. Nieco inną definicję proponuje Pika (2008). Definiuje ona gesty jako ekspresyjne ruchy kończyn, głowy i postawy ciała:

- skierowane do odbiorcy,
- pozbawione skutków mechanicznych – jak np. popchnięcie kogoś – co odróżnia gesty od ruchów instrumentalnych, których celem nie jest komunikacja<sup>4</sup>,
- prowokujące zamierzoną odpowiedź,
- wykonywane intencjonalnie (celowo).

*Intencjonalność* stanowi w przedstawionych definicjach kryterium kluczowe. W zamierzeniu pozwala ona różnicować zachowania gestowe od innych zachowań komunikacyjnych o wrodzonym, a więc nieintencjonalnym, charakterze, które w różnej postaci spotyka się w systemach komunikacyjnych zwierząt (Tinbergen, 1951). O intencjonalnym użyciu gestu decyduje:

- względna niezależność od kontekstu – ten sam gest używany jest do osiągnięcia różnych celów komunikacyjnych, jak też różne cele komunikacyjne osiągnięte są za pomocą tego samego gestu (de Waal i Pollick, 2011; Tomasello, 2008),
- sprawdzanie obecności i stanu uwagi odbiorcy komunikatu (*audience-checking*),
- oczekiwanie na jego reakcję (*response-waiting*) oraz
- powtarzanie gestu przy braku spodziewanej reakcji (*persistence*).

Ponieważ aspekt modalności zmysłowej jest kluczowy w kontekście całości artykułu, wspomnieć należy podział gestów dokonany właśnie pod tym względem, zaproponowany przez Simone Pikę (2008):

---

<sup>4</sup> Powstaje tu pewna trudność, ponieważ kryterium to komplikuje rozpoznanie ważnej klasy gestów – *gestów dotykowych* – które wymagają pewnego oddziaływania mechanicznego na odbiorcę.

- *gesty dźwiękowe* – ich wykonywaniu towarzyszy wydawanie dźwięku, np. klaskanie lub charakterystyczne uderzanie w klatkę piersiową u goryli;
- *gesty dotykowe* – angażują kontakt dotykowy z odbiorcą, np. *directed scratch* u szympansov, uważany za pierwszy stwierdzony u dziko żyjących małp gest referencyjny, tj. mający odniesienie przedmiotowe (zob. Pika i Mitani, 2009);
- *gesty wizualne* – angażują wyłącznie modalność wzrokową.

W naszej pracy przyjmujemy szeroką definicję gestów. Wychodzimy od intuicyjnego rozumienia gestów jako intencjonalnych, komunikacyjnych a więc nieinstrumentalnych ruchów rąk, ale rozszerzamy je o większość sygnałów cielesnych działających w modalności wizualnej, czyli odbieranych wzrokowo.

*Centralnymi, prototypowymi przykładami gestów są zatem ruchy ramion, dłoni i palców wykonywane z zamiarem przekazania określonego komunikatu. Przykłady bardziej peryferyjne obejmują inne sygnały wizualne, takie jak elementy proksemiczne, postawę i orientację ciała, świadome ekspresje twarzy, a nawet kierunek spoglądania.* Ruchy artykulacyjne związane z mową, chociaż mogą tworzyć continuum z ruchami wyszczególnionymi wyżej, zwłaszcza mimiką, są włączane do zbioru gestów tylko o tyle, o ile mogą być postrzegane wzrokowo.

## 2. ARGUMENTY NA RZECZ HIPOTEZ PIERWSZEŃSTWA GESTÓW

Pierwsze spekulacje o roli gestów w ewolucji języka pojawiły się na długo przed tym, zanim problem ten stał się przedmiotem badań naukowych. Dopiero w latach siedemdziesiątych ubiegłego wieku hipotezę pierwszeństwa gestów oparto na szerszych podstawach empirycznych. Obecnie jej różne formy składają się na jedno z najbardziej wpływowych stanowisk w dyskusji nad filogenezą języka i nadal są twórczo rozwijane. W niniejszej sekcji zarysowujemy historię hipotez gesturalnych oraz najważniejsze dowody, które je wspierają.

### 2.1. GESTY I POCHODZENIE JĘZYKA – KRÓTKI RYS HISTORYCZNY

Przez wiele stuleci – od starożytności aż do XVIII wieku – dominował pogląd, że gesty są naturalną formą porozumiewania się ludzi, tworzącą autonomiczny i uniwersalny język. Propagatorami tej idei byli: retor rzymski Kwintylian, a później Giovanni Bonifacio (1547–1645) i John Bulwer (1606–1656). Również twórcy i nauczyciele pierwszych systemów języka migowego m.in. Charles-Michel de l'Épée (1712–1789) i Roch-Ambroise Cucurron Sicard (1742–1822) wierzyli, że komunikacja gestowa stanowi doskonałą podstawę do stworzenia uniwersalnego języka, niezależnego od mowy. Gesty



używane przez głuchoniemych zajmowały także Diderota (1713–1784), który uznawał je za wolne od zniekształcających wpływów konwencji i tradycji, co umożliwia bezpośredni wgląd w procesy myślowe.

Gesty pojawiały się także w historycznych spekulacjach dotyczących powstania języka. Condillac (1715–1780) zakładał, że zanim ludzie nauczyli posługiwać się mową, komunikowali się między sobą za pomocą gestów i ruchów ciała. Pierwszeństwo gestów postulował także Giambattista Vico (1688–1744), widząc w nich najdoskonalszy sposób przedstawiania wyobrażeń wzrokowych. Edward Tylor (1832–1917) podkreślał wielość form wyrażania myśli, włączając w to języki gestowe, pismo obrazkowe i słowo pisane, uznając jednocześnie, że w zrozumieniu początków języka pomocne mogą być zwłaszcza studia nad gestami i pismem obrazkowym. Wilhelm Wundt (1832–1920) początków języka upatrywał w podstawowych ruchach ekspresyjnych, zorganizowanych w odrębnych dla poszczególnych doświadczeń emocjonalnych wzorcach. W pierwszej połowie XX wieku zainteresowanie podobnymi problemami osłabło, jednak na początku lat 70. nastąpiła zmiana wyznaczająca początek współczesnej debaty na temat roli gestów w ewolucji języka.

## 2.2. ARGUMENTY HEWESA I WZNOWIENIE ZAINTERESOWANIA ROLĄ GESTÓW W EWOLUCJI JĘZYKA

W nowoczesnej postaci hipoteza pierwszeństwa gestów została sformułowana przez amerykańskiego antropologa Gordona W. Hewesa. W 1973 roku opublikował on artykuł *Primate Communication and the Gestural Origin of Language* (*Komunikacja naczelnych i gesturalne źródła języka*), w którym syntetyzując dane pochodzące z różnych dziedzin nauki, przedstawił zbiór argumentów wspierających gesturalny scenariusz pochodzenia języka. Niektóre argumenty przedstawione przez Hewesa nadal zachowują ważność i są przywoływane w toczącej się obecnie debacie dotyczącej ewolucji języka (zob. np. Corballis, 2002; Tomasello, 2008), inne natomiast zostały zgodnie z wynikami nowszych badań albo zaktualizowane, albo odrzucone. W niniejszej sekcji ograniczymy się do przedstawienia najważniejszych argumentów zaproponowanych przez Hewesa.

Głównym argumentem Hewesa były względne *sukcesy w uczeniu małp języka migowego* (np. Gardner i Gardner, 1969), co stanowiło wyraźny kontrast w porównaniu z całkowicie nieudanymi próbami nauczania ich języka mówionego. Można na tej podstawie sądzić, że wczesne człowiekowate, których zdolności poznawcze prawdopodobnie nie odbiegały zasadniczo od tych, które posiadają współcześnie żyjące małpy człekokształtne, także mogły być

zdolne do wytworzenia protojęzyka<sup>5</sup> gestowego. Chociaż wczesne człowiekowie, podobnie jak inne naczelne, posługiwały się komunikacją głosową, nie mogła być ona według Hewesa punktem wyjścia dla rozwoju komunikacji typu językowego, tj. arbitralno-konwencjonalnej, kompozycjonalnej, propozycjonalnej, itd. Główną przeszkodą ku temu był brak wystarczającej kontroli wolicjonalnej nad głosem. Opierając się na dostępnych wówczas badaniach, Hewes zauważał, że reakcje wokalne u małp są wyzwalane przez określone bodźce emocjonalne, z którymi są sztywno powiązane, a ponadto wokalizacje nie są adresowane do określonego odbiorcy, działają zatem nieselektywnie, a nawet pod nieobecność jakiegokolwiek odbiorcy. Cechom tym przeciwstawił komunikacją gestową małp, którą charakteryzuje wolicjonalność i oparcie na wyższych procesach kognitywnych, co umożliwiła większą plastyczność i otwartość na innowacje (w tym miejscu konieczna jest jednak uwaga, że wyniki obecnie dostępnych badań – choć nie kwestionują wspomnianej różnicy jakościowej – wskazują na znacznie bardziej złożoną niż wcześniej sądzono naturę komunikacji głosowej naczelnych<sup>6</sup>).

Hewes powołał się także na wyniki analiz wskazujące, że *ludzki przewód głosowy jest przystosowaniem stosunkowo późnym*, obecnym jedynie u *Homo sapiens* (zob. jednak sekcja 3.1.). Uważał przy tym, że gatunki człowiekowatych żyjące przed *Homo sapiens* wykazywały umiejętności, które według niego wymagały istnienia choćby rudymen tarnej formy języka: posługiwanie się ogniem, wytwarzanie narzędzi, a zwłaszcza organizowanie grupowych polowań na duże zwierzęta. Zakładał, że do realizacji tych zadań oraz podtrzymywania ich jako tradycji kulturowych protojęzyk gestowy mógł być w pełni wystarczający.

Hewes (1973, 1981, 1996) poruszył także kilka innych istotnych zagadnień, które podejmowane są we współczesnej debacie nad rolą gestów w ewolucji języka. Podkreślał możliwą rolę gestu wskazywania na początkowych etapach kształtowania się języka, zwrócił uwagę na zjawisko imitacji zachowań i czynności manualnych jako potencjalnej drogi ustanawiania znaków językowych, wskazywał też na znaczenie lateralizacji i kontroli ruchowej w kształtowaniu się języka gestowego. Ciekawym argumentem było spostrzeżenie o depigmentacji wnętrza dłoni u ludzi o ciemnym kolorze skóry –

<sup>5</sup> Hewes jako pierwszy użył terminu *protojęzyk*, odnosząc go do stadium przejściowego pomiędzy niejęzykową komunikacją małp, a komunikacją z wykorzystaniem rozwiniętego języka. Jednak obecne rozumienie tego terminu, który w uproszczeniu oznacza język pozbawiony 'gramatyki' – w sensie reguł morfologicznych czy składniowych – rozpropagował Derek Bickerton (1990).

<sup>6</sup> Cechują ją m.in. funkcjonalne rozumienie odniesienie (*functional reference*), efekty widowni (*audience effects*), produktywność (jednakże bez kompozycjonalności), możliwość celowej manipulacji (*tactical deception*) – przegląd nowych danych np. w Slocombe (2011).

Hewes (1996) sugerował, że cecha ta może służyć lepszej widoczności dłoni, a więc mogła wyłonić się jako przystosowanie do komunikacji gesturalnej.

Na koniec wypada zaznaczyć, że Hewes zdawał sobie sprawę z problemu 'zmiany modalności' i proponował własne rozwiązanie, oparte na ideach 'gestów ust' (*mouth gestures*) i symbolizmu dźwiękowego jako mechanizmów ewolucyjnego przejścia od gestu do mowy<sup>7</sup> (zob. także sekcja 4.5.1.).

### 2.3. WSPÓŁCZESNE HIPOTEZY GESTOWE

Obecnie hipotezy gestowe są wciąż rozwijane i zajmują ważne miejsce w toczącej się debacie nad filogenezą języka. W niniejszej sekcji przedstawimy kilka bardziej współczesnych – licząc od momentu ukazania się publikacji Hewesa – przesłanek przemawiających na korzyść hipotez gesturalnych. Eksponują one między innymi ikoniczny potencjał gestów, jako najbardziej intuicyjny sposób ekspresji językowej, nowe ustalenia odnośnie funkcji mózgu oraz charakterystyczną dla ludzi zdolność do naśladownictwa, czy szerzej: *mimezy*.

#### 2.3.1. IKONICZNOŚĆ GESTÓW

William Stokoe, jeden z pionierów nowoczesnych badań nad językami migowymi, wykazał równoprawność języków migowych względem mowy oraz zaproponował rozwiązania dla kilku kluczowych problemów ewolucji języka (Stokoe, 1960). Pierwszy z nich dotyczy przyporządkowania znaczeń do arbitralnych dźwięków mowy, który można rozwiązać, odwołując się do gestowego stadium w ewolucji języka – ikoniczność gestu, a więc wizualne podobieństwo do tego, co oznacza, tworzy pomost pomiędzy dźwiękiem a jego odniesieniem. Drugi problem związany jest z pochodzeniem gramatyki i tutaj także zdaniem Stokoe (1991; potem także Armstrong, Stokoe i Wilcox, 1995) odnieść się można do ikoniczności gestu, który z jednej strony może być odczytywany jako pojedynczy znak, z drugiej zaś może reprezentować np. oprócz samej czynności także podmiot, który ją wykonuje oraz przedmiot, na który ukierunkowane jest działanie. W ten sposób pojedynczy gest może reprezentować działanie jako złożoną całość: ręka funkcjonuje jako prototypowy rzeczownik, a jej działanie to prototypowy czasownik, co razem składa się na prototypowe zdanie (Armstrong i Wilcox, 2007). Przestrzenna natura gestu pozwala także na intuicyjnie zrozumiałą wizualizację ról semantycznych (dłoń uderza w dłoń lub kreśli ścieżkę) oraz relacji przestrzennych i czasowych. Podsumowując, odwołanie się do stadium gestowe-

<sup>7</sup> „A plausible theory of the primacy of gestural language over speech must, as has already been noted, account for its general replacement by spoken language” (Hewes 1996: 587).

go w ewolucji języka ułatwia wytłumaczenie zarówno pochodzenia arbitralności mowy, jak i źródła gramatyki.

Oprócz argumentów teoretycznych wspierających taką linię rozumowania zaczynają się pojawiać także interesujące dane empiryczne, w szczególności z badań prowadzonych przez grupę Susan Goldin-Meadow. W jednym z eksperymentów (Goldin-Meadow i in. 2008) badani opisywali proste czynności za pomocą mowy, gestów oraz przezroczy. Opisy słowne formowane były zgodnie z szykiem zdania właściwym językowi ojczystemu badanych, jednak już opisy przy użyciu gestów lub przezroczy, niezależnie od języka badanych, cechowała stabilna kolejność „działający (aktor) – obiekt (patiens) – działanie (akt)”, odpowiadająca szykowi SOV; prowadzi to do wniosku, że komunikacja gesturalna ujawnia naturalny, niejako ‘pierwotny’, format mentalnego reprezentowania struktury wydarzeń. W innym badaniu (Fay, Arbib i Garrod, 2013) uczestnicy za pomocą gestów lub niewerbalnych wokalizacji komunikowali kolejne ‘znaczenia’ z zamkniętego repertuaru znaczeń – np. emocji, przedmiotów lub działań. Komunikacja gestowa okazała się zarówno skuteczniejsza, jak i wydajniejsza czasowo od głosowej. Autorzy badania uznali to za świadectwo, że modalność wizualna, z natury swej lepiej nadająca się do wyrażania znaczeń motywowanych, mogła być bardziej użyteczna w początkowych stadiach rozwoju języka.

Należy jednak zaznaczyć, że status ikoniczności jako pomocy poznawczej ułatwiającej przejście do komunikacji opartej na symbolach jest daleko nieoczywisty. Na przykład w ontogenezie reprezentacje ikoniczne nie wydają się uprzywilejowane względem arbitralnych, tj. dzieci nie nabywają ich szybciej czy łatwiej (zob. np. Tomasello, 2008: 147). Także pod względem realizacji mózgowej systemu przetwarzania znaków ikonicznych oraz systemu przetwarzające symbole wykazują względem siebie daleko idącą rozłączność (Niederhut, 2012).

### 2.3.2. RĘCZNOŚĆ I LATERALIZACJA

Przez pewien okres wskazywano na ręczność i lateralizację jako cechy przemawiające za gesturalnym scenariuszem pochodzenia języka. U zdecydowanej większości ludzi lewa półkula mózgu odpowiada zarówno za większość zadań związanych z przetwarzaniem języka (to w niej lokują się tak zwane ‘ośrodki mowy’), jak i kontrolę motoryczną dominującej ręki (praworęczność cechuje około 90% globalnej populacji ludzi). Źródła tej korelacji są nie do końca wyjaśnione, jednak dostępne dane dobitnie wskazują na jej systematyczny charakter – np. częstość udziału lewej półkuli w przetwarzaniu języka jest wprost proporcjonalna do stopnia preferencji względem prawej ręki (Knecht i in., 2000). Sformułowano kilka hipotez odnoszących się do lateralizacji i języka. Hewes (1973) zakładał, że lateralizacja dla precyzyjnej kontroli ruchu, a potem gestykulacji, pojawiła się przed wyłonieniem się

mowy. Według Corballisa (2003) reprezentacje ruchów komunikacyjnych rąk były stopniowo absorbowane przez wokalizację, która już przedtem była reprezentowana w półkuli lewej, jak to ma miejsce u wielu innych gatunków porozumiewających się głosowo. Natomiast zgodnie z hipotezą rzutową sformułowaną przez Calvina (1982, 1983; zob. też Calvin i Bickerton, 2000) jednym z pierwszych przejawów lateralizacji funkcji mózgu i ręczności była czynność rzucania kamieniami w celu upolowania drobnej zwierzyny. Celne rzucanie wymaga obliczenia trajektorii lotu pocisku i sporządzenia odpowiadającego jej złożonego planu motorycznego, precyzyjnie koordynującego czasowo ruchy wielu części ciała, od palców i nadgarstka po tułów. Według Calvina rzucanie jedną ręką doprowadziło do wyselekcjonowania w mózgu dedykowanego obwodu neuronalnego pozwalającego sprostać tym wymaganiom obliczeniowym; obwód ten następnie objął swym zasięgiem kontrolę ruchową potrzebną do wytwarzania narzędzi, gestykulacji i ostatecznie artykułowanej mowy oraz składni.

Obecnie znaczenie ręczności i lateralizacji funkcji mózgu jako dowodów przemawiających za hipotezami gesturalnymi maleje. Przede wszystkim najnowsze badania podważają dawny pogląd, że ręczność na poziomie populacyjnym jest cechą wyłącznie ludzką. Choć dane nadal są niejednoznaczne, wydaje się, że wyróżnikiem naszego gatunku jest nie tyle sama obecność tego zjawiska, ile jego skala i systematyczność (np. Cashmore, Uomini i Chapelain, 2008). Jeżeli chodzi o małpy człekokształtne, część badań (np. Harrison, 2008) nie stwierdza jakiegokolwiek preferencji na poziomie populacji, jednakże inne (np. Hopkins, 2006) odnotowują wyraźną jej obecność, najczęściej praworęczności, przynajmniej w przypadku niektórych gatunków i niektórych czynności. Podobne dane istnieją dla małp nieczłekokształtnych, na przykład pawiany cechuje preferencja do wykonywania gestów komunikacyjnych, ale nie instrumentalnych, za pomocą prawej ręki (zob. Meguerditchian, Cochet i Vauclair, 2011). Ponadto lewa półkula kontroluje komunikację głosową u wielu gatunków zwierząt, w tym daleko spokrewnionych z człowiekiem, np. żab i ptaków (zob. Corballis, 2003). Obserwacje te podważają przyjmowane wcześniej założenie, że cecha ta występuje jedynie u człowieka i jest efektem działania presji selekcyjnych, takich jak np. potrzeba precyzyjnej kontroli ruchowej niezbędnej do posługiwania się narzędziami. W badaniach nad neuroobrazowaniem mózgu wykazano także, że pojęcie „lewej półkuli językowej” jest daleko idącym uproszczeniem, a funkcje związane z przetwarzaniem języka angażują wiele obszarów rozsianych po całym mózgu (np. Deacon, 1997; Lieberman, 2003).

### 2.3.3. OBSZAR BROKI I NEURONY LUSTRZANE

Badania mózgu podsunęły także kilka innych interesujących tropów prowadzących w stronę hipotez gesturalnych: dotyczą one przede wszystkim re-

wizji poglądu na funkcję obszaru Broki i odkrycia neuronów lustrzanych. Definicja funkcji obszaru Broki wiązane pierwotnie z produkcją mowy, ulega w świetle nowych badań znaczącej zmianie (Fadiga, Craighero i D'Ausilio, 2009). Wykazano, że obszar ten jest zaangażowany również w rozumienie języka, wykonywanie i obserwację czynności manualnych, wykonywanie i słuchanie muzyki oraz reprezentację abstrakcyjnych struktur hierarchicznych. Na tej podstawie niektórzy autorzy (Fadiga, Craighero i D'Ausilio, 2009) wyciągają wniosek, że funkcja obszaru Broki polega na detekcji i reprezentowaniu złożonych hierarchicznie/syntaktycznie zależności, bez względu na modalność i użycie (produkcja lub percepcja). Zakłada się, że ewolucyjnym prekursorem tych zdolności jest aktywność motoryczna związana z wykonywaniem czynności, a dzięki funkcji neuronów lustrzanych – także ich rozumieniem. Dane te wskazują na równoważność modalności wizualnej i głosowej jako potencjalnych sposobów transmisji informacji językowej, jednocześnie przyznając w tym względzie filogenetyczne pierwszeństwo systemowi motorycznemu.

Ciekawą i wciąż rozwijaną koncepcję ewolucji języka zaproponowali Rizzolatti i Arbib (1998). Jej podstawą są *neurony lustrzane*, czyli grupa komórek nerwowych aktywnych zarówno podczas np. wykonywania czynności takiej jak chwytanie, jak i podczas obserwacji tej samej czynności wykonywanej przez kogoś innego. Neurony lustrzane pierwotnie odkryto w korze przedruchowej makaków (pole F5), którego położenie odpowiada mniej więcej obszarowi Broki w mózgu ludzkim (Rizzolatti i in. 1996). Późniejsze badania potwierdziły istnienie systemu lustrzanego także u człowieka (Iacoboni i in., 1999). Hipoteza neuronów lustrzanych identyfikuje neuronalny mechanizm odpowiadający za równoważność i odwracalność ról pomiędzy nadawaniem a odbiorem komunikatu (*parity requirement*), jako wstępny i zarazem niezbędny warunek komunikacji. Małpa, widząc, jak inna małpa (lub człowiek) sięga po orzeszek, rozumie sens tej czynności dzięki kopii reprezentacji motorycznej tego działania we własnym mózgu. W ten sposób pomiędzy „nadawcą” (wykonującym czynność) a „odbiorcą” (obserwującym tę czynność) wytwarza się mentalny związek polegający na jednoczesnym reprezentowaniu tej samej informacji.

Arbib (2002, 2005, 2012) z czasem zmodyfikował pierwotną koncepcję, podkreślając, obok funkcji neuronów lustrzanych, rolę imitacji oraz wolicjonalnego systemu kontroli na ruchami komunikacyjnymi. Na nowo opracowany został również stadialny model filogenezy języka: trzy pierwsze stadia obejmują wszystkie naczelne do momentu pojawienia się wspólnego przodka szympansa i człowieka, kolejne zaś dotyczą zmian zachodzących po rozdzieleniu się linii obu gatunków. Cały model przedstawia się następująco (Arbib, 2005):

S1: chwytanie

S2: system lustrzany dla chwytania

S3: pojawienie się zdolności do prostej imitacji (tylko u szympanśów)

S4: pojawienie się zdolności do złożonej imitacji (po rozdzieleniu się linii człowieka i szympanśów)

S5: proto-znak (kluczowa zmiana prowadząca do powstania otwartego repertuaru znaków)

S6: proto-mowa (kluczowa zmiana prowadząca do rozszerzenia kontroli korowej nad głosem)

S7: rozwinięty język

Chociaż podstawę modelu wciąż stanowią czynności manualne i system lustrzany, jednak kluczowe znaczenie odgrywa w nim obecna jedynie u człowiekowatych zdolność do pantomimicznej imitacji, złożonej z doraźnych proto-znaków reprezentujących całe sytuacje i działania (holistyczny protojęzyk). Dalej, dzięki rozszerzeniu kontroli ruchowej na język i krtań, stopniowo inkorporowana byłaby modalność głosowa i porozumiewanie się oparte na użyciu symboli arbitralnych.

#### 2.3.4. MIMEZA I PANTOMIMA

Do imitacji nawiązuje także wpływowa linia argumentacji wywodząca się z mimetycznej koncepcji ewolucji języka, rozwijanej początkowo przez Merlina Donalda (1991, 2001). Według Donalda, za fundamentalną różnicę między zdolnościami poznawczymi małp a ludzi odpowiada *mimeza* (*mimesis*), czyli zdolność do reprezentacji mentalnych, które są uświadomione, samodzielnie inicjowane i intencjonalne, ale nie mają charakteru językowego. Zdolność mimezy pozwala na zapamiętanie schematów ruchowych – takich jak podskok, rzucenie przedmiotu, czy krok w tańcu – jako reprezentacji *oderwanych od konkretnego podmiotu oraz samego wykonania tych czynności*. Takie reprezentacje nie muszą aktywować się wyłącznie za zasadzie bezwarunkowej reakcji na bodziec – podmiot może je w dowolnym momencie świadomie, wolicjonalnie przywoływać z pamięci (*autocueing*). Dzięki temu może w każdej chwili wykonać daną czynność, np. krok taneczny, czy to podczas faktycznego tańca, czy to tylko dla treningu tanecznego lub zabawy. Taką sekwencję ruchów można również przywołać jedynie do pamięci roboczej, czyli wyobrazić sobie „off-line”, aby ją dokładniej zaplanować, czy udoskonalić. Podmiot zdolny do reprezentacji mimetycznych może wyodrębnić sekwencje ruchowe (np. rzucanie kamieniem) zaobserwowane u innych, a następnie je zapamiętać, wyobrazić sobie, czy samodzielnie naśladować. W ten sposób mimeza leży u podstaw imitacji. Reprezentowana mimetycznie czynność – np. naśladowanie rzutu kamieniem – jest rozpoznawana jako ta sama przez osobę wykonującą i obserwującą, a więc można jej użyć do komunikacji. Warto zauważyć, że takie znaki mimetyczne nie są arbitralne, konwencjonalne ani kompozycjonalne, a ich natura jest w fundamentalny sposób

cielesna; mimo że mimeza jest zdolnością multimodalną, domena wzrokowa odgrywa tu największą rolę.

Jordan Zlatev (np. 2008)<sup>8</sup> przedstawia najbardziej zaawansowaną kontynuację koncepcji Donalda. Postuluje on hierarchię zdolności mimetycznych odpowiadającą także stadiom ich filogenetycznego rozwoju:

- *proto-mimeza* – opierająca się na odwzorowaniu spostrzeżeń zewnętrznych (eksterocepcja) na ruchy własnego ciała (propriocepcja), obecna u małp nieczłękokształtnych a ontogenetycznie u noworodków, manifestująca się w zachowaniach takich jak kontakt wzrokowy czy prosta koordynacja zachowań;
- *mimeza diadyczna* – opierająca się na wolicjonalnej reprezentacji, obecna w podstawowej formie u małp człękokształtnych, leży u podstaw np. imitacji, czy wyobraźni, w tym wyobrażania sobie przyszłości;
- *mimeza triadyczna* – opierająca się na intencji komunikacyjnej, właściwa ludziom i tylko do pewnego stopnia ukulturowionym małpom (*encultured apes*), objawiająca się np. celowo wykonywanymi gestami ikonicznymi lub wskazującymi w sposób deklaracyjny;
- *post-mimeza I stopnia* – opierająca się na normatywności i konwencji, obecna u ludzi (i w minimalnym stopniu małp uczonych komunikacji symbolicznej), znajdująca wyraz w komunikacji za pomocą symboli;
- *post-mimeza II stopnia* – opierająca się na systematycznym i kompozycyjnym użyciu symboli zarówno w komunikacji jak i wewnętrznie w procesach myślowych, obecna tylko u ludzi, leży u podstaw języka oraz np. zdolności do rozumienia fałszywych przekonań.

Podobne poglądy przedstawia Daniel Hutto (2008) – pierwotną formą komunikacji było według niego mimetyczne odgrywanie wydarzeń nieposegmentowane na poszczególne jednostki znaczące, lecz mające postać całościowej, cielesnej narracji. Jak przypuszcza Hutto, regularne odgrywanie (*re-enactment*) mogło pełnić istotną funkcję społeczną, tworząc podstawę do formowania się zwyczajów, zacieśniając więzy i stopniowo zastępując iskanie. Michael Tomasello (2008) uznaje pantomimę i wskazywanie za naturalne i pierwsze specyficznie ludzkie formy komunikacji, stanowiące fazę przejściową pomiędzy komunikacją małp a konwencjonalnym językiem. Ich podstawę (podobnie jak wykształconych później języków konwencjonalnych) stanowią w większości właściwe jedynie człowiekowi formy poznania społecznego i motywacji: dzielenie intencji oparte na rekursywnym *mindreading* oraz kooperacyjne motywy porozumiewania się.

---

<sup>8</sup> Mimetyczne stanowisko Zlateva (np. 2008, 2013) jest jednak bliższe teoriom multimodalnym (zob. sekcja 5), niż *stricte* gesturalnym.



## 2.3.5. INNE ARGUMENTY

Wacewicz i Żywicznyński (2008) zwracają uwagę na większą *sekretność* (tajność) komunikacji w kanale wzrokowym, który dzięki temu lepiej nadaje się na przekazanie wiadomości tylko wybranym adresatom, a nie wszystkim będącym w pobliżu. Z logiki ewolucyjnej wynika, że mogło to mieć zasadnicze znaczenie na wczesnych etapach rozwoju komunikacji, kiedy była ona systemem ewolucyjnie niestabilnym – nie miała jeszcze właściwej ludziom informatywnej natury, lecz naturę manipulatywną, typową dla zwierząt. Pewnym potwierdzeniem tej koncepcji są najnowsze badania polowe szympanów, które zwiększają udział gestów w komunikacji w sytuacjach wymagających tajności (Hobaiter i Byrne, 2012).

Alternatywne podejście do roli gestów w ewolucji języka proponuje David McNeill (2012). Jego zdaniem teorie zakładające pierwszeństwo gestów nie są w stanie w zadowalający sposób wyjaśnić głębokiej i wielopoziomowej integracji mowy i gestów, jaką obserwujemy obecnie podczas ekspresji językowej. Jednocześnie koncepcje te umniejszają rolę gestów w ucieleśnianiu języka, zakładając ich marginalizację przez mowę. Wychodząc z założenia, że gesty są nieodłącznym komponentem mowy, nie zaś tylko jej uzupełnieniem, McNeill uważa, że oba rodzaje semiozy wywodzą się z umysłowych jednostek wyobrazeniowo-językowych, które określa mianem *Growth Points* (GPs). Tworzą one niepodzielne pakiety dające początek ideom wyrażanym jednocześnie w gestach i mowie. Ta sama idea jest zatem w naturalny sposób wyrażana w tym samym czasie za pomocą kodu wyobrazeniowego i językowego. Według McNeilla nabycie tej właśnie zdolności jest krytycznym momentem w poznawczym rozwoju człowieka, prowadzącym do pojawienia się języka. Decydująca jest tu zatem zmiana poznawcza, która już od samego początku formowała ścisły związek pomiędzy myślą, językiem i gestem. Do propozycji tej wracamy w sekcji 5.

## 3. PROBLEM – PRZEJŚCIE DO MOWY

Kluczowy problem zmiany modalności można podsumować w następujący sposób: *jeżeli język powstał jako system w przeważającej mierze gesturalny, jak wyjaśnić jego przejście do formy obecnej, czyli w przeważającej mierze głosowej?*

Jakkolwiek przejście od hipotetycznego języka głównie gesturalnego do obecnego systemu komunikacji językowej, która ma naturę głównie głosową, dostrzegane było jako poważny problem przez wcześniejszych autorów (Hewes, 1973), najpełniejszej krytyki dokonuje Fitch (2010). Podkreśla on zwłaszcza dwie trudności, a są nimi:

- skala anatomicznych i neuronalnych przystosowań człowieka do języka mówionego,
- kompletność i funkcjonalność obecnie istniejących języków migowych.

3.1. PRZYSTOSOWANIA *HOMO SAPIENS* DO MOWY

Po pierwsze, ludzie zasadniczo różnią się od blisko z nimi spokrewnionych małp człekokształtnych pod względem budowy i kontroli motorycznej przewodu głosowego. Cechami charakterystycznymi dla ludzi są m.in.:

- obniżona pozycja krtani,
- brak worów powietrznych (*air sacs*),
- lepsze unerwienie mięśni klatki piersiowej,
- rozwinięta zdolność imitacji głosowej.

Dla części z tych różnic nie są wykluczone alternatywne wyjaśnienia, np. obniżenie krtani może przynajmniej teoretycznie być wynikiem pionowej postawy, rekonfiguracji twarzy lub wyolbrzymienia rozmiaru ciała podczas wokalizacji (omówienie – np. Fitch, 2000); zanik worów powietrznych mógłby wynikać z ich dużej podatności na infekcje; lepsze unerwienie klatki piersiowej mogłoby być przystosowaniem do kontroli oddechu przy pokonywaniu długich dystansów; zdolność imitacji głosowej można by wiązać z muzyką. Jednak całość tych różnic trudno jest zinterpretować inaczej niż jako przystosowania do artykułowanej mowy.

Podsumowując, w ocenie komentatorów takich jak Fitch (2010) rozległość adaptacji do mowy stanowi bardzo mocny argument za stosunkowo wczesnym rozwinięciem się jej w ewolucji człowiekowatych, sporo przez pojawieniem się *H. sapiens*. Jest to zarazem argument przeciwko hipotezom gesturalnym tylko o tyle, o ile zakładamy ‘późne’ wykształcenie się języka. Fakty dotyczące adaptacji do mowy można uspoźnić z hipotezami gesturalnymi np. dopuszczając ‘wczesne’ pojawienie się mowy na bazie powstałej jeszcze wcześniej protojęzykowej komunikacji wizualnej, np. według scenariusza zakładającego gesturalny protojęzyk u *H. erectus* z późniejszym stopniowym rozwojem artykułowanej mowy. Dawne wyłonienie się języka jest zgodne z poglądem większości badaczy ewolucji języka oraz całością nowo dostępnych danych empirycznych. Na przykład najnowsze opracowania (Johansson, 2012; Dediu i Levinson, 2013) jednoznacznie konkludują, że rekonstrukcje genomu, anatomii (w tym anatomii przewodu głosowego), ontogenezy oraz kultury materialnej neandertalczyków – choć w żadnym razie nie dowodzą obecności języka u tego gatunku – są w całej rozciągłości spójne z taką możliwością. W świetle tych ustaleń hipoteza o niedawnym i nagłym pojawieniu się mowy jest mało prawdopodobna, natomiast nie kolidują one ze scenariuszami gesturalnymi, o ile te ostatnie zakładają odpowiednio dużą skalę czasową.

## 3.2. JĘZYKI MIGOWE JAKO PEŁNOPRAWNE JĘZYKI

Znacznie poważniejszą trudność stanowi *kompletność i funkcjonalność obecnie istniejących języków migowych*. Choć mogłoby się wydawać inaczej, jest to argument *przeciwko* hipotezom gesturalnym. Żeby zobaczyć dlaczego, konieczne jest przyjrzenie się naturze języków migowych.

Jezyki migowe są 'pełnoprawnymi' językami, które są równoważne językom mówionym (Stokoe, 1960; Stokoe i in., 1965; Emmorey, 2002). Dotyczy to zarówno własności kodu, implementacji mózgowej, rozwoju w ontogenezie (nabywania przez dzieci) oraz glottogenezie (powstania i rozwoju samego systemu językowego), jak i przede wszystkim aspektu funkcjonalnego. Ponieważ ich pełnoprawny status jest oficjalnie uznawany przez językoznawców, języki migowe podobnie jak te mówione figurują w klasyfikacjach czy bazach danych języków świata, np. *Ethnologue* lub *World Atlas of Linguistic Structures*<sup>9</sup>.

Jezyki migowe, podobnie jak mówione, posługują się głównie znakami o naturze arbitralnej i konwencjonalnej, które potrafią wyrażać dowolne abstrakcje czy znaczenia niedosłowne – czego ilustracją może być np. poezja migowa (np. Sutton-Spence, 2005). Podobnie jak mowa posiadają one kombinatoryczną strukturę na poziomie morfologii i składni, a nawet 'fonologii', a zatem jest w nich obecna 'podwójna artykulacja' (*duality of patterning*).

Ośrodki mózgowe przetwarzające język migowy są w dużej części tożsame z tymi przetwarzającymi mowę (Corina i in., 1992), a ich specjalizacja wydaje się dotyczyć przetwarzania języka jako takiego – niezależnie od modalności – a nie aspektów niejęzykowych, takich jak np. kontrola motoryczna artykulatorów (Emmorey, 2002). Są to więc ośrodki różne od tych odpowiadających za spontaniczne gestykulacje. Z tego względu w przypadku uszkodzenia ośrodków językowych u głuchoniemych obserwuje się objawy charakterystyczne dla afazji, łącznie z jej selektywnością, a więc zdarzają się przypadki afatycznego zaburzenia migania przy zachowanej zdolności do gestykulacji (np. Bellugi i Klima 2001; Hickok i in., 1996) lub zachowanej zdolności do pantomimy (Emmorey, 2002).

Nabywanie języka migowego przebiega na bardzo podobnej skali czasowej co nabywanie języka mówionego i charakteryzuje się obecnością tych samych stadiów rozwojowych i zjawisk, takich jak gaworzenie (jego manualny odpowiednik) czy nadmierna regularyzacja form gramatycznych; co więcej, dzieci nie wydają się wykazywać preferencji wobec języka mówionego względem migowego przy równoległej ekspozycji na nie (Petitto i Marentette, 1991).

<sup>9</sup> *Ethnologue*: <<http://www.ethnologue.com/>>, *World Atlas of Linguistic Structures*: <<http://wals.info/>>.

Niedawny przykład ISN (*Idioma de Signos Nicaragüense*), nikaraguańskiego języka migowego, ilustruje zjawisko wyłonienia się *de novo* kompletnego systemu językowego – proces którego stadia można mimo pewnych różnic porównać do pidginizacji i kreolizacji w językach mówionych (Kegl i in. 1999).

Co najważniejsze, badacze migów uznają, że pod względami potencjału komunikacyjnego i wydajności języki migowe nie ustępują mowie (Stokoe, 1960; Emmorey, 2002).

Mimo wspomnianych cech i mimo równoważności funkcjonalnej migów z mową, komunikacja werbalna u ludzi ma postać głosową, języki migowe pełnią zaś funkcje wtórne bądź pomocnicze<sup>10</sup>. Wyjątek stanowią jedynie te populacje, w których wydajna komunikacja głosowa jest znacznie utrudniona ze względu na powszechność występowania wad słuchu, jak to ma miejsce w wielu małych społecznościach na świecie, np. w Al-Sajid w Izraelu, w Adamorobe w Ghanie, w Kata Kolok na Bali, czy też niektórych miejscowościach w Jukatanie w Meksyku. We wszystkich pozostałych społecznościach językowych prymarną formą językowego porozumiewania się jest mowa. Fakt ten oznacza, że – jeżeli zgodnie z hipotezami gesturalnymi komunikacja językowa miała swoje początki w modalności wizualnej – w ewolucji człowiekowatych musiała zajść zmiana dominującej modalności; zmiana trudna do wyjaśnienia przy głębszej analizie. Fitch stawia zatem pytanie o powody ewolucyjne (presje selekcyjne) oraz mechanizm takiej zmiany<sup>11</sup>; liczni komentatorzy (np. Burling, 2005; Corballis, 2003; Kendon, 1991, 2008; MacNeillage, 2008; Tallerman, 2011) zgadzają się, że jest to najpoważniejszy problem hipotez gesturalnych.

---

<sup>10</sup> Użycie migów lub gestów w zastępstwie mowy zwykle motywowane jest względami *religijnymi*, jak w przypadku niektórych średniowiecznych zakonów monastycznych; względami *obyczajowymi*, jak w przypadku migów australijskich aborygenów Warlpiri bądź też względami *praktycznymi*, jak w przypadku południowoafrykańskich zbieracko-łowickich plemion San (Buszmenów) podczas polowania.

<sup>11</sup> Fitch (2010: 434): „[A] significant disadvantage of gestural models is their difficulty in explaining the virtually complete transition to vocal, spoken language in modern *Homo sapiens*... Whatever their virtues, models of gestural protolanguage are incomplete without a detailed and compelling model of the transition to spoken language, as most gestural proponents have recognized (Hewes, 1973; Corballis, 2002; Arbib, 2005)”.

(2010: 442): „but the lack of a plausible selective force to drive signed language into vocal language remains a compelling argument against a fully gestural, and fully linguistic, protolanguage”.

## 4. ROZWIĄZANIA

Przed omówieniem propozycji rozwiązań problemu zarysowanego w sekcji 3. należy zaznaczyć, że są one logicznie niezależne od argumentów za lub przeciw hipotezom gesturalnym, czy ogólniej: za lub przeciw użyteczności gestów – nawet jeśli są z nimi częściowo zbieżne. Poniżej proponujemy odpowiedzi na pytanie postawione w punkcie 3. – odpowiedzi biologicznie realistyczne, tj. pozwalające uzgodnić możliwość zmiany głównej modalności języka ze stanem wiedzy na temat anatomii, ewolucji oraz funkcjonowania mózgu.

Możliwe są dwa ogólne rodzaje takich odpowiedzi. Po pierwsze można wskazać na potencjalne presje selekcyjne faworyzujące rozwój porozumiewania się w kanale głosowym mimo ukształtowanej już bazy gesturalnej. Druga, bardziej interesująca i rozważana przez nas w sekcji 5., możliwość polega na zakwestionowaniu samego pojęcia 'zmiany modalności'. Zgodnie z tą propozycją, rozłączność komunikacji wzrokowej i dźwiękowej jest tylko pozorna, a ewolucja języka mogła przebiegać w obu tych modalnościach równocześnie. Mielibyśmy więc do czynienia nie tyle z zerojedynkową zamianą, ile zmieniającym się w filogenezie rozłożeniem akcentów oraz rodzaju przekazywanej informacji (niewerbalna, analogowa, holistyczna *versus* werbalna, symboliczna, kombinatoryczna) pomiędzy obie modalności.

## 4.1. ARGUMENTY TRADYCYJNE

Przegląd możliwych rozwiązań rozpoczynamy od propozycji już obecnych w literaturze i omawianych w niej w kontekście hipotez gesturalnych; spostrzeżenia te uznajemy za interesujące, ale niewystarczające jako rozwiązania rozpatrywanego problemu. Warto odnotować fakt, że poniższe punkty sprowadzają się niejako do wykazania *wad* komunikacji w kanale wzrokowym, zatem przy powierzchownej analizie mogłyby zostać użyte jako argumenty *przeciwko* hipotezom gesturalnym.

- Mowa jest bardziej *ekonomiczna* (np. Knight, 2000) – ruchy artykulacyjne wymagają mniej czasu i energii niż ruchy rąk i ciała;
- mowa umożliwia porozumiewanie się przy *gorszej widoczności lub w ciemności* (Rousseau, 1775);
- głos skuteczniej *przyciąga uwagę* (Rousseau, 1775);
- mowa *nie angażuje rąk*, które podczas porozumiewania się mogą być wykorzystane do celów praktycznych: praca, przenoszenie (np. Carstairs-McCarthy, 1996);
- mowa pozwala na *nauczanie czynności manualnych*, np. wytwarzania narzędzi (Armstrong i Wilcox, 2007);
- *nabywanie mowy* rozpoczyna się w *życiu płodowym* człowieka, co daje tej modalności przewagę rozwojową (Hewes, 1996);

- komunikacja głosowa umożliwia ciągle *monitorowanie położenia małego dziecka*, co mogło być istotne u człowiekowatych przy ich zbierackim trybie życia i w związku z tym częstym brakiem ciągłego fizycznego kontaktu matki z małym dzieckiem, jak to ma miejsce u innych małp (Falk, 2009);
- głos pozwala na kierowanie komunikatu *do ogółu* zamiast tylko do pojedynczego odbiorcy (Tomasello, 2008).

Do większości z wymienionych wyżej argumentów krytycznie odnosi się Fitch (2010). Wskazuje on, że trudno mówić o absolutnej wyższości mowy nad gestem w którymś z tych aspektów, ponieważ zależy ona od ekologii i innych czynników zewnętrznych, a dla wszystkich wskazanych zalet mowy łatwo znaleźć równoważące je przewagi komunikacji wizualnej. Gestów nie widać w ciemności, ale widać je w świetle ognia, a także można się nimi posłużyć w modalności dotykowej, co praktykują użytkownicy języków migowych. Kanał wzrokowy zyskuje przewagę w komunikacji na dalekie dystanse czy w warunkach hałasu, w tych sytuacjach również skutecznie przyciąga uwagę. Jak zauważa Fitch, choć modalność głosowa uwalnia ręce, modalność wizualna uwalnia jamę ustną, co musiało mieć zasadnicze znaczenie w paleolicie – dane kopalne wskazują, że człowiekowate bardzo intensywnie używały zębów do bardzo długotrwałego żucia twardych pokarmów, a także jako narzędzi do wykonywania innych czynności mechanicznych. Także argument o efektywności energetycznej nie jest przekonujący, ponieważ – jak wskazuje Fitch – mowie standardowo towarzyszy spontaniczna gestykulacja, co w końcowym efekcie czyni ten sposób komunikacji co najmniej równie kosztownym.

Również pozostałe, nie analizowane przez Fitcha argumenty okazują się niewystarczająco silne. Przy nauczaniu czynności manualnych od instrukcji werbalnej zdecydowanie efektywniejsza jest demonstracja czy fizyczne pokierowanie rękami adepta. Argument Hewesa jest sam w sobie zdecydowanie zbyt słaby, zwłaszcza w świetle danych rozwojowych wskazujących na co najmniej równie szybkie nabywanie języka migowego co mowy (zob. sekcja 3.2.). Spostrzeżenie Falk jest ciekawe, jednak nie wymaga założenia o języku mówionym – artykułowanym i przenoszącym propozycjonalną treść – a jedynie wydawaniu dźwięków. Równie interesująca jest propozycja Tomasello, ale ją także łatwo zrównoważyć. Wspominaną już zaletą gestu jest większa tajność komunikacji, pozwalająca na lepszy dobór adresatów przekazu, a także na ograniczenie groźby wykrycia przez wrogów oraz drapieżniki.

## 4.2. DWOISTOŚĆ INFORMACJI

Susan Goldin-Meadow (np. 2011) zauważa, że w modalności wizualnej można efektywnie przekazać zarówno informację *kombinatoryczną-posegmentowaną*, jak i *mimetyczną* (holistyczno-wyobrażeniową). Tę pierwszą możliwość realizują migi – odrębne jednostki o charakterze dyskretnym i arbitralnym, z których kompozycjonalnie zestawiane są dłuższe ciągi (frazy, zdania). Z drugą możliwością mamy do czynienia w przypadku gestykulacji czy pantomimy, w których komunikat nie składa się z kombinacji dyskretnych jednostek, lecz ma naturę całościową i wyobrażeniową. Goldin-Meadow zauważa jednocześnie, że modalność głosowa nadaje się do przekazywania jedynie informacji kombinatorycznej-posegmentowanej, polegającej na składaniu ze sobą dyskretnych jednostek, jakimi są np. fonemy i morfemy. Natomiast możliwości przekazywania głosem informacji mimetycznej są ograniczone do elementów prozodii oraz zjawisk typu onomatopei czy symbolizmu dźwiękowego, ich rola jest więc drugorzędna.

W naturalnych warunkach, tj. w rozmowie dwóch lub kilku osób, następuje *jednoczesny* wydajny przekaz obu powyższych rodzajów informacji. Upraszczając, ludzie mówiąc do siebie praktycznie zawsze jednocześnie gestykulują (Goldin-Meadow, 2003; Kendon, 2004) – robią tak nawet osoby rozmawiające przez telefon oraz niewidomi. Modalność wizualna mogłaby przekazywać kod posegmentowany, ale głos nie przekaże mimetycznej informacji holistyczno-wyobrażeniowej. Dlatego efektywna komunikacja w naturalnej konwersacji ma postać mowy (która koduje informacje w sposób arbitralny i posegmentowany) wraz z towarzyszącą gestykulacją, która koduje informację w sposób całościowy. Goldin-Meadow (2008) zauważa, że właśnie ta przewaga modalności wizualnej mogła paradoksalnie stanowić przyczynę przejścia od hipotetycznego protojęzyka gestów do mowy.

Zbliżoną propozycję znajdujemy u Erin Brown (za: Zlatev, 2013). Według niej modalność głosowa zaczęła wyrażać głównie kod symboliczny, ponieważ dźwięk z natury swej słabo nadaje się do przenoszenia znaczeń niearbitralnych (motywowanych) – a to automatycznie sprawia, że wokalizacji łatwiej przypisać jest znaczenie arbitralne. W podobnym duchu Kendon<sup>12</sup> zauważa, że gest to znak nieodzownie mający wyraźną przestrzenną formę i lokalizację, a więc swego rodzaju 'przestrzenną konkretność', która z kolei nie cechuje mowy. Jest to istotne przy przekazywaniu znaczeń 'oderwanych od tu i teraz' (*displaced*), czyli np. odnoszeniu się do przyszłości i przeszłości, a także znaczeń abstrakcyjnych. Konkretność przestrzenna gestu może być trudna do zignorowania i w ten sposób może mentalnie blokować dotarcie do jego abstrakcyjnego znaczenia, natomiast znak dźwiękowy, niema-

<sup>12</sup> Prezentacja na konferencji *Protolang 2*, 19.09.2011, Toruń.

jący wyrazistych własności przestrzennych, łatwiej daje się interpretować w kategoriach abstraktu. Pokrewna intuicja – o większej arbitralności mowy oraz nadmiernej konkretności ‘przytłaczającej’ tym samym odbiór abstrakcji – obecna jest także już u Hewesa (1973, 1996: 587), który proponuje symbolizm dźwiękowy jako mechanizm ‘przejęciowy’, wprowadzający przynajmniej częściowe umotywowanie znaku.

#### 4.3. NABYWANIE JĘZYKA MIGOWEGO I MÓWIONEGO PRZEZ DZIECI

Małe dzieci nabywają języki migowe z łatwością porównywalną do akwizycji języka mówionego. Jak wspomniano w sekcji 3.2., w procesie tym daje się również zauważyć paralelne stadia, m.in. manualny odpowiednik stadium gaworzenia. Petitto i Marentette (1991) stwierdzają, że łatwość nabywania migów dotyczy również *słyszących* dzieci, które jeżeli mają równoczesny dostęp do języka migowego *oraz* mowy, przyswajają oba rodzaje języka równie sprawnie, nie wykazując widocznych preferencji w kierunku mowy i osiągając te same etapy rozwoju jednocześnie w obu modalnościach. Ontogeneza dostarcza więc argumentów za tym, że ‘przejście’ do modalności głosowej jest w pewnym sensie pozorne, gdyż dotyczy poziomu E-języków (w sensie Chomsky’ego), a nie samej zdolności językowej *language faculty* jako (zestawu) biologicznych adaptacji. Właściwa człowiekowi biologiczna gotowość do nabycia języka wydaje się mieć naturę amodalną i pozwalającą na realizację tak w jednej, jak i w drugiej modalności – obie są równodostępne w ontogenezie. Wskazuje to na jedynie powierzchowny charakter ‘przejścia na głos’. Znajduje to swoje poparcie także w danych neurofizjologicznych, ponieważ jak wspominaliśmy wyżej informacja językowa, bez względu na jej modalność (słowo mówione, znak migowy), przetwarzana jest w tych samych częściach mózgu. Dotyczy to zarówno produkcji znaków, jak też ich percepcji.

Powiązanie obu modalności widoczne jest także podczas nabywania języka mówionego i występuje także w przypadku pojedynczych gestów. Gaworzeniu (6–8 miesiąc życia) towarzyszą rytmiczne ruchy rąk. Przed wypowiedzeniem pierwszego słowa dzieci używają już gestów deiktycznych, a czasami także ikonicznych (zwykle w okolicach 10 m. ż.). W późniejszym okresie zaczynają stosować kombinacje słowo-gest, przenoszące to samo znaczenie, a jeszcze później kombinacje słowo-gest przenoszące różne znaczenia (np. *daj*, wskazując na owoc). Udział obu modalności uwidacznia się także w toku rozwoju poznawczego w późniejszym wieku. Goldin-Meadow (2003) wykazała, że przynajmniej niektóre nowo nabywane pojęcia (stałość ilości, równoważność), zanim znajdą reprezentację werbalną, najpierw mogą być reprezentowane w gestach.

Co ciekawe, liczne badania eksperymentalne, prowadzone głównie przez zespół Susan Goldin-Meadow, świadczą, że wczesne użycie gestów przez



małe dzieci stanowi predyktor szerokiej gamy ich późniejszych osiągnięć językowych. Dla przykładu, wzorzec użycia gestów przez 18-miesięczne dzieci pozwala przewidzieć rozmiar ich zasobów leksykalnych oraz złożoność wypowiedzanych zdań (Rowe i Goldin-Meadow, 2009). Innym przykładem są wspomniane w poprzednim akapicie kombinacje słowo-gest: wiek, w którym dzieci potrafią zastosować to połączenie może być dobrym predyktorem wieku, w którym dziecko zacznie posługiwać się wypowiedziami dwuwyrzowymi (np. *daj jabłko*).

#### 4.4. NATURALNE ZWIĄZKI RĘKA-USTA

Podział komunikacji na przebiegającą w kanale głosowym oraz wizualnym jest użyteczną idealizacją, jednak *de facto* modalności te są ze sobą powiązane. Ma to miejsce na wielu płaszczyznach: tak 'powierzchniowym' poziomie przekazu sygnału, jak i na 'głębokim' poziomie kognitywnej oraz neuronalnej implementacji odpowiednich systemów.

Ruchy ust i ruchy dłoni zdają się w dużym stopniu zawiadywane przez wspólny i stary filogenetycznie system kontroli motorycznej. Wskazuje na to zarówno sąsiedztwo obszarów korowych kontrolujących ruchy tych narządów, jak i dane eksperymentalne. Na przykład stymulacja elektryczna pola 44 (homolog części obszaru Broki u ludzi) u rebusów wywołuje ruchy kończyny górnej oraz warg (Petrides i in., za: Meguerditchian, Cochet i Vauclair, 2011; zob. też Corballis, 2003). Kora przedruchowa makaków (F5) zawiera oprócz neuronów lustrzanych także neurony aktywujące się na widok obiektu, który może zostać chwycony, kodując jego wielkość w celu przygotowania odpowiedniego planu wykonania chwytu jednocześnie za pomocą dłoni i ust (Murata i in., 1997, Rizzolatti i in. 1998). Naturalne związki motoryczne między dłonią a ustami u ludzi dokumentują Gentilucci i Corballis (2006), opisując m.in. różnice w rozwarciu ust i spektrum głosowym u badanych wypowiadających zadaną sylabę w zależności od tego, czy podczas wypowiadania chwytają mały, czy duży obiekt. Podobna prawidłowość zachodzi także wtedy, kiedy badani tylko obserwują czynność chwytania wykonywaną przez inną osobę, co świadczy o zaangażowaniu systemu lustrzanego.

Przypuszcza się, że naturalny związek ręka-usta ma swe źródła w zachowaniach związanych z podawaniem pokarmu do ust, natomiast później został egzaptacyjnie wyzyskany do zadań językowych. Obwód ten mógł odegrać rolę w transformacji komunikacji gestowej: przejścia od gestów rąk do gestów ust (Gentilucci i Corballis, 2006). Koncepcja związków ręka-usta ma zaskakująco długą historię, której początków należy szukać już w dość egzotycznej propozycji „gestów języka” Richarda Pageta<sup>13</sup>.

<sup>13</sup> Według Pageta (1930, zob. Kendon, 2011) język i inne artykulatory są w naturalny sposób związane motorycznie z rękami – dzięki temu mogły nieświadomie podążać

## 4.5. RUCHY ARTYKULACYJNE JAKO RODZAJ GESTÓW

Interesujące możliwości rozwiązania problemu przejścia od gestów do mowy daje motoryczna teoria percepcji mowy (Lieberman i in., 1967; Liberman i Mattingly, 1985; Liberman i Whalen, 2000), utrzymująca, że system ruchowy odpowiada nie tylko za produkcję mowy, ale i za jej rozumienie. Według tej teorii nasze mentalne reprezentacje fonemów mają postać nie parametrów fizycznych dźwięku, lecz programów motorycznych, a rozpoznawanie danego fonemu zachodzi nie poprzez jego wartość dźwiękową, lecz przez odwołanie do sekwencji ruchowej języka i innych artykulatorów, która pozwoliłaby nam na wyprodukowanie danego dźwięku. Oznacza to prymarność aspektu ruchowego wobec akustycznego i pozwala konceptualizować mowę *jako gest*, tj. jako system gestów okolic twarzy i ust, który niejako dodatkowo produkuje dźwięk.

Konceptualizacja mowy jako rodzaju gestu pozwala zauważyć, że ta forma komunikacji obok oczywistego komponentu głosowego w istotny sposób wykorzystuje także komponent wizualny. Najbardziej oczywistą tego manifestację stanowi umiejętność czytania z ruchu warg, która daje możliwość odtworzenia znacznej części komunikatu nawet bez dostępu do dźwięku (Summerfield, 1992). Świadectwem głębokiej integracji informacji wzrokowej z dźwiękiem przy percepcji mowy jest także efekt McGurka (McGurk i MacDonald, 1976): podłożenie dźwięku sylaby *ga* pod zapis video, na którym widoczne są usta wymawiające sylabę *ba* powoduje, że odbiorca słyszy w efekcie tego zabiegu sylabę *da*. Dane te jednoznacznie wskazują na multimodalny charakter percepcji mowy i ważną w niej rolę komponentu ruchowo-wizualnego.

## 4.5.1. GESTY OKOLICY TWARZY I UST

Gesty okolic twarzy i ust (*orofacial gestures*) zgodnie z założeniami teorii pierwszeństwa gestów mogły pierwotnie pełnić funkcję komunikacyjną w modalności wizualnej, po czym dźwięki towarzyszące ich produkcji mogły same nabrać znaczenia komunikacyjnego, w ten sposób umożliwiając przejście do mowy. Zauważmy, że taki pomysł sam w sobie nie identyfikuje presji selekcyjnej odpowiedzialnej za takie przejście, ale jest cenny ze względu na propozycję biologicznie realistycznego mechanizmu. Potencjał *orofacial gestures* nie umknął uwadze badaczy ewolucji języka, stając się ważnym elementem licznych scenariuszy, zwłaszcza nowych koncepcji o profilu gesturalnym.

---

za rękami wykonującymi gesty i same wykonywać „gesty języka” (*tongue gestures*), z których z kolei mogły powstać dźwięki mowy. Już współcześni Pagetowi komentatorzy krytycznie odnieśli się do tej propozycji (zob. Kendon, 2011; Fitch, 2010).

- Hewes (np. 1973), jak wspominaliśmy, uważał gesty ust (*mouth gestures*) za drugi, obok symbolizmu dźwiękowego, z możliwych łączników między gestem a mową.
- Corballis (np. 2003) jest zdania, że za włączenie komponentu akustycznego mogła odpowiadać potrzeba dokładniejszego różnicowania podobnych gestów twarzowych, które mogły być lepiej identyfikowane poprzez dodanie właśnie dźwięku.
- Studdert-Kennedy (2005) zakłada, że ekspresje mimiczne odegrały główną rolę w nabyciu, niezbędnej do artykulacji, kontroli nad poszczególnymi organami szlaku głosowego. Powtarzające się akty naśladownictwa wyrazów mimicznych doprowadziły do zdobycia niezależnej kontroli nad różnymi częściami twarzy, co z racji istniejących już wcześniej powiązań pomiędzy mimiką a wokalizacją przyczyniło się do rozszerzenia kontroli także na artykulatory głosowe.
- MacNeilage (1998, 2008) nie będąc zwolennikiem teorii gesturalnych, widzi jednak ważną rolę gestów twarzowych. Wychodząc od podobieństwa pomiędzy mową a wytwarzającymi dźwięk gestami twarzowymi u małp (cmokanie, mlaskanie), proponuje następujący scenariusz: naprzemienne otwieranie i zamykanie ust połączone z żuciem, ssaniem i lizaniem zaczęło spełniać funkcje komunikacyjne i przybrało postać gestów twarzowych, a w dalszej kolejności cykle otwieranie-zamykanie ust uległy transformacji do sylab oraz głosek.
- Meguerditchian, Cochet i Vauclair (2011) podkreślają istotność gestów twarzowych głównie w oparciu na przeglądzie danych neurobiologicznych z badań na małpach, ujawniających głębokie związki między ręką a ustami, które opisywaliśmy w sekcji 4.4.
- Orzechowski, Wacewicz i Żywiczyński (w druku), zwracają uwagę na kluczową rolę słuchowej informacji zwrotnej (*auditory feedback*). Autorzy ci rozszerzają propozycję Corballisa, postulując, że korzyść z dodania komponentu dźwiękowego nie ogranicza się jedynie do odbiorcy, ale dotyczy też nadawcy komunikatu, któremu łatwiej jest rozróżnić odpowiednie gesty przy ich produkcji. Stanowisko to znajduje wsparcie w badaniach empirycznych, wskazujących, że zakłócona (np. opóźniona) słuchowa informacja zwrotna lub jej brak prowadzi do poważnych zaburzeń wymowy.

## 5. KONKLUZJA – W STRONĘ HIPOTEZ MULTIMODALNYCH?

W każdym ze skrajnych stanowisk dotyczących źródeł języka występuje kilka zasadniczych trudności. Teorie wywodzące język z wokalizacji muszą zmierzyć się ze słabością wyjściowych przesłanek dla późniejszych przystosowań językowych: przy głębszej analizie wokalizacje naczelných oraz język to

dwa skrajnie różne systemy komunikacji, mające niewiele wspólnego poza właśnie wykorzystaniem modalności dźwiękowej. Z kolei radykalna perspektywa gesturalna, zakładająca istnienie najpierw języka gestowego, boryka się z problemem przejścia od modalności wizualnej do głosowej. Ponadto żadne ze stanowisk nie potrafi dostarczyć przekonującego wyjaśnienia przyczyn głębokiej integracji gestykulacji z językiem.

Na rozwiązanie tych trudności pozwala przyjęcie perspektywy multimodalnej: postrzeganie gestu i głosu nie jako dwóch odrębnych systemów komunikacyjnych, lecz jako odmiennych realizacji jednego, wspólnego systemu. Jej konsekwencją jest wizja ewolucji języka od początku angażującej obie modalności w sposób silnie zintegrowany. W jej ramach można dopuszczać czasową specjalizację lub przewagę jednej z modalności, wynikającą np. ze stopnia rozwoju poszczególnych przystosowań językowych do panujących wymagań środowiskowych. Może to stanowić siłę hipotez gesturalnych, gdyż uprawdopodobnia scenariusz, według którego w początkowych stadiach rozwoju protojęzyka znak w modalności wzrokowej był istotniejszym źródłem informacji niż towarzyszące mu wokalizacje.

Perspektywa multimodalna pozwala na najbardziej naturalną integrację argumentów wymienione w sekcjach 4.2.–4.5. właśnie w sposób bliski podejściu gesturalnemu, tj. jednak uwypuklając znaczenie modalności wizualnej na wczesnych etapach kształtowania się ludzkiej zdolności językowej.

Zwolennikami perspektywy multimodalnej są przede wszystkim gesturologi, tacy jak Kendon (2011), McNeill (2012) czy Sandler (2013). Wynika to z faktu, że ich badania unaocniają wspomniany już bardzo głęboki związek gestykulacji z mową (zob. sekcja 1.3.1.). Gestykulacje są integralną częścią konwersacji, a zatem gestykulujemy np. w trakcie rozmowy telefonicznej, mimo pełnej świadomości, że rozmówca nie może nas dostrzec (Bavelas i in. 2008). Szczególne znaczenie mają dane z badań nad ludźmi niewidomymi od urodzenia, a więc takimi, którzy nie mieli możliwości nabycia wzorców gestykulacji poprzez obserwację wzrokową. Jak się okazuje, takie osoby również używają gestykulacji i to nawet jeśli ich rozmówcą jest inna niewidoma osoba (Iverson i Goldin-Meadow, 1997). O stabilności integracji gestykulacji z mową świadczą też dane neuronaukowe, np. dysocjacja względem migów wspomniana w sekcji 3.2., czy dysocjacja względem ruchów instrumentalnych, tj. mających efekt mechaniczny, jak np. chwytanie. To ostatnie ilustruje przypadek I.W. (McNeill, 2005), który wobec utraty proprioceptywnego schematu ciała nie jest w stanie wykonywać ruchów instrumentalnych, nie patrząc jednocześnie na ręce – jednakże jego gestykulacje nie podlegają podobnym ograniczeniom.

Perspektywa multimodalna zyskuje jednak rosnącą popularność także w innych kręgach badaczy ewolucji języka. Na multimodalność komunikatów naturalnie występujących u małp wskazują prymatolodzy. Z jednej

strony gestom, czy szerzej zdefiniowanym czynnościom o charakterze komunikacyjnym (np. bębnieniu czy potrząsaniu gałęzią), często towarzyszy zamierzona lub niezamierzona produkcja dźwięku (np. Hobaiter i Byrne, 2012). Z drugiej strony, jak przypomina choćby Falk (2009), wokalizacje naczelnych często sztywno sparowane są zarówno z określonym stanem emocjonalnym, jak i z określoną konfiguracją twarzy (ekspresją mimiczną), która może być postrzegana wzrokowo. Także na ekspresje mimiczne jako ważne, ale zaniedbany badawczo komponent „całościowego komunikatu” zwracają szczególną uwagę Slocombe i in. (2011). W swym przeglądzie literatury na temat komunikacji naczelnych autorki te identyfikują zdecydowaną tendencję do przyjmowania uproszczonej perspektywy unimodalnej i apelują o zwrot w kierunku badań multimodalnych.

Do przyjęcia perspektywy multimodalnej skłaniają także ustalenia neuronauk. Punktem wyjścia jest spostrzeżenie, że podział na modalności głosową i wizualną nie pokrywa się z kryterium 'językowości' przekazu. Przypominają o tym języki migowe, ale także naturalna komunikacja niemająca natury językowej. Np. Gonseth, Vilain i Vilain (2012) przekonują, że wskazywanie jest *de facto* czynnością rozłożoną pomiędzy modalności, w której indeksykarność aktu komunikacyjnego jest najczęściej osiągnięta w wyniku współpracy wokalizacji ze wskazaniem manualnym. Wydaje się to potwierdzać poziom mózgowy, tj. dla implementacji neuronalnej danego systemu decydujący czynnik stanowi nie modalność, lecz typ przetwarzania, np. ikoniczność czy konwencjonalność (Niederhut, 2012). Przyjęcie możliwości, że najważniejsze z punktu widzenia języka dysocjacje nie przebiegają według podziału na modalności, jest szczególnie obiecujące dla całego kierunku badań.

Publikacja powstała w ramach projektu sfinansowanego z grantu UMO-2012/07/E/HS2/00671 Narodowego Centrum Nauki.

#### LITERATURA CYTOWANA

- Arbib, M.A. (2002). *The Mirror System, Imitation, and the Evolution of Language*. W: K. Dautenhahn, Ch. Nehaniv (red.). *Imitation in Animals and Artifacts*. London: A Bradford Book.
- Arbib, M.A. (2005). From monkey-like action recognition to human language: an evolutionary framework for neurolinguistics. *Behavioral and Brain Sciences*, 28, s. 105–167.
- Arbib, M.A. (2012). *How the brain got language*. Oxford: Oxford University Press.
- Armstrong, D.F., Wilcox, S.E. (2007). *The gestural origin of language*. Oxford: Oxford University Press.
- Armstrong, D.F., Stokoe, W.C., Wilcox, S.E. (1995). *Gesture and the nature of language*. Cambridge: Cambridge University Press.

- Bavelas, J., Gerwing, J., Sutton, C., Prevost, D. (2008). Gesturing on the telephone: Independent effects of dialogue and visibility. *Journal of Memory and Language*, 58, s. 495–520.
- Bellugi, U., Klima, E. (2001). Sign Language. W: N. Smelser, P. Baltes (red.) *International Encyclopedia of the Social and Behavioral Sciences vol. 21* (s. 14066–14071). Oxford: Elsevier Science Publishers.
- Bickerton, D. (1990). *Language and species*. Chicago: University of Chicago Press.
- Burling, R. (2005). *The talking ape: How language evolved*. Oxford: Oxford University Press.
- Campbell, L. (1998). *Historical Linguistics: An Introduction*. Edinburgh: Edinburgh University Press.
- Calvin, W.H. (1982). Did throwing stones shape hominid brain evolution?. *Ethology and Sociobiology*, 3, s. 115–124.
- Calvin, W.H. (1983). *The throwing Madonna: Essays on the brain*. New York: McGraw-Hill.
- Calvin, W.H., Bickerton, D. (2000). *Lingua ex machine: reconciling Darwin and Chomsky with the human brain*. Cambridge: MIT Press.
- Carstairs-McCarthy, A. (1996). Review of Armstrong, Stokoe i Wilcox 'Gesture and the nature of language'. *Lingua*, 99, s. 135–138.
- Cashmore, L., Uomini, N., Chapelain, A. (2008). The evolution of handedness in humans and great apes: a review and current issues. *Journal of Anthropological Sciences*, 86, s. 7–35
- Corballis, M.C. (2002). *From hand to mouth: The origins of language*. Princeton, NJ: Princeton University Press.
- Corballis, M.C. (2003). From mouth to hand: gesture, speech, and the evolution of right-handedness. *Behavioral and Brain Sciences*, 26 (2), s. 199–208.
- Corina, D.P., Vaid, J. i Bellugi, U. (1992) The linguistic basis of left hemisphere specialization. *Science*, 255, s. 1258–1260.
- Deacon, T.W. (1997). *The symbolic species: The coevolution of language and the brain*. New York: W.W. Norton and Company/London: Penguin Press.
- Dediu, D., Levinson, S.C. (2013). On the antiquity of language: the reinterpretation of Neandertal linguistic capacities and its consequences. *Frontiers in Psychology*, 4, s. 397.
- de Waal, F.B.M., Pollick, A. S. (2011). Gesture as the most flexible modality of primate communication. W: K.R. Gibson i M. Tallerman (red.), *The Oxford handbook of language evolution* (s. 82–89). Oxford: Oxford University Press.
- Donald M. (1991). *Origins of the modern mind: Three stages in the evolution of culture and cognition*. Cambridge: Harvard University Press.
- Donald, M. (2001). *A Mind So Rare. The Evolution of Human Consciousness*. New York: Norton. Emery, N.J.
- Dunbar, R. (1998). *Grooming, gossip and the evolution of language*. Harvard University Press.
- Ekman, P., Friesen, W.V. (1969). The repertoire of nonverbal behavior: Categories, origins, usage, and coding. *Semiotica*, 1, s. 49–98.
- Emmorey, K. (2002). *Language, cognition, and brain: Insights from sign language research*. Hillsdale, NJ: Lawrence Erlbaum.

- Fadiga L., Craighero L., D'Ausilio A. (2009) Broca's Area in language, action and music. *Annals of the New York Academy of Sciences*, 1169, s. 448–458.
- Falk, D. (2009). *Finding our tongues: Mothers, infants and the origins of language*. New York: Basic Books.
- Fay, N., Arbib, M., Garrod, S. (2013). How to Bootstrap a Human Communication System. *Cognitive Science*, 37, s. 1356–1367.
- Fisiak, J. (1985). *Wstęp do współczesnych teorii lingwistycznych*. Warszawa: WSiP.
- Fitch, W.T. (2010). *The evolution of language*. Cambridge: Cambridge University Press.
- Fogassi, L., Ferrari, P. F. (2004). Mirror neurons, gestures and language evolution. *Interaction Studies*, 5, 345–363.
- Gardner, R.A., Gardner, B.T. (1969). Teaching sign language to a chimpanzee. *Science*, 165, s. 664–672.
- Gentilucci, M., Corballis, M.C. (2006). From manual gesture to speech: A gradual transition. *Neuroscience and Biobehavioral Reviews*, 30, s. 949–960.
- Goldin-Meadow, S. (2003). *Hearing Gesture: How Our Hands Helps Us Think*. Cambridge: The Belknap Press of Harvard University Press.
- Goldin-Meadow, S. (2005). What language creation in the manual modality tells us about the foundations of language. *The Linguistic Review*, 22, s. 199–225.
- Goldin-Meadow, S. (2008). Gesture, Speech, and Language. W: A. Smith, K. Smith, R. Ferrer-i-Cancho (red.) *Proceedings of the 7th International Conference on the Evolution of Language* (s. 427–428). London: World Scientific.
- Goldin-Meadow, S. (2011). What modern-day gesture can tell us about language evolution. W: M. Tallerman i K.R. Gibson (red.), *The Oxford handbook of language evolution* (s. 545–557). Oxford: Oxford University Press.
- Goldin-Meadow, S., So, W.C., Özyürek, A., Mylander, C. (2008). The natural order of events: How speakers of different languages represent events nonverbally. *Proceedings of the National Academy of Sciences of the USA*, 105, s. 9163–9168.
- Gonseth, C., Vilain, A., Vilain, C. (2012). Ontogeny of two communicative tools: Distance encoding and multimodality in deictic pointing. W: T. Scott-Phillips, M. Tamariz, E. Cartmill, J. Hurford (red.), *The Evolution of Language. Proceedings of the 9th International Conference (EVOLANG 9)* (s. 150–157). New Jersey: World Scientific.
- Harrison, R. (2008). What can the study of handedness in nonhuman apes tell us about the evolution of language. W: A. D. M. Smith, K. Smith, R. Ferrer-i-Cancho (red.), *Proceedings of the 7th International Conference on the Evolution of Language* (s. 431–432). World Scientific.
- Hauser, M.D., Chomsky, N., Fitch W.T. (2002). The faculty of language: What is it, who has it, and how did it evolve?. *Science*, 298, 1569–1579.
- Hewes, G.W. (1973). Primate communication and the gestural origins of language. *Current Anthropology*, 14, s. 5–24.
- Hewes, G.W. (1981). Pointing and language. W: T. Myers, J. Laver, J. Anderson (red.), *The cognitive representation of speech* (s. 263–269). Amsterdam: North-Holland.
- Hewes, G.W. (1996). A history of the study of language origins and the gestural primacy hypothesis. W: A. Lock, C.R. Peters (red.), *Handbook of human symbolic evolution* (s. 571–595). Oxford: Oxford University Press.

- Hickok, G., Bellugi, U., Klima, E.S. (1996) The neurobiology of sign language and its implications for the neural basis of language. *Nature*, 381, s. 699–702.
- Hobaiter, C., Byrne, R. W. (2012). Gesture use in consortship. *Developments in Primate Gesture Research*, 6, 129–146.
- Hockett, C.F. (1960). The Origin of Speech. *Scientific American*, 203, 88–96.
- Hopkins, W.D. (2006). Comparative and familial analysis of handedness in great apes. *Psychological Bulletin*, 132, s. 538.
- Hurford, J. (2007). *The origins of meaning. Language in the light of evolution*. Oxford: Oxford University Press.
- Hutto, D.D. (2008). First communions: Mimetic sharing without theory of mind. W: J. Zlatev, T. Racine, C. Sinha, E. Itkonen (red.), *The Shared Mind: Perspectives on intersubjectivity* (s. 246–276). Amsterdam: John Benjamins Publishing.
- Iacoboni, M., Woods, R.P., Brass, M., Bekkering, H., Mazziotta, J.C., Rizzolatti, G. (1999). Cortical mechanisms of human imitation. *Science*, 286, s. 2526–2528.
- Iverson, J., Goldin-Meadow, S. (1997). What communication got to do with it? Gesture in congenitally blind children. *Developmental psychology*, 33, 453–467.
- Johansson, S. (2005). *Origins of language. Constraints on hypothesis*. Amsterdam: John Benjamins Publishing.
- Johansson, S. (2012). The case for Neanderthal language – how strong is it?. W: T. Scott-Phillips, M. Tamariz, E. Cartmill, J. Hurford (red.), *The Evolution of Language. Proceedings of the 9th International Conference (EVOLANG 9)* (s. 173–180). New Jersey: World Scientific.
- Kendon, A. (1991). Some considerations for a theory of language origins. *Man*, 26, s. 199–221.
- Kendon, A. (2004). *Gesture: Visible action as utterance*. Cambridge: Cambridge University Press.
- Kendon, A. (2008). Signs for Language Origins? *The Public Journal of Semiotics*, II(2), s. 2–29.
- Kendon, A. (2011). Some modern considerations for thinking about language evolution: A discussion of The Evolution of Language by Tecumseh Fitch. *The Public Journal of Semiotics*, III (1), s. 79–108.
- Kenneally, Ch. (2007). *The first word*. New York: Viking.
- Kegl, J., Senghas, A., Coppola, M. (1999). Creation through contact: Sign language emergence and sign language change in Nicaragua. W: M. DeGraff (red.), *Language Creation and Language Change: Creolization, Diachrony, and Development* (s. 179–237). Cambridge: MIT Press.
- Knecht, S., Dräger, B., Deppe, M., Bobe, L., Lohmann, H., Flöel, A., Ringelstein E.-B., Henningsen, H. (2000). Handedness and hemispheric language dominance in healthy humans. *Brain*, 123, s. 2512–2518.
- Knight, C. (2000). Play as precursors of phonology and syntax. W: Ch. Knight, M. Studdert-Kennedy, J. Hurford (red.), *The evolutionary emergence of language* (s. 99–119). Cambridge: Cambridge University Press.
- Krauss, R.M., Chen, Y., Gottesman, R.F. (2000). Lexical gestures and lexical access: A process model. W: D. McNeill (red.), *Language and gesture* (s. 261–283). New York: Cambridge University Press.



- Lieberman, A.M., Cooper, F.S., Shankweiler, D.P., Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological review*, 74, s. 431–461.
- Lieberman, A.M., Mattingly, I.G. (1985). The motor theory of speech perception revisited. *Cognition*, 21, s. 1–36.
- Lieberman, A.M., Whalen, D.H. (2000). On the relation of speech to language. *Trends in cognitive sciences*, 4, s. 187–196.
- Lieberman, Ph. (2003). Motor control, Speech, and the Evolution of Human Language. W: M.H. Christiansen, S. Kirby (red.), *Language Evolution* (s. 255–271). New York: Oxford University Press.
- MacNeilage, P.F. (1998). The frame/content theory of the evolution of speech production. *Behavioral and brain sciences*, 21, s. 499–556.
- MacNeilage, P.F. (2008). *The Origin of Speech*. Oxford: Oxford University Press.
- Maestriperi, D. (2007). Gestural communication in three species of macaques (*Macaca mulatta*, *M. nemestrina*, *M. arctoides*): Use of signals in relation to dominance and social context. W: K. Liebal, C. Muller, S. Pika (red.), *Gestural Communication in Nonhuman and Human Primates* (s. 53–68). Amsterdam: John Benjamins Publishing.
- McGurk, H., MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, 264, s. 746–748.
- McNeill, D. (1992). *What Gestures Reveal About Thought*. Chicago: The University of Chicago Press.
- McNeill, D. (2005). *Gesture and Thought*. Chicago: The University of Chicago Press.
- McNeill, D. (2012). *How Language Began: Gesture and Speech in Human Evolution*. Cambridge: Cambridge University Press.
- Meguerditchian, A., Cochet, H., Vauclair, J. (2011). From gesture to language: ontogenetic and phylogenetic perspectives on gestural communication and its cerebral lateralization. W: A. Vilain, J. Schwartz, Ch. Abry, J. Vauclair (red.), *Primate communication and human language: vocalisation, gestures, imitation and deixis in humans and non-humans* (s. 89–118). Amsterdam: John Benjamins Publishing.
- Mithen, S. (2005). *The singing Neanderthals: The origins of music, language, mind and body*. London: Weidenfeld and Nicholson.
- Murata, A., Fadiga, L., Fogassi, L., Gallese, V., Raos, V., Rizzolatti G. (1997). Object representation in the ventral premotor cortex (area F5) of the monkey. *Journal of neurophysiology*, 78(4), s. 2226–30.
- Niederhut, D. (2012). Gesture and the origin of language. *The Evolution of Language. Proceedings of the 10th International Conference (EVOLANG 10)* (s. 266–273).
- Orzechowski, S., Waciewicz, S., Żywicznyński, P. (w druku). Orofacial gestures in language evolution. The auditory feedback hypothesis.
- Petitto, L.A., Marentette, P.F. (1991). Babbling in the Manual Mode: Evidence for the Ontogeny of Language. *Science*, 251, s. 1493–1496.
- Pika, S. (2008). What is the nature of the gestural communication of great apes? W: J. Zlatev, T. Racine, C. Sinha, E. Itkonen (red.), *The Shared Mind: Perspectives on intersubjectivity* (s. 165–186). Amsterdam: John Benjamins Publishing.
- Pika, S., Liebal, K. (2006). Differences and similarities between the natural gestural communication of the great apes and human children. W: *Proceedings of the 6th International Conference on the Evolution of Language (Evolang 6)* (s. 267–274).

- Pika, S., Mitani, J.C. (2009). The directed scratch: Evidence for a referential gesture in chimpanzees? W: R. Botha, Ch. Knight (red.), *The prehistory of language* (s. 166–180). Oxford: Oxford University Press.
- Pika, S., Liebal, K., Call, J., Tomasello, M. (2005). The gestural communication of apes. *Gesture*, 5, s. 41–56.
- Pinker, S., Jackendoff, R. (2005). The faculty of language: what is special about it? *Cognition*, 95, s. 201–236.
- Pollick, A.S., de Waal, F.B.M. (2007). Ape gestures and language evolution. *Proceedings of the National Academy of Sciences of the USA*, 104, s. 8184–8189.
- Rizzolatti, G., Arbib, M.A. (1998). Language within our grasp. *Trends in Neurosciences*, 21, s. 188–194.
- Rizzolatti, G., Fadiga, L., Fogassi, L., Gallese, V. (1996). Premotor cortex and the recognition of motor actions. *Cognitive brain research*, 3, s. 131–141.
- Rowe, M.L., Goldin-Meadow, S. (2009). Early gesture selectively predicts later language learning. *Developmental Science*, 12, s. 182–187.
- Sandler, W. (2013). Vive la différence: Sign language and spoken language in language evolution. *Language and Cognition*, 5(2–3), s. 189–203.
- Stokoe, W.C. (1960). *Sign language structure*. Silver Spring: Linstok Press.
- Stokoe, W.C. (1991). Semantic phonology. *Sign language studies*, 71, s. 107–114.
- Stokoe, W.C. (2001). *Language in hand: Why sign came before speech*. Washington, DC: Gallaudet University Press.
- Stokoe, W.C., Casterline, D.C., Croneberg, C.G. (1965). *A dictionary of American Sign Language on linguistic principles*. Silver Spring: Linstok.
- Studdert-Kennedy, M. (2005). How did language go discrete?. W: M. Tallerman (red.), *Language origins – perspectives on evolution* (s. 48–67). Oxford: Oxford University Press.
- Slocombe, K. (2011). Have we underestimated great ape vocal capacities?. W: K. R. Gibson, M. Tallerman (red.) *The Oxford Handbook of Language Evolution* (s. 90–95). Oxford: Oxford University Press.
- Slocombe, K., Waller, B., Liebal, K. (2011). The language void: the need for multimodality in primate communication research. *Animal Behaviour*, 81(5), s. 919–924.
- Summerfield, A.Q. (1992). Lipreading and audio-visual speech perception. *Philosophical Transactions of the royal society B*, 335, s. 71–78.
- Sutton-Spence, R. (2005). *Analysing Sign Language Poetry*. New York: Palgrave Macmillan.
- Tallerman, M. (2011). Protolanguage. W: K.R. Gibson, M. Tallerman (red.) *The Oxford Handbook of Language Evolution* (s. 479–491). Oxford: Oxford University Press.
- Tallerman, M., Gibson, K.R. (red.). (2011). *The Oxford handbook of language evolution*. Oxford: Oxford University Press.
- Tinbergen, N. (1951). *The study of instinct*. New York: Oxford University Press.
- Tomasello, M. (2008). *Origins of human communication*. Cambridge, MA: MIT Press.
- Waciewicz, S. (2008). Ewolucja języka: horyzont metodologiczny. W: Stalmaszczyk, P. (red.) *Metodologie językoznawstwa. Współczesne tendencje i kontrowersje* (s. 27–42). Kraków: Wydawnictwo LEXIS, s. 27–42.

- Wacewicz, S. (2013). Ewolucja języka – współczesne kontrowersje. W: Stalmaszczyk, P. (red.) *Metodologie językoznawstwa. 1. Ewolucja języka. Ewolucja teorii językoznawczych* (s. 11–26). Łódź: Wydawnictwo Uniwersytetu Łódzkiego.
- Wacewicz, S., Żywicznyński, P. (2008). Broadcast Transmission, Signal Secrecy And Gestural Primacy Hypothesis. W: A. Smith, K. Smith, R. Ferrer-i-Cancho (red.), *The Evolution of Language. Proceedings of the 7th International Conference (EVO-LANG 7)* (s. 354–361). Singapore: World Scientific.
- Zlatev, J. (2008). The co-evolution of intersubjectivity and bodily mimesis. W: J. Zlatev, T. Racine, C. Sinha, E. Itkonen (red.), *The Shared Mind: Perspectives on intersubjectivity* (s. 215–244). Amsterdam: John Benjamins Publishing.
- Zlatev (w druku). Bodily mimesis and the transition to speech.
- Żywicznyński, P., Wacewicz, S. (2008). *Ewolucja języka. W stronę hipotez gesturalnych*. Toruń: Wydawnictwo Naukowego Uniwersytetu Mikołaja Kopernika.

*Ethnologue*: <http://www.ethnologue.com/>

*World Atlas of Linguistic Structures*: <http://wals.info/>