

Hipotezy powstania języka a problem prozodii¹

Language origin hypotheses and prosody

Żywiczyński, Przemysław. 2020. Hipotezy powstania języka a problem prozodii. W: Prozodia. *Przyswajania, badanie, zaburzenia, terapia*, red. Marta Wysocka, Barbara Kamińska, Stanisław Milewski, str. 38-64. Gdańsk: Wydawnictwo Uniwersytetu Gdańskiego. DRAFT

1.1. Wprowadzenie

W ramach nauki o ewolucji języka (ang. *the science of language evolution*; Żywiczyński, 2018, s. 11–18) nie podjęto jeszcze szeroko zakrojonych badań nad prozodią, chociaż refleksja nad zjawiskami prozodycznymi pojawia się w wielu wpływowych hipotezach powstania języka. Refleksja ta dotyczy przede wszystkim prozodii emocjonalnej, ale podejmowane są również próby wyjaśnienia pochodzenia prozodii jako części systemu gramatycznego języka. Celem tej pracy jest opisanie tych prób w odpowiednich kontekstach badawczych. Należy wyraźnie stwierdzić, że problematyka prozodii zajmuje miejsce marginalne w badaniach nad ewolucją języka. Z tego względu w niniejszym rozdziale pojawia się refleksja odnośnie do tego, jak podjęcie tematu prozodii może stać się istotnym motywem badawczym w nauce o ewolucji języka. Z tego powodu wzmianki na temat prozodii w hipotezach glottogenetycznych zostały ukazane na szerokim historycznym i teoretycznym tle, tak aby wskazać te miejsca, gdzie omawiana kwestia może być szczególnie interesująca w kontekście ogólnych badań nad pochodzeniem języka. Na końcu natomiast przedstawiono najbardziej obiecujące sposoby rozwinięcia tej problematyki badawczej w obrębie bądź w nawiązaniu do nauki o ewolucji języka.

¹ Publikacja powstała w ramach projektu sfinansowanego z grantu nr UMO-2017/27/B/HS2/00642.

1.2. Protojęzyk, hipotezy powstania języka i problem prozodii

Mimo, że nauka o ewolucji języka rozwija się od niedawna, wypracowała już pewien zasób pojęć metateoretycznych. Jednym z najważniejszych jest „**protojęzyk**” (zob. np. Smith, 2008, s. 99). Termin ten oznacza **system komunikacji prostszy od języka, ale posiadający niektóre cechy dla języka charakterystyczne, których nie posiadają systemy komunikacyjne zwierząt**, a szczególnie systemy, którymi posługują się małe człokształtne (zob. np. Bickerton, 1990, s. 122–125; Fitch, 2010, s. 399–400; Żywiczyński i Wacewicz, 2015, s. 184–188). Ten sens terminu należy oczywiście odróżnić od tego, jaki ma on w językoznawstwie historycznym, gdzie określa hipotetyczny system językowy, z którego wywodzą się spokrewnione ze sobą języki (zob. np. Bynon, 1977, s. 70–75). Oba znaczenia łączy więc hipotetyczny status protojęzyka, a dzieli – jego natura. W językoznawstwie historycznym protojęzyk odnosi się do w pełni ukształtowanego języka, a w ewolucji języka – do systemu, który posiada jedynie pewne cechy języka. Dochodząc do metateoretycznej funkcji, jaką to pojęcie pełni w tej drugiej dyscyplinie, w literaturze przeglądowej często wyzyskuje się szczegółowe poglądy badaczy na protojęzyk w celach klasyfikacyjnych. Jak zobaczymy, podstawowe osie debaty nad protojęzykiem przedstawione poniżej często służą określeniu poszczególnych scenariuszy powstania języka (stąd mowa np. o scenariuszach holistycznych, muzycznych czy gesturalnych).

Tabela 1.1. Główne osie debaty nad naturą protojęzyka*

Struktura		
kombinatoryczna (Bickerton, 1990; Jackendoff, 2002)		holistyczna (Wray, 1998; Mithen, 2005; Arbib, 2012)
Funkcja		
semantyczna		reprezentacyjna (Bickerton, 1990)
komunikacyjna (Wray, 1998; Jackendoff, 2002)	muzyczna (Mithen, 2005; Fitch, 2010)	

Modalność			
wokalna (Dunbar, 1996; Burling, 2005; MacNeilage, 2008)	gesturalna (Corballis, 2003; Armstrong i Wilcox, 2007)	multimodalna (Kendon, 2011; Arbib, 2012; McNeill, 2012; Sandler i in., 2014)	pantomimiczna (Zlatev i in., 2017)

* Należy zauważyć, że problem modalności protojęzyka powstaje tylko w przypadku założenia, iż pełnił on (też) funkcję komunikacyjną.

Źródło: Żywiczyński, Gontier i Wacewicz, 2017; zob. też: Johansson, 2005; Fitch, 2010.

Jak wynika z powyższej tabeli, teorie dotyczące protojęzyka charakteryzowane są w odniesieniu do jego **struktury, funkcji i modalności komunikacyjnej**. Jeśli chodzi o tę pierwszą oś debaty, to dominują dwa stanowiska. Zwolennicy pierwszego z nich, **kombinatorycznego** (zob. wyżej; także: Dennett, 2017), zakładają, że protojęzyk składał się z elementów quasi-leksykalnych, a jego przemiana we w pełni ukształtowany język polegała na wykształceniu składni. Z kolei zwolennicy podejścia **holistycznego** uważają, że protojęzyk tworzyły odpowiadające wypowiedziom holofrazy, a te następnie uległy segmentacji na jednostki leksykalne (zob. przede wszystkim: Wray, 1998; Arbib, 2012). Jeśli chodzi o funkcjonalność protojęzyka, większość badaczy przyjmuje, że służył on przede wszystkim szeroko pojętym celom **komunikacyjnym**. Mniejszość, choć dość wpływowa ze względu na odwołanie do poglądów Noama Chomsky'ego², dowodzi, że pierwotną funkcją języka było **reprezentowanie** oraz łączenie pojęć – w tym ujęciu protojęzyk był na początku narzędziem służącym do myślenia, a dopiero później uzyskał funkcję komunikacyjną (zob. przede wszystkim: Bickerton, 1990; 1998; 2003; 2009). Z kolei wśród badaczy związanych ze stanowiskami komunikacyjnymi dominuje pogląd, że jednostki protojęzyka wyrażały treści **semantyczne** – jako protoleksemę (w stanowiskach kombinatorycznych) bądź jako holofrazy od-noszące się do całych zdarzeń (w stanowiskach holistycznych). Inni badacze przyjmują, że protojęzyk nie służył (przynajmniej na początku) przekazowi treści

² Należy wyjaśnić, że choć Chomsky twierdzi, iż język pojawił się jako narzędzie służące do myślenia, a nie komunikacji, to według niego pojawił się on nagle. W tym kontekście Chomsky opowiada się przeciwko gradualistycznej, darwinowskiej wizji ewolucji języka, a za zmianą skokową, która miała doprowadzić do pojawienia się języka, bez pośredniego etapu protojęzykowego (np. Chomsky, 1995; Berwick i Chomsky, 2011).

propozycjonalnych, ale emocjonalnych, szczególnie za pomocą wokalizacji (zob. przede wszystkim: Mithen, 2005; Fitch, 2010). Takie stanowiska zazwyczaj określa się mianem **hipotez muzycznego protojęzyka**. Ostatnią osią debaty jest modalność protojęzyka. Wyznaczają ją dwie skrajne pozycje – stanowiska **pierwszeństwa mowy**, według których protojęzyk miał naturę wokalną (np. Dunbar, 1996; Bur-ling, 2005; MacNeilage, 2008), oraz stanowiska **pierwszeństwa gestów**, według których protojęzyk rozwinął się w kanale motoryczno-wizualnym, a następnie dokonało się jego przejście do kanału wokalno-słuchowego (np. Hewes, 1973; Armstrong i Wilcox, 2007)³. W dużym uproszczeniu: stanowiska **multimodalne** i **pantomimiczne** lokują się między tymi dwiema brzegowymi propozycjami (zob. Zlatev i in., 2017).

Refleksja nad prozodią występuje w kilku – często odmiennych – konfiguracjach stanowisk dotyczących protojęzyka, jednak **przede wszystkim problem prozodii jest zaznaczony w teoriach muzycznych**. Jak już wspomniano, w sensie ogólnym zakładają one, że w pełni ukształtowany język poprzedzał system komunikacji, który opierał się na wokalizacjach, a ich podstawową funkcją był przekaz treści niesemantycznych o charakterze emotywnym (np. Mithen, 2005; Zlatev i in., 2017). Oprócz stanowisk muzycznych w sensie ogólnym istnieją też stanowiska muzyczne w sensie silnym, które oprócz poglądu o niesemantycznej funkcji protojęzyka opierają się na założeniu, że miał on formę pieśni, to jest składał się z dyskretnych elementów dźwiękowych ułożonych w skalę, która była powtarzalna (Fitch, 2010, s. 474–481; zob. też: Nettle, 2000). Omówimy teraz te dwa rodzaje stanowisk muzycznych, wskazując miejsca, w których odwołują się one do problemu powstania prozodii.

³ Z jednej strony wielu zwolenników hipotez pierwszeństwa mowy przyznaje, że w ewolucyjnym rozwoju języka wokalizacjom z pewnością towarzyszyły gesty (np. MacNeilage, 2008, s. 283), a z drugiej wielu prominentnych przedstawicieli hipotez pierwszeństwa gestów to samo twierdzi o wokalizacjach (np. Armstrong i Wilcox, 2007, s. 68). Logika podziału na te dwa stanowiska dotyczy przede wszystkim identyfikacji tych zachowań, które były główną siłą powstania i rozwoju języka – zwolennicy hipotez pierwszeństwa mowy identyfikują tę siłę jako zachowania wokalne, a zwolennicy hipotez pierwszeństwa gestów jako zachowania gesturalne (zob. Żywiczyński, Gontier i Waciewicz, 2017).

1.3. Język naturalny – wokalizacje i gesty

Hipotezy określane jako muzyczne zarówno w sensie ogólnym, jak i silnym odwołują się do rozwiązań holistycznych (np. Mithen, 2005; Fitch, 2010; Zlatev i in., 2017). Jeśli chodzi o modalność protojęzyka, wydawać by się mogło, że stanowiska muzyczne będą zdominowane przez hipotezy wokalne, to jest takie, które zakładają, że język wyłaniał się przede wszystkim w kanale wokarno-słuchowym. Podejścia tego typu są obecne w literaturze, a wśród nich istnieje koncepcja prozodycznego protojęzyka przedstawiona przez W. Tecumseha Fitcha (2005, 2006, 2010). Jednak **wiele muzycznych scenariuszy występuje w koncepcjach multimodalnych i pantomimicznych**. Mowa tu zwłaszcza o tych scenariuszach, które odpowiadają definicji protojęzyka muzycznego w sensie ogólnym.

Takie stanowiska mają długą historię i odnajdujemy je już u oświeceniowych autorów, którzy twierdzili, że język rozwinął się z bimodalnej komunikacji opartej na pantomimie oraz wokalizacjach (zob. Żywiczyński, 2018, s. 100–108). Na ówczesnych myślicieli największy wpływ wywarł scenariusz przedstawiony najpierw przez Bernarda de Mandeville’a w drugim wydaniu poematu ekonomiczno-społecznego *Bajka o pszczołach* (*The fable of the bees*, 1729), a później w poczytnym traktacie filozoficznym *O pochodzeniu poznania ludzkiego* (*Essai sur l’origine des connaissances*, 1756), którego autorem była jedna z najbardziej wpływowych postaci epoki – biskup Étienne Bonnot de Condillac. Zarówno u Mandeville’a, jak i Condillaca rozważania na temat początków języka mają formę eksperymentu myślowego, w którym para dzieci niewładających jeszcze żadnym językiem zostaje na trwałe oddzielona od reszty ludzkości. Według tych myślicieli oraz wielu ich zwolenników z okresu oświecenia (zob. Żywiczyński, 2018, s. 119–122) pierwsze próby komunikacji między dziećmi dotyczyły ich potrzeb (np. głodu) i miały formę okrzyków oraz pantomimy: „ten, który cierpiał z powodu utraty obiektu istotnego dla swoich pragnień, nie ograniczał się tylko do krzyków czy wydawania dźwięków, ale angażował całe ciało w zasygnalizowanie cierpienia – poruszał głową, ramionami i innymi częściami ciała” (de Condillac, 1756, s. 169–176). Szczególnie ze scenariusza przedstawionego przez Condillaca wynikał „podział pracy” między komunikacją zachodzącą w dwóch

modalnościach. **Okrzyki miały przede wszystkim wyrażać walencję emocjonalną** (szczęście, cierpienie, satysfakcję, rozterkę), a **pantomima i gesty – treści semantyczne**, które ikonicznie albo indeksykalnie⁴ odnosiły się do zewnętrznych obiektów (de Condillac, 1756, s. 169). Należy zauważyć, że postulat Condillaca o bimodalnej naturze pierwotnego systemu komunikacji i „podziale pracy” między modalnościami skutkuje połączeniem stanowiska muzycznego z semantycznym. Dla dwojga „dzikich dzieci” **emocjonalne wokalizacje stanowiły protojęzyk muzyczny, a gesty i pantomima – protojęzyk semantyczny**.

Dalej Condillac twierdzi, że dzięki powtarzającym się próbom komunikacji między dziećmi wokalizacje stopniowo łączyły się z pojęciami pierwotnie wyrażanymi przez gesty (de Condillac, 1756, s. 170–171), a mówiąc językiem semiotyki: wokalizacje nie były już indeksami stanów emocjonalnych, ale uzyskiwały „komunikacyjną funkcję znaku” (ang. *communicative sign function*), dzięki której nadawca używa znaku z intencją, że znak odsyła do czegoś innego niż on sam, a odbiorca rozumie tę intencję (Zlatev, Persson i Gärdenfors, 2005). W kolejnych pokoleniach dzieci pierwszej pary dysponowały rosnącymi zdolnościami artykulacyjnymi i język mógł w coraz większym stopniu opierać się na symbolach wokalnych, a w coraz mniejszym – na gestach i pantomimie (de Condillac, 1756, s. 171–172). Jednak Condillac zauważa, że nawet w zachowaniu współczesnych ludzi występują pozo-stałości, które świadczą o bimodalnych korzeniach komunikacji. Z jednej strony są nimi „taniec gestów”, który towarzyszy mowie, wzmacniając niesione przez nią treści, oraz „taniec kroków” – czyli taniec w sensie potocznym, który według Condillaca służy osiągnięciu przyjemności (de Condillac, 1756, s. 116). Z drugiej strony **ludzie cały czas potrafią wyzyskiwać pierwotny charakter mowy, wyrażając emocje za pomocą odpowiedniej modulacji głosu**. Niestety Condillac nie rozwija tej obserwacji, którą można uznać za pierwszy naturalistyczny scenariusz powstania cech prozodycznych języka, w tym wypadku ograniczonych do prozodii emocjonalnej (ang. *emotional prosody*; zob. Bachorowski, 1999). W kontekście omawianych w niniejszym rozdziale zagadnień byłoby to interesujące.

⁴ W języku współczesnej semiotyki znaki tego typu należałoby nazwać umotywowanymi (ang. *motivated signs*), tj. takimi, w których forma jest uzasadniona przez treść, w odróżnieniu od znaków niemotywowanych, czyli arbitralnych (np. Sonesson, 2007; Zlatev, 2008).

Jak już wspomniano, propozycja Mandeville'a i Condillaca stała się dominującą hipotezą powstania języka w oświeceniu, a siła jej oddziaływania wyszła daleko poza tę epokę. Wokalno-pantomimiczny protojęzyk zaczęto traktować jako rodzaj komunikacji przyrodzony ludziom, to jest taki, którym ludzie są w stanie skutecznie komunikować się, nawet jeśli brakuje im socjalizacji (zob. np. pracę Jeana Itarda z 1801 r. o wilczym dziecku, które francuski badacz rehabilitował) albo wspólnego języka (przeгляд relevantnych źródeł przedstawia Hewes, 1973). Pojawiły się nawet postulaty, żeby zbudować uniwersalny system komunikacji oparty na pantomimie i emocjonalnych wokalizacjach (pomysły te były szczególnie popularne wśród ideologów – grupy myślicieli i pedagogów, którzy na początku XIX wieku działali w Paryżu; zob. np. Laromiguière, 1826). W założeniu naturalności tkwi właśnie siła propozycji Mandeville'a i Condillaca, którzy w swoim eksperymencie myślowym uchwycili w przekonujący sposób to, jak ludzie mogliby się komunikować, gdyby nie udało się im nabyć mowy. Ich intuicja została, przynajmniej w jakimś stopniu, uprawdopodobniona przez badania empiryczne. I tak twórca antropologii kulturowej Edward Tylor na podstawie analizy języków migowych, sposobów komunikacji używanych przez „dzikie” (ang. *savage*) ludy oraz przypadków re-habilitacji dzieci wilczych doszedł do wniosku, że **sposób porozumiewania się za pomocą ikonicznych gestykulacji i emotywnych wokalizacji jest u ludzi wrodzony** (Tylor, 1870, s. 23; 1871, s. 145; zob. też: Mocerino, 2016). Pisząc o języku naturalnym (ang. *natural language*), podobnie jak Condillac, uważał on, że **emocjonalna prozodia** może pochodzić z tego wrodzonego systemu komunikacji. Już w czasach najnowszych Gordon W. Hewes, również antropolog, prekursor nauki o ewolucji języka i badań nad gestami, przenalizował przekazy europejskich podróżników z epoki wielkich odkryć geograficznych i doszedł do wniosku, że właśnie za pomocą niejęzykowych wokalizacji, gestów i pantomimy byli oni w stanie uzyskać od tubylców informacje dotyczące nawet stosunkowo abstrakcyjnych zagadnień, na przykład topografii, niebezpieczeństw czyhających na przybyszów, a nawet systemów politycznych plemion (Hewes, 1973; zob. też: Żywiczyński i Waciewicz, 2015, s. 139).

1.4. Współczesne koncepcje pantomimiczne i multimodalne

Jedną z podstawowych słabości wymienionych powyżej propozycji jest założenie, że zdolności komunikacyjne współczesnych ludzi nie różnią się zasadniczo od tych, które posiadali nasi przodkowie uczestniczący w procesie tworzenia się języka. Nauka o ewolucji języka stara się radzić sobie z tym problemem, odwołując się do metodologicznego postulatu zbieżnych świadectw (ang. *converging evidence*; Żywiczyński i Wacewicz, 2015, s. 158–159), który uzależnia ocenę danego scenariusza powstania języka od zgromadzenia wspierających się nawzajem danych z relewantnych obszarów badań. Stanowiskiem najbardziej zbliżonym do koncepcji języka naturalnego, z którego miał wyłonić się w pełni ukształtowany język, są współczesne koncepcje pantomimiczne (zob. przede wszystkim Zlatev i in., 2017; Gärdenfors, 2017). Postulat, że pantomima stanowiła istotny przejściowy etap w ewolucyjnym kształtowaniu się języka, występuje u bardzo wpływowych badaczy zajmujących się ewolucją języka, takich jak Michael Arbib (np. 2012), Michael Tomasello (np. 2008) oraz Merlin Donald (1991; 2001), Jordan Zlatev (np. 2008) czy Peter Gärdenfors (np. 2017), z których ostatni odwołują się do teorii mimezy. W rysowanych przez nich scenariuszach atrakcyjność pantomimy jako systemu komunikacji protojęzykowej polega na tym, że pozwala ona na przekaz semantycznie otwartego zasobu znaczeń (ang. *open-endedness*), które mogą być oderwane od „tu i teraz” (ang. *displacement*; zob. kryterialne cechy języka – Hockett, 1960) bez wymogu istnienia konwencji komunikacyjnych (Żywiczyński, Wacewicz i Sibierska, 2018).

Wspólną cechą koncepcji pantomimicznych jest założenie, że **w komunikacji protojęzykowej znaczenia miały charakter umotywowany**, a ze względu na **większy potencjał ikoniczny kanału motoryczno-wizualnego** to on był przede wszystkim odpowiedzialny za przekaz treści semantycznych (Zlatev i in., 2017; Żywiczyński, Wacewicz i Sibierska, 2018). W przeciwieństwie do hipotez pierwszeństwa gestów koncepcje pantomimiczne dopuszczają, że pantomimę wspierały inne zasoby semiotyczne, głównie **niejęzykowe wokalizacje** (Żywiczyński, Wacewicz i Sibierska, 2018), choć odgrywały one drugorzędną rolę i zazwyczaj **nie służyły przekazowi treści semantycznych**. W tym zakresie, podobnie jak dawne koncepcje

języka naturalnego, współczesne hipotezy pantomimiczne postulują holistyczny protojęzyk, który częściowo ma charakter semantyczny – za sprawą modalności wizualnej, a częściowo muzyczny – przede wszystkim za sprawą wokalizacji. Idąc dalej, tak jak Condillac czy Tylor, zwolennicy pantomimicznego protojęzyka dostrzegają, że **proces przejścia tego systemu do kanału wokalnego, który charakteryzuje współczesną komunikację językową, stwarza pole do namysłu nad powstaniem prozodii i jej rolę w ewolucji języka**. Należy też zauważyć, że założenie istnienia „podziału pracy” w protojęzyku między komunikację zachodzącą w kanałach motoryczno-wizualnym i wokalno-słuchowym (de Condillac, 1756) odróżnia koncepcje pantomimiczne od koncepcji multimodalnych (albo gestowo-głosowych), które – opierając się na ścisłych związkach między mową i gestykulacją we współczesnej komunikacji – kładą nacisk na założenie, że język od początku swojej ewolucyjnej historii rozwijał się przy równym współdziałaniu wokalizacji i gestów (np. Kendon, 1991; Goldin-Meadow, 2011; McNeill, 2012; zob. też: Żywiczyński i Wacewicz, 2015, s. 236–238).

Sytuując się w taki sposób na osiach debaty dotyczącej protojęzyka, stanowiska pantomimiczne wydają się z powodzeniem wykorzystywać podstawowe argumenty na rzecz hipotez pierwszeństwa gestów, z których najważniejsze to (za: Żywiczyński i Wacewicz, 2015, s. 246–255):

- Status komunikacji gestowej u małych człekokształtnych. Ta linia argumentacji dotyczy zarówno **udanych prób uczenia małych człekokształtnych systemów komunikacji motoryczno-wizualnej** (np. Gardner i Gardner, 1969; Patterson, 1981; zob. też: Hewes, 1973; Corballis, 2002; Tomasello, 2008) w zestawieniu z całkowicie nieudanymi próbami uczenia ich mowy (np. Kellogg i Kellogg, 1933; Hayes i Hayes, 1952), jak i **komunikacji gesturalnej u wolno żyjących małych człekokształtnych**, w szczególności szympanów i goryli, którą charakteryzuje wolicjonalność, plastyczność i otwartość na innowacje (np. Hewes, 1973; Pika, 2008) w porównaniu z ich komunikacją wokalną, która – ze względu na niższą korykalną kontrolę tych zachowań – ma charakter bardziej instynktowny (np. Ghazanfar i Rendall, 2008).
- Ikoniczność gestów. Argumenty te podkreślają **większy potencjał ikoniczny gestów niż wokalizacji**, co miało być użyteczne w początkowych etapach

rozwoju języka. Wiele argumentów odwołujących się do ikoniczności wywodzi się z badań nad językami migowymi (Stokoe, 1991; Armstrong i Wilcox, 2007), również nowo powstałymi, takimi jak nikaraguański język migowy (Kegl, Senghas i Coppola, 1999) czy język migowy Al-Sayyid Bedouin (Sandler i in., 2005), a ostatnio również z badań empirycznych nad skutecznością improwizowanej komunikacji gesturalnej i wokalne (Fay, Arbib i Garrod, 2013; Fay i in., 2014; Motamedi-Mousavi i in., 2017).

- Argumenty neuropoznawcze. Na przykład badania nad ośrodkiem Broki wskazują, że nie uczestniczy on jedynie w generowaniu mowy, ale **odpowiada za ogólną zdolność do rozpoznawania hierarchicznych zależności**, bez względu na to, czy dotyczą one mowy, muzyki czy też czynności manualnych (Fadiga, Craighero i D'Ausilio, 2009).

W odróżnieniu od hipotez pierwszeństwa gestów koncepcje pantomimiczne nie zmagają się z tak wyraźnym **problemem zmiany modalności** z gesturalnej na wokalną, dlatego że wokalizacje jako protojęzyk muzyczny wbudowane były już w wyłaniający się system komunikacji językowej. Koncepcje pantomimiczne w sposób przekonujący muszą wyjaśnić, **jak ten protojęzyk muzyczny zmienił się w dominującą modalność językową**. Jak zobaczymy, odnoszą się one do tego problemu w różny sposób, co czasami sprzyja podjęciu zagadnień związanych z prozodią (prezentacja stanowisk pantomimicznych na podstawie: Wacewicz i Żywiczyński, 2018).

I tak Arbib umieszcza pantomimę w swojej propozycji zwanej hipotezą lu-strzanych neuronów (ang. *mirror neuron hypothesis*; Arbib, 2005; 2012), która jest zdecydowanie najbardziej kompletnym scenariuszem powstania języka we współczesnych badaniach. Jak wskazuje jej nazwa, punktem wyjścia jest tutaj system neuronów lustrzanych (Rizzolatti i Arbib, 1998) – grupa komórek nerwowych aktywnych zarówno podczas wykonywania czynności, jak i podczas jej obserwacji. Istnienie takiej grupy stwierdzono w linii naczelnych (zob. też: Żywiczyński i Wacewicz, 2015, s. 252–254). Arbib argumentuje (2005; 2012), że system ten stanowi istotne przystosowanie do przekazu informacji, gwarantując **równoważność** (ang. *parity*) i **przemienność** (ang. *interchangeability*) **ról nadawcy i odbiorcy**. Dzięki

temu u szympanсів i, jak twierdzi Arbib, u wspólnego przodka szympanсів (łac. *Pan*) i ludzi (ang. *last common ancestor with chimpanzee*, LCA-c) rozwinął się system prostej imitacji umożliwiający powtarzanie czynności, a na jego podstawie – już tylko w linii człowiekowatych – system złożonej imitacji (ang. *complex action recognition and complex imitation*, CAR&IM), który umożliwia **dzielenie czynności na podczynności i rekombinację tych podczynności w nowe sekwencje** (Arbib, 2005; 2012). System złożonej imitacji z kolei promował rozwój zdolności symbolicznych i w ten sposób torował drogę **pantomimie**, którą Arbib definiuje jako **system komunikacji posiadający otwartą semantykę**, to jest taką, która pozwala na ustanawianie nowych związków między elementami CAR&IM (Arbib, 2012, s. 261).

Słabość pantomimy jako systemu komunikacji wynika z braku norm komunikacyjnych, co powoduje, że jej zrozumienie i replikacja łączą się ze znacznymi kosztami czasowymi i kognitywnymi. Według Arbiba ten stan rzeczy prowadził do **stopniowej konwencjonalizacji komunikatów pantomimicznych**, czyli **utruty przez nie właściwości ikonicznych na rzecz cech arbitralnych** (Arbib, 2012, s. 226). Ten element koncepcji Arbiba uwiarygodniają badania z zakresu iterowanego uczenia się, które ukazują, jak **na drodze ewolucji kulturowej szybko zachodzi konwencjonalizacja komunikatów ikonicznych** (np. Schouwstra i de Swart, 2014; Motamedi-Mousavi i in., 2017). W końcu dochodzimy do problemu przejścia protojęzykowej komunikacji do kanału głosowego. Powodem tej zmiany było **stopniowe zapelnianie się przestrzeni semantycznej** (ang. *crowding of the semantic space*), czyli wzrost ogólnej liczby sygnałów językowych, które odnosiły się do coraz bardziej podobnych do siebie znaczeń. Według dobrze eksperymentalnie zweryfikowanej hipotezy (Tomasello, 2008; Fay, Arbib i Garrod, 2013; Fay i in., 2014) w takich warunkach sygnały głosowe są bardziej użyteczne niż, obciążone ikonicznością, sygnały gesturalne. Według Arbiba ten powód wystarczył, aby wyjaśnić przejście wyłaniającego się systemu językowego do kanału wokalnospuchowego.

Arbib wskazuje na mechanizm neuronalny, który jego zdaniem jest odpowiedzialny za zmianę modalności (opis tego mechanizmu na podstawie: Fitch, 2010, s. 456–464). Jego uwaga skupia się na polu 5 Brodmanna, gdzie ulokowana jest

drugorzędna kora somatosensoryczna odpowiedzialna za kojarzenie bodźców wzrokowych z czynnościami ruchowymi i gdzie po raz pierwszy stwierdzono istnienie neuronów lustrzanych (Arbib, 2005). Postuluje on, że **pantomimie i gestom zawsze towarzyszyły wokalizacje**, co – jak już się przekonaliśmy – stanowi jedno z fundamentalnych założeń hipotez pierwszeństwa pantomimy. Następnie zaś Arbib stwierdza, że taki stan rzeczy promował **rozszerzenie pola działania** (kolateralizację) **neuronów lustrzanych na przyległe rejony oromotoryczne**, które u ludzi odpowiedzialne są za kontrolę wokalizacji (zob. też punkt 5). To właśnie ta zmiana spowodowała **wzrost umiejętności artykulacyjnych oraz fonacyjnych** u naszych przodków i w konsekwencji umożliwiła przejście komunikacji językowej do kanału wokalnie-słuchowego. Stanowisko to łączy się z refleksją nad cechami prozodycznymi języka. Tak jak w innych scenariuszach pantomimicznych **prozodia emocjonalna** jest pozostałością pantomimicznego protojęzyka i **wyzyskuje stare ścieżki neuronalne odpowiedzialne za zachowania wokalne**. Natomiast **prozodia jako część systemu języka korzysta już z infrastruktury kortykanej, która wyłoniła się za sprawą ekspansji systemu neuronów lustrzanych**.

Stanowisko Tomasello dotyczące pantomimy (2008; 2009), bardzo wpływowego badacza w obszarze nauki o ewolucji języka, nie dotyczy spraw tutaj omawianych. W dużym skrócie: Tomasello skupia się na warunkach wstępnych, które umożliwiły pojawienie się komunikacji pantomimicznej u człowiekowatych. Najważniejsze z nich mają charakter społeczno-kognitywny, jak na przykład „wspólna płaszczyzna uwagowa” (ang. *common ground*; Tomasello, 2008, s. 73–88), która pozwala jednostkom współpracować podczas wykonywania wspólnych zadań (np. zabawa, w której uczestnicy odgrywają odpowiednie role, a później tymi rolami się zamieniają), albo typowo ludzka dyspozycja do dzielenia się uczciwymi informacjami z niespokrewnionymi osobnikami (Tomasello, 2008, s. 85–97, zob. też: Wacewicz i Żywiczyński, 2018).

Wątki związane z tymi problemami znajdują się też w tych koncepcjach pantomimicznych, które odwołują się do **teorii mimezy cieleśnej** (ang. *bodily mimesis*). Jest to propozycja sformułowana przez Donalda, który stwierdza, że kluczową adaptacją w linii człowiekowatych było pozyskanie zdolności mimetycznych,

czyli zdolności do **wykorzystywania własnego ciała tak, aby w zamierzony i kontrolowany sposób wyrażać treści symboliczne** (Donald, 1991), na przykład aby za pomocą pantomimy „opisywać” przeszłe zdarzenia (por. Żywiczyński, Wacewicz i Sibierska, 2018). Zdolność ta oddzieliła przodków człowieka (Donald łączy pojawienie się zdolności mimetycznych z *Homo erectus*; zob. Donald, 1991) od kultury epizodycznej (ang. *episodic culture*) małp człekokształtnych, która w znacznym stopniu ogranicza poznanie oraz komunikację do „tu i teraz”. Według Donalda i jego zwolenników mimeza umożliwiła pojawienie się szeregu typowo ludzkich umiejętności. Chodzi tu na przykład o **umiejętność świadomego rekonstruowania i planowania działań** (co jest szczególnie ważne w przypadku produkcji narzędzi), **wzrost umiejętności demonstracyjnych**, który ukształtował drogę pedagogice (Gärdenfors i Högberg, 2017), a – na planie komunikacyjnym – **rozwój zdolności symbolicznych** (powstanie tzw. mimezy triadycznej opartej na trójkącie semiotycznym nadawca–referent–odbiorca; Zlatev, 2008; por. Peirce, 1868) oraz **powstanie otwartości semantycznej i oderwania od „tu i teraz” w opisywaniu zdarzeń** (por. Hockett, 1960).

Ważnym postulatem stanowisk mimetycznych jest integracja ekstrocepcji (percepcji świata zewnętrznego) i propriocepcji (percepcji własnego ciała; Zlatev, 2008). Integracja ta pozwala na jednoczesne wyzyskiwanie różnych modalności do konstrukcji komunikatów, na przykład użycia pantomimy i emocjonalnej wokalizacji do skonstruowania komunikatu dotyczącego jakiegoś zdarzenia (Zlatev i in., 2017). Prowadzi to do spekulacji, w jaki sposób mogła rozwijać się komunikacja w kanale słuchowo-wokalnym. Znany fakt z badań porównawczych jest niska kontrola małp człekokształtnych nad wokalizacjami⁶. Ma to związek z uzyskaniem w toku ewolucji człowiekowatych bezpośrednich połączeń nerwowych (z bocznej powierzchni kory nowej) z mięśniami odpowiedzialnymi za wokalizacje, które nie występują u innych ssaków, łącznie z innymi naczelnymi (Fitch, 2010, s. 349;

⁵ Według badań etologicznych komunikacja małp człekokształtnych ma przede wszystkim charakter diadyczny (czego przykładem są gesty zapraszające do iskania; np. Pika i in., 2005).

⁶ Istnieją również badania, które wskazują, że małpy posiadają pewną kontrolę nad wokalizacjami.

Deacon, 1992; Jürgens, 1994)⁷. Dzięki zwiększeniu intencjonalnej kontroli nad wokalizacjami można było zacząć wykorzystywać je do przenoszenia treści symbolicznych, na przykład **zmiana głośności mogła wyrażać zmianę rozmiaru, a iteracje dźwięków – powtarzalne działanie**. Niektórzy badacze wskazują, że **cechy prozodyczne, takie jak głośność, iloczys czy intonacja, mogły być wykorzystywane do przenoszenia informacji symboliczno-gramatycznych, kładąc w ten sposób podwaliny współczesnej prozodii** (zob. np. Zlatev, 2014).

Podobny do pantomimicznych scenariusz ewolucji zdolności komunikacyjnych znajdujemy w hipotezach **multimodalnych**. Kładą one nacisk na integrację modalności – szczególnie gesturalnej i głosowej – w ewolucji komunikacji językowej u człowiekowatych. Ich punktem wyjścia jest **multimodalny charakter najbardziej typowego użycia języka, to jest rozmowy twarzą w twarz** (ang. *face-to-face conversation*; np. Goffman, 1967; Levinson, 2006; Kendon, 2004). Na przykład Gabriella Vigliocco, Pamela Perniss i David Vinson piszą (2014, s. 2): „Mowa i gest są nieodłączne i stanowią ściśle zintegrowaną jednostkę – ten fakt uzasadnia potrzebę rozwijania multimodalnego podejścia (do problemu komunikacji językowej)”⁸. Zdaniem wpływowych badaczy gestów, takich jak Adam Kendon czy David McNeill, **mowa i gesty stanowią dwa równoważne zasoby semiotyczne**, z których wspólnego użycia rodzi się komunikacja językowa (albo „językowanie”, ang. *linguaging* – pojęcie, którego używa Kendon, żeby podkreślić dynamiczny, *in statu nascendi* charakter komunikacji językowej; zob. Kendon, 2004). Według nich **ten rodzaj integracji semiotycznej nastąpił jedynie u człowieka i jego przodków** (np. McNeill, 2012). Na przykład Kendon łączy powstanie języka z wyłonieniem się „duetu mowno-kinetycznego” (ang. *speech-kinesic ensemble*; Kendon, 2014). Za ten proces miała odpowiadać rosnąca koordynacja między rękoma a ustami, która wiązała się z wykonaniem coraz

⁷ Inne gatunki ssaków mają jedynie połączenie pośrednie między tymi strukturami, które przechodzi przez pień mózgu (Kuypers, 1958; Fitch, 2010, s. 351).

⁸ Oryginalny tekst: *Speech and gesture are part and parcel of the same system and constitute a tightly integrated unit, thus underscoring the need for a multimodal approach.*

bardziej złożonych działań praktycznych, szczególnie dotyczących wytwarzania i używania narzędzi⁹.

W odróżnieniu od scenariuszy pantomimicznych, które podkreślają semiotyczną odrębność komunikacji zachodzącej w kanale motoryczno-wizualnym i wokально-słuchowym, **scenariusze multimodalne wskazują, że „podział pracy” między modalnościami w czasie ewolucji języka rozkładał się równomiernie między zachowania gesturalne i wokalne.** I tak McNeill twierdzi, że kluczowym momentem w powstaniu języka było zorganizowanie procesów myślowych i komunikacyjnych w tak zwane punkty wzmocnienia¹⁰ (ang. *growth points*; McNeill, 1992) składające się zarówno z gestów, jak i z wokalizacji, z których te pierwsze służyły przekazowi treści wyobraźniowych, a drugie – propozycjonalnych, podobnie jak to ma miejsce we współczesnej komunikacji (McNeill, 1992; 2005). Jeśli chodzi o prozodię, to ze scenariuszy multimodalnych – a w szczególności z tego, który rysuje McNeill – wynika, że wokalizacje (proto)językowe już na wczesnym etapie integracji oddzieliły się od wokalizacji emocjonalnych. Podobnie należy założyć, że **prozodyczne cechy języka, w odróżnieniu od prozodii emocjonalnej, zaczęły kształtować się już na początkowych etapach jego ewolucji.**

1.5. Prozodia – pozostałość muzycznego protojęzyka

Przejdziemy teraz do omówienia **scenariuszy muzycznych w sensie silnym**, czyli takich, które zakładają, że **na wczesnych etapach rozwoju protojęzyka miał on formę pieśni.** To stanowisko, podobnie jak scenariusze pantomimiczne i multimodalne (czyli teorie muzyczne w sensie słabym; zob. wyżej), ma długą i bogatą historię. Motywy podkreślające rolę pieśni w powstaniu języka pojawiają się na przykład u Jeana-Jacques’a Rousseau, który – choć akceptował pogląd Condillaca o bimodalnym charakterze pierwszych form protojęzykowej komunikacji (Rousseau, 1781/1998, s. 305; zob. Żywiecziński, 2018, s. 108–112) – twierdził,

⁹ Usta i zęby często były wykorzystywane do produkcji oraz używania narzędzi (np. Clement, Hillson i Aiello, 2012).

¹⁰ Pojęcie punktu wzmocnienia zapożyczone od Lwa Wygotskiego (Vygotsky, 1934/1964) służy McNeillowi do określenia najważniejszych treściowo momentów wypowiedzi (1992).

że na kolejnym etapie język rozwijał się przede wszystkim w kanale wokalnospłuchowym, ponieważ ten lepiej służył koordynacji działań rosnących liczebnie grup (1781/1998, s. 305–317). Według Rousseau **pierwsze języki wokalne były językami tonalnymi**. Znakomicie nadawały się do przekazywania emocji. Cały czas można je odnaleźć w znajdujących się na wczesnych etapach rozwoju cywilizacji społecznościach żyjących w ciepłych regionach świata (fr. *sociétés naissantes*). Według francuskiego filozofa późniejsza migracja na północ wymusiła powstanie jeszcze liczniejszych społeczności, które zwiększały szanse w walce

o przetrwanie we wrogim środowisku (1781/1998, s. 305–317). Języki kształtujące się w takich warunkach zaczęły być wykorzystywane przede wszystkim do przekazywania idei, tracąc swoją pierwotną melodyjność i tonalny charakter. Według tego scenariusza **prozodia – językowe użycie tonu i melodii – miała być pozostałością po muzycznych językach *sociétés naissantes***. Tego wątku Rousseau nie rozwija, ale – jak zobaczymy – jest on typowy dla muzycznych (w sensie silnym) teorii języka.

Podobne pomysły do tego, co zaproponował Rousseau, odnajdujemy w XIX wieku, na przykład u Ludwiga Noirégo albo Friedricha Maxa Müllera (zob. Żywi-czyński, 2018, s. 143–144, 159). Jednak w owym czasie najbardziej przełomową muzyczną koncepcję powstania języka zaprezentował sam Karol Darwin. Aby nie drażnić przedstawicieli kręgów konserwatywnych, w swoim pierwszym dziele *O powstawaniu gatunków* (*On the origin of species*, 1859) nie wspominał o ewolucji człowieka. Uwagi na ten temat przedstawił w pracy *O pochodzeniu człowieka* (*Descent of man*, 1871) i częściowo w późniejszej książce *O wyrażaniu emocji u człowieka i zwierząt* (*The expression of the emotions in man and animals*, 1872). W dziele *O pochodzeniu człowieka* Darwin rozwinął słynny scenariusz powstania języka w oparciu o **mechanizm doboru seksualnego**, który opisuje sukces reprodukcyjny w odniesieniu do problemu doboru partnera, przede wszystkim przez rywalizację między samcami i wybory samców dokonywane przez samice (Darwin, 1871, s. 248–256)¹¹. Według Darwina **w fazie początkowej protojęzyk spełniał**

¹¹ Darwin pojęciowo odróżniał dobór naturalny od doboru płciowego, podczas gdy współczesny ewolucjonizm traktuje dobór płciowy jako podtyp doboru naturalnego (szczegółowe wyjaśnienie w: Żywczyński i Waciewicz, 2015, s. 90–95).

funkcję podobną do pieśni ptaków – służąc zarówno płciowej rywalizacji, jak i podkreśleniu więzów oraz umacnianiu relacji między płciami. Na tym etapie dobór preferował osobniki, które miały lepsze umiejętności muzyczno-artykulacyjne (1871, s. 56–57). Rozwój tego systemu komunikacji był skorelowany z rozwojem mózgu, co miało umożliwić naszym przodkom **budowanie długich ciągów myślowych, które stopniowo łączyły się z pieśniami w umysłach nadawców i odbiorców, tworząc w ten sposób pierwsze symboliczne jednostki języka** (1871, s. 57). Powołując się na ustalenia współczesnego mu językoznawstwa historycznego, Darwin zakładał, że taki język mógł rozwinąć się we w pełni ukształtowane języki, którymi dzisiaj mówimy. Podobnie jak u Rousseau **prozodia miała być pozostałością pierwotnego muzycznego protojęzyka** i została zaadaptowana na potrzeby nowopowstałego systemu komunikacyjnego.

Inną ważną postacią dla koncepcji muzycznych był Otto Jespersen, który poświęcił ostatni, obszerny rozdział swojego dzieła *Język: jego natura, rozwój i pochodzenie* (*Language: Its nature, development, and origin*, 1922, rozdział XXI) problemowi powstania języka. Jespersen, jak sam podkreśla, używa metody retrospektywnej – analizę **zaczyna od współczesnych języków i stara się zrekonstruować trajektorię zmian językowych aż do początkowych form komunikacji językowej**. Celem tego badacza nie jest jednak rekonstrukcja pierwszego języka¹², ale sformułowanie ogólnych praw, które rządzą dynamiką ewolucji języka (1922, s. 418). Najważniejszym z nich jest zasada zmniejszania wysiłku mięsniowo-węgo (ang. *the principle to lessen the muscular effort*; Jespersen, 1922, s. 418). Stąd Jespersen wyprowadza hipotezę, że **pierwsze języki były artykulacyjnie trudniejsze niż języki współczesne** (np. wykorzystywały strumień ingresywny płucny). Posiadały też rozbudowane wzorce tonalne i intonacyjne, co upodabniało je do pieśni. Jespersen pisze (1922, s. 420): „był czas, kiedy mowa była pieśnią albo raczej że mowa i pieśń były od siebie nierozdzielne”¹³. Podobnie zatem jak Darwin, Jespersen opowiada się za hipotezą muzycznego protojęzyka w sensie

¹² Niektórzy XIX-wieczni filologowie porównawczy (np. Alexander Murray) starali się zrekonstruować pierwszy język ludzkości za pomocą metody porównawczej (Żywiczyński i Wacewicz, 2015, s. 67). Krytykę takiego podejścia odnajdujemy u Williama Whitneya (1872, s. 279).

¹³ Oryginalny tekst: *there once was a time when all speech was song, or rather when there once was a time when all speech was song, or rather when these two actions were not yet differentiated.*

silnym. Również w podobnym tonie wypowiada się o funkcji tego **protojęzyka, który nie miał służyć do przekazu treści propozycjonalnych** (tj. do „transferu myśli”; 1922, s. 433, 437), **ale do wokalne** *rozrywki* (ang. *vocal play*) **i ekspresji emocji**, na przykład miłości (1922, s. 432–433).

Podczas gdy Darwin skupia się na mechanizmie powstania takiego protojęzyka, Jespersen stara się przede wszystkim wyjaśnić procesy, które zmieniły muzyczny protojęzyk we współczesne formy komunikacji językowej. Podkreśla on holistyczny charakter protojęzyka, którego **morfologia miała przypominać języki syntetyczne**, z zastrzeżeniem że (proto)morfemy protojęzyka miały:

- kumulować znacznie więcej funkcji gramatycznych niż ma to miejsce we współczesnych językach (Jespersen, 1922, s. 422);
- robić to w bardzo nieregularny sposób;
- nie podlegać jednoznacznej segmentacji (Jespersen, 1922, s. 425).

Jespersen podkreśla, że **trajektorią rozwojową języka było przechodzenie od systemu opartego na jednostkach słabo analizowanych i stosunkowo długich do systemu opartego na jednostkach derywowanych za pomocą reguł z łatwo analizowanych, krótkich elementów** (1922, s. 429). Jeśli chodzi o to, jak muzyczny, emotywny protojęzyk stopniowo pozyskiwał funkcję referencyjną, początkowo jego jednostki odnosiły się do jednostkowych zdarzeń, później do typów zdarzeń, a w końcu wyłaniające się z tych jednostek protomorfemy zaczynały tworzyć jednostki leksykalne (1922, s. 440). Wśród procesów, które odpowiadały za te zmiany Jespersen wymienia metaforyzację oraz **sekrecję** (ang. *secretion*) – jego własne pojęcie stworzone na oznaczenie takiego rodzaju gramatykalizacji, w której tylko część słowa uzyskuje funkcję gramatyczną (1922, s. 384–386)¹⁴.

Badacz nie wypowiada się wprost o ewolucji prozodii, ale w jego rozważaniach znajduje się istotny trop. **Analogicznie do wyodrębniania się jednostek gramatycznych, proces gramatykalizacji mógł doprowadzić do wyodrębnienia się prozodii**. Według tego scenariusza z bogatej, ematywnej warstwy intonacyjnej protojęzyka wykształciło się kilka schematycznych wzorców intonacyjnych charakterystycznych dla współczesnych języków.

¹⁴ Zazwyczaj cała jednostka leksykalna ulega gramatykalizacji (por. Heine i Kuteva, 2007).

1.6. Koncepcja „nagiej prozodii” (ang. *bare phonology*)

We współczesnej nauce o ewolucji języka scenariusz powstania języka mający największy wpływ na badania zaproponował W. Tecumseh Fitch. O sukcesie tej propozycji do pewnego stopnia decyduje naukowa pozycja autora, który wraz z Noamem Chomskym i Markiem Hauserem napisał bodaj najważniejsze teksty nowoczesnego językoznawstwa – o zdolności językowej i jej ewolucji (Hauser, Chomsky i Fitch, 2002; Fitch, Hauser i Chomsky, 2005). Fitch jest światowej sławy ekspertem zajmującym się zwierzęcymi wokalizacjami – zarówno ich podłożem anatomicznym, jak i neuronalnym (np. Fitch, 2002; Fitch i in., 2016). Jednak koncepcja broni się sama, będąc obok pomysłu Arbiba jednym z najlepiej opracowanych i kompletnych scenariuszy ewolucyjnego powstania języka.

U jego podłoża znajduje się pogląd, że muzyka i język są uniwersalne dla wszystkich kultur (Nettl, 2000) i że posiadają taką samą elementarną organizację, to jest struktura zarówno muzyki, jak i języka opiera się na rozbudowanych, hierarchicznie zorganizowanych sygnałach, które podlegają międzypokoleniowej transmisji kulturowej (Fitch, 2010, s. 474–475). Kluczowe jest tutaj przekonanie o **generatywnym charakterze muzycznego protojęzyka** albo – jak pisze Fitch – prozodycznego protojęzyka. Teoria ta zakłada, że **ze stosunkowo niewielkiej liczby jego jednostek podstawowych za pomocą reguł można było derywować wielopiętrowe, zagnieżdżone struktury, podobnie jak ma to miejsce we współczesnej muzyce i we współczesnym języku** (np. Browman i Goldstein, 1989). Dlatego, podkreśla Fitch, ów prozodyczny protojęzyk mógł być zarówno źródłem muzyki, jak i języka – stąd alternatywne określenie „**protomuzyka**”. Z protojęzyka rozwinął się system fonologiczny mowy oraz forma muzyczna (gdzie podstawowe jednostki – nuty – łączą się w muzyczne frazy, motywy, tematy itd.; zob. Smith, 1996).

Gdy chodzi o mechanizmy, które doprowadziły do powstania muzycznego protojęzyka albo protomuzyki (por. Miller, 2000), Fitch zastanawia się nad udziałem w tym procesie **doboru seksualnego** wskazanego przez Darwina. Propozycja ta jest atrakcyjna między innymi dlatego, że stwierdzono udział doboru seksualnego w powstaniu podobnych systemów komunikacji u zwierząt – istnienie motywowanych

doborem seksualnym **skomplikowanych, hierarchicznie zbudowanych pieśni**, które muszą być przekazane przez tutorów w czasie okresu krytycznego (Thorpe, 1958; Brenowitz, Margoliash i Nordeen, 1997), zaobserwowano u wielu gatunków pochodzących z różnych taksonów zwierząt, na przykład u ptaków śpiewających (łac. *Oscines*), waleni z rzędu fiszbinowców (łac. *Mysticeti*) czy fok (łac. *Phocidae*) (Fitch, 2010, s. 490). Jednak scenariusz powstania języka i muzyki jedynie za sprawą doboru seksualnego ma dwie podstawowe wady. Przede wszystkim w znakomitej większości wypadków u śpiewających zwierząt występuje **dysproporcja płciowa** – zazwyczaj pieśni wykonują samce rywalizujący o samice¹⁵, co znajduje się w sprzeczności z tym, jak używany jest język, gdy na poziomie populacji nie notuje się dysproporcji w kompetencji językowej (np. Hyde i Linn, 1988; Lange, Euler i Zaretsky, 2016). Fitch wskazuje, że język mógł rozwinąć się w warunkach, gdzie dwie płcie rywalizują o partnera (ang. *mutual mate choice*), co zdają się potwierdzać pewne obserwacje etologiczne (Langmore i in., 1996), ale trudno sobie wyobrazić, jak dobór seksualny mógł doprowadzić nie tylko do wyłonienia się pierwotnego muzycznego protojęzyka, ale także do przekształcenia takiego protojęzyka w system zdolny do komunikacji znaczeń propozycyjalnych, a nie tylko emotywnych (Fitch, 2010, s. 492).

Fitch rozważa inny kontekst, który mógł być napędową ewolucji języka. Składać by się na niego miały **interakcje zachodzące w grupach rodzinnych**, a przede wszystkim interakcje wokalne między dzieckiem a matką. Powołując się na badania antropologiczne i porównawcze (Dissanayake, 2001; Falk, 2004; zob. też: Hrdy, 1981), badacz stwierdza, że ewolucyjna trajektoria człowiekowatych promowała wyłonienie się **wokalnej komunikacji służącej do utrzymywania kontaktu z potomstwem** (Fitch, 2010, s. 492–494). Czynniki, których pojawienie się w naszej linii ewolucyjnej miało się do tego przyczynić, to po pierwsze rozwój **dwunożności**, która uniemożliwiła ciągły kontakt z oseskiem¹⁶ charakterystyczny dla innych małp człekokształtnych, oraz **mała samodzielność potomstwa** (ang.

¹⁵ Istnieją też przypadki, w których samice ptaków konkurują o samców, używając pieśni (Fitch, 2010, s. 490).

¹⁶ Na przykład od momentu narodzin aż do końca pierwszego roku życia potomstwo szympansov doświadcza ciągłego fizycznego kontaktu z matką (Bard, 1994).

altriciality), która przede wszystkim wynikała z rosnącej encefalizacji człowiekowatych¹⁷ (Lewin i Foley, 2004). W związku z długim okresem dojrzewania pojawił się nacisk selekcyjny na tworzenie **par monogamicznych**, co znowu wyróżnia człowieka na tle poligamicznych małp człekokształtnych, i **uczestnictwo ojca w opiece na potomstwem** (Deacon, 1997). Według Fitcha stwarzało to warunki, aby komunikacja wokalna nie ograniczała się do interakcji matek z dziećmi, ale dotyczyła również ojców. W ten sposób rozwiązanie oparte na języku rodzicielskim omija problem nierównej dystrybucji umiejętności językowych, który przewiduje scenariusz odwołujący się do doboru seksualnego (Fitch, 2010, s. 493).

Jest też inny ważny moment dotyczący wyłonienia się z fatycznej komunikacji między dziećmi a rodzicami systemu zdolnego przenosić znaczenia propozycyjne. Otóż **kooperatywne normy użycia języka**, które każą dzielić się uczciwymi sygnałami (również) z niespokrewnionymi osobnikami (tj. przewidywania grice'owskie; Grice, 1975; por. Scott-Phillips, 2014), są kłopotliwe wobec założeń ewolucyjnej teorii sygnalizacji, gdzie podstawowym punktem odniesienia jest **cyniczny model komunikacji** Johna Krebsa i Richarda Dawkinsa (1984). Podstawową cechą sygnalizacji zwierzęcej jest manipulacja, a pojawienie się języka, który ma naturę kooperatywną (w wyżej opisanym sensie), wymagało specjalnych warunków. Założenie, że język najpierw ewoluował w **niszy krewniaczej** (tj. w warunkach doboru krewniaczego; zob. Hamilton, 1964a; 1964b), a dopiero później dzielenie się uczciwymi informacjami stało się normą obowiązującą członków całej populacji. Stanowi to dość atrakcyjne wyjaśnienie faktu, jak takie specjalne warunki mogły się pojawić (Fitch, 2010, s. 492–494).

Ostatecznie Fitch nie stwierdza, że motorem napędowym rozwoju języka był albo dobór seksualny, albo krewniaczy, sugerując, że jeden i drugi mechanizm – w różny sposób i na różnych etapach – mógł przyczynić się do powstania muzycznego protojęzyka i jego dalszego rozwoju (Fitch, 2010, s. 494). Finalnym osiągnięciem propozycji Fitcha, podobnie jak w przypadku Arbiba, jest wskazanie etapów rozwoju zdolności językowej, poczynając od pierwotnego muzycznego

¹⁷ Ze względu na pokaźne rozmiary mózgu ludzkie dzieci – ale też, mamy prawo przypuszczać, dzieci innych człowiekowatych, poczynając od *Homo erectus* – wymagały wydłużonego okresu dojrzewania, który potrzebny jest na pełne ukształtowanie się mózgu (Żywieczyński i Waciewicz, 2015, s. 231–232).

protojęzyka, a skończywszy na w pełni ukształtowanym języku. Kluczowym etapem w tym scenariuszu było wykształcenie się z muzycznego protojęzyka „**nagiej fonologii**” – systemu komunikacji wokalne, w którym jednostki łączyły się w hierarchicznie złożone, ale pozbawione propozycjonalnego znaczenia sygnały. Dla Fitcha te struktury stanowiły załączek **składni**, który należałoby jednak określić mianem protoskładni, ponieważ jej reguły nie służyły łączeniu jednostek (proto)leksykalnych, a jedynie formalnych elementów muzycznych. Opisując powstanie znaczenia, Fitch odwołuje się do scenariusza Jespersena i twierdzi, że stopniowo – w niszy krewniaczej – **poszczególne frazy muzycznego protojęzyka zaczynały być używane do oznaczania typów aktywności, zdarzeń, rytuałów i tym podobnych, co w końcu miało doprowadzić do powstania holistycznego protojęzyka** (zob. Wray, 1998). Badacz zwraca przy tym uwagę, że scenariusz, w którym (proto)język od początku ma naturę wokalną, promuje powstawanie arbitralnych związków między formą a znaczeniem, bez konieczności zakładania etapu ikonizacji w rozwoju (proto)języka, co głoszą różnego rodzaju teorie gesturalne. Poza tym ten holistyczny protojęzyk, choć stopniowo zyskiwał semantykę, cały czas był przede wszystkim narzędziem emocjonalnej ekspresji (por. Mithen, 2005). Podobnie jak w innych scenariuszach muzycznych u Fitcha **emocjonalną prozodię należy uznać za pozostałość tego właśnie systemu komunikacji**.

Kolejnym etapem ewolucji było wyłonienie się z holistycznego protojęzyka **znaczeń analitycznych łączonych ze sobą za pomocą składni**, której załączki pochodziły z systemu „nagiej fonologii”. Mechanizmem odpowiedzialnym za ten proces była przede wszystkim **ewolucja kulturowa**, rozumiana jako przekaz znaczeń w układzie horyzontalnym między uczestnikami interakcji oraz wertykalnym – między pokoleniami użytkowników (np. Kirby, 2017). Fitch powołuje się na ustalenia semiotyki eksperymentalnej (Galantucci, 2009), szczególnie uprawianej w paradygmacie iterowanego uczenia się (Kirby, 2017)¹⁸, które pokazują, że

¹⁸ Iterowane uczenie się jest zaczerpniętą z modelowania komputerowego metodą badania sposobu, w jaki komunikacja zmienia się w warunkach przekazu kulturowego. Za każdym razem podczas międzypokoleniowej transmisji języka (zazwyczaj jest to prosty język wymyślony dla celów eksperymentu) produkcja językowa (ang. *output*) starszego pokolenia P₁ stanowi dane wejściowe (ang. *input*) dla młodszego pokolenia P₂, następnie produkcja językowa pokolenia P₂ staje się danymi wejściowymi dla pokolenia P₃ itd. (Żywiczyński i Waciewicz, 2015, s. 324).

w warunkach przekazu kulturowego niezorganizowane systemy komunikacji stają się usystematyzowane, zyskując na przykład kompozycjonalność (Kirby, Cornish i Smith, 2008; Winters, Kirby i Smith, 2015). Na tym etapie zjawiska prozodyczne mogły zacząć być wyzyskiwane do pełnienia funkcji gramatycznych i pragmatycznych, na przykład określania siły illokucyjnej wypowiedzi.

Ostatnim rozdziałem scenariusza proponowanego przez Fitcha jest powstanie współczesnego języka. Zakłada on, że na zasadzie koewolucji zmian kulturowych i biologicznych powstała silna presja selekcyjna, aby kolejne pokolenia użytkowników języka były coraz lepiej kognitywnie wyposażone do nabywania znaczeń analitycznych i reguł składniowych, przez co język stawał się jeszcze bardziej kompozycjonalny, aż osiągnął swoją współczesną formę (Fitch, 2010, s. 504).

Fitch podkreśla, że na tym etapie ewolucji język przede wszystkim zaczął służyć przekazywaniu uczciwych informacji i rozwijał się w grupach krewniaczych.

Żaden z innych scenariuszy muzycznych (w sensie silnym) nie dorównuje kompletnością teorii Fitcha – propozycje te mają w najlepszym wypadku walor przyczynkarski. Z tych powodów warto ponownie wspomnieć o **integracyjnym scenariuszu Mithena** (2005), który oprócz własności multimodalnych i mimetycznych proponowanego przez siebie protojęzyka, podkreślał jego muzyczny charakter. Najistotniejszym osiągnięciem jego pracy wydaje się postulat metodologiczny, że badania neurokognitywne, w tym porównawcze, nad językiem i muzyką mogą dać wgląd w ewolucyjną przeszłość obu systemów. Przedstawieniu tych zagadnień Steven Mithen poświęcił całą pierwszą część książki *The singing Neanderthals* (Śpiewający neandertalczyki) (Mithen, 2005, s. 11–104), w której podjął **problem prozodii – między innymi intonacji i rytmu w kontekście przypadków**

afazji i amuzji¹⁹ – dochodząc do, wydaje się, pochoptych wniosków, że **prozodia może stanowić ewolucyjny łącznik między językiem a muzyką** (por. Ahlsén, 2006, s. 55–63). Innym rozpoznawalnym pomysłem jest koncepcja **muzyjęzyka** (ang. *musilanguage*) zaproponowana przez muzykologa Stevena Browna (2000). Wychodząc od scenariusza Darwina, Brown, podobnie jak Fitch, twierdzi, że u podstaw współczesnego języka i muzyki leżał jeden system – muzyjęzyk. Jego

¹⁹ Wskazał na przykład na odporność kompetencji prozodycznej w nawet bardzo poważnych przypadkach afazji (Mithen, 2005, s. 46–61).

podstawową właściwością było użycie dyskretnych wysokości dźwięku do wyrażania treści emotywno-referencyjnych. W propozycji Browna brakuje jednak wyjaśnień dotyczących tego, co miało zadecydować o powstaniu takiego systemu i jak jego ewolucja doprowadziła do powstania współczesnego języka z jednej strony, a muzyki – z drugiej.

1.7. Perspektywy rozwoju badań nad prozodią w nauce o ewolucji języka

Powyższy przegląd uwidacznia brak zorganizowanej refleksji nad prozodią w nauce o ewolucji języka. Pokazuje on jednak zarazem, że istnieje duża potrzeba podjęcia takich badań, bo w wielu znaczących stanowiskach na temat ewolucji języka pojawia się problem prozodii, a w niektórych z nich – przede wszystkim w stanowiskach muzycznych, takich jak koncepcja Fitcha – jest on jednym z problemów kluczowych.

Najważniejsze z tych obszarów, które mogą rzucić światło na ewolucję prozodii, to **badania neuropoznawcze nad prozodią, badania porównawcze nad komunikacją wokalną zwierząt**, a przede wszystkim małp człekokształtnych, oraz **badania z zakresu semiotyki eksperymentalnej**, których przedmiotem będzie wyłanianie się cech przypominających prozodię w warunkach transmisji kulturowej.

Jeśli chodzi o ten pierwszy kierunek badań, dotychczasowe ustalenia nie dają jasnego obrazu, jak silny jest związek prozodii z innymi aspektami języka oraz z zachowaniami muzycznymi. Z jednej strony są badania, które wskazują, że wśród deficytów językowych u pacjentów z uszkodzeniami lewej półkuli móz-gowej (ang. *left hemisphere damage*, LHD), a w szczególności z uszkodzeniami w obszarze Broki, występuje dysprozodia objawiająca się na przykład problemami z odpowiednim kładzeniem akcentu lub intonacją (Ahlsén, 2006, s. 63). Z drugiej strony stwierdzono też przypadki nawet głębokiej afazji, w których produkcja i rozumienie prozodii były niezaburzone (np. Yaqub i in., 1988; Takahashi i in., 1992). Istnieją też intrygujące przypadki afazji, w których wprawdzie pojawia się dysprozodia, ale pacjenci są w stanie poprawnie identyfikować walencję

emocjonalną wypowiedzi (Mendez, 2001), oraz przypadki amuzji, w których u części pacjentów stwierdzano deficyty prozodyczne (np. Piccirilli i in., 1994), a u innych – nie (np. Peretz i in., 1994). Dogłębne poznanie skomplikowanych relacji między prozodią, językiem i muzyką na poziomie neuronalnym z pewnością rzuci nowe światło na ewolucyjną przeszłość tych systemów.

Inny obszar badań – kluczowy dla zrozumienia, jak powstała prozodia – stanowi ewolucyjny punkt wyjścia dla wszystkich rozważań o powstaniu mowy. Chodzi tu o **wokalną komunikację małp człekokształtnych**. Zakłada się, że komunikacja ta w dużym stopniu odzwierciedla zdolności wokalne wspólnego przodka małp i ludzi, a w szczególności ostatniego wspólnego przodka szympan-sów i ludzi, który żył około siedmiu milionów lat temu (Fitch, 2010, s. 234–249; Lewin i Foley, 2004). Dominuje pogląd, że wokalizacje małp, w tym małp człeko-kształtnych, mają charakter sztywny i automatyczny, a najczęstszym czynnikiem je wywołującym jest pobudzenie emocjonalne (Deacon, 1997; Fitch, 2000). Wo-licjonalny i innowacyjny charakter ludzkich wokalizacji ma być gwarantowany przez bezpośrednie połączenie między korą a aparatem wokalnym, o którym była mowa wcześniej. Jednak niektóre z nowszych badań etologicznych wskazują, że małpy posiadają większą kontrolę nad wokalizacjami, niż pierwotnie sądzono.

I tak wolno żyjące szympanasy modyfikują swoje wokalizacje w zależności od tego, kto może je usłyszeć (tzw. „efekt publiczności”, np. Slocombe, 2011), a ich wokalne wymiany mają charakter interaktywny, który pod pewnymi względami przypomina kolejność zabierania głosu w konwersacji (ang. *turn-taking*, np. Fröhlich i in., 2016). W końcu rośnie liczba danych dotyczących umiejętności zmiany przez małpy cech spektrotympralnych (takich jak rytm albo głośność) wydawanych przez siebie wokalizacji – na przykład pigmejki karłowate (łac. *Cebuella pygmaea*) dostosowują głośność wydawanych przez siebie okrzyków kontaktowych do otoczenia, w którym się właśnie znajdują (wykonując głośniejsze okrzyki w miejscach gęściej zaludnionych; de la Torre i Snowdon, 2002). Podobne zjawiska zaobserwowano również u przedstawicieli innych taksonów prymatów, w tym szympan-sów (Notman i Rendall, 2005). Reasumując: ustalenie, jaki jest faktyczny dystans między ludzkimi a małpimi zachowaniami wokalnymi, w tym zachowaniami prozodycznymi, jest warunkiem koniecznym do podjęcia

empirycznie wiarygodnych prób rekonstrukcji ewolucji mowy, włączając w to ewolucję prozodii.

Semiotyka eksperymentalna to pole, które dostarcza ważnego dla ewolucji języka materiału empirycznego, szczególnie jeśli chodzi o wpływ przekazu kulturowego na ewolucję systemów komunikacyjnych. Niektóre z tych badań skoncentrowane są na zachowaniach wokalnych. Marcus Perlman prowadzi szeroko zakrojone badania nad ikonicznym potencjałem niejęzykowych wokalizacji (np. Perlman, Dale i Lupyan, 2015; Perlman, 2017), w tym ikonicznego wykorzystywania prozodii (Perlman, Clark i Johansson Falck, 2015). Warto też wspomnieć o pionierskich wysiłkach zbadania, jak w reżymie symulowanego przekazu kulturowego powstają elementy struktury muzycznej. Na przykład Ravignani i współautorzy (2016) pokazali, jak z losowo wygenerowanych sekwencji perkusyjnych uczestnicy eksperymentu – w kolejnych iteracjach – tworzyli coraz bardziej uporządkowane wzorce rytmiczne, a finalnym rezultatem badania było wyłonienie się sześciu wzorców, które najczęściej powtarzają się w różnych kulturach muzycznych. Z pewnością rozwój badań semiotyczno-eksperymentalnych nad powstawaniem elementów prozodii wniesie duży wkład w zrozumienie, jak prozodia rozwija się w obrębie kulturowej ewolucji języka.

Istnieje olbrzymia dysproporcja między uwagą, jaką badacze ewolucji języka poświęcają prozodii, a uwagą poświęcaną innym elementom systemu językowego. Przy braku prac, które koncentrowałyby się na ewolucji prozodii jako osobnym problemie badawczym, powstały już dziesiątki, jeśli nie setki, opracowań dotyczących ewolucji fonologii, semantyki i składni, a także gestów towarzyszących mowie czy języków migowych. Taki stan rzeczy trwa pomimo tego, że wiele scenariuszy powstania języka zwraca uwagę na prozodię i potrzebuje ustaleń dotyczących jej ewolucji oraz roli, jaką ta ewolucja odegrała w całościowym procesie wyłaniania się zdolności językowej. Perspektywy rozwoju badań nad tymi problemami wskazane powyżej każą mieć nadzieję, że wkrótce pojawią się przedsięwzięcia naukowe im dedykowane, dzięki czemu wzrośnie nasza wiedza nie tylko o ewolucji języka, ale również – sądząc po wpływie, jaki badania ewolucyjne miały na zrozumienie innych aspektów języka – o samej prozodii.

Streszczenie

Nauka o ewolucji języka nie podjęła jeszcze szeroko zakrojonych badań nad prozodią, choć refleksja nad zjawiskami prozodycznymi pojawia się w wielu teoriach powstania języka. Celem rozdziału jest wskazanie tych stanowisk, w których problem prozodii albo został już podjęty, albo wskazano jego miejsce w procesie kształtowania się zdolności językowej. Według pantomimicznych oraz multimodalnych scenariuszy prozodia miała wyłonić się z emocjonalnych wokalizacji, które wraz z komunikacją gestową tworzyły pierwotny język. Natomiast zwolennicy stanowisk muzycznych utrzymują, że cechy prozodyczne leżały u podstaw wyłaniającego się systemu komunikacji językowo-muzycznej. W rozdziale wskazano też te obszary badań, które wydają się kluczowe dla rozwoju badań nad ewolucją prozodii.

Słowa kluczowe

ewolucja języka, prozodia, protojęzyk, pantomima, gesty, ewolucja muzyki, komunikacja małp człekokształtnych i ptaków, dobór płciowy, ewolucja kulturowa

Summary

The science of language evolution has not yet undertaken extensive research on prosody, although the reflections on prosodic phenomena can be found in many hypotheses of language origin. The purpose of this chapter is to identify the research which either directly addresses prosody or at least identifies its place in language origin scenarios. According to pantomimic and multimodal conceptions, prosody emerged from emotional vocalisations, which, together with gestures, comprised the early form of language. The advocates of musical origins of language, on the other hand, maintain that prosodic features were at the very foundation of the emerging linguistic-musical communication system. The chapter highlights the areas of investigation which appear crucial to the development of research into the evolution of prosody.

Key words

evolution of language, prosody, protolanguage, pantomime, gestures, evolution of music, ape and bird communication, sexual selection, cultural evolution

Bibliografia

- Ahlsén E. (2006). *Introduction to neurolinguistics*. Amsterdam - Philadelphia, PA: John Benjamins.
- Arbib M.A. (2005). From monkey-like action recognition to human language: An evolutionary framework for neurolinguistics. *Behavioral and Brain Sciences*, 28, 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.
- Bachorowski J.A. (1999). Vocal expression and perception of emotion. *Current Directions in Psychological Science*, 8(2), 53-57, DOI: 10.1111/1467-8721.00013.
- Bard K. (1994). Similarities and differences in the neonatal behavior of chimpanzee and human infants. W: G. Eder, E. Kaiser, F. King (red.), *The role of the chimpanzee in research* (s. 43-55). Basel: Karger.
- Barrett M.J. (1977). Masticatory and non-masticatory uses of teeth. W: R.V.S.Wright (red.), *Stone tools as cultural markers: change, evolution and complexity* (s. 18-23). Canberra: Australian Institute of Aboriginal Studies.
- Berwick R., Chomsky N. (2011). The biolinguistic program: The current state of its development. W: A.M. Di Sciullo, C. Boeckx (red.), *The biolinguistics enterprise: New perspectives on the evolution and nature of the human language faculty* (s. 19-41). Oxford: Oxford University Press.
- Bickerton D. (1990). *Language and species*. Chicago, IL: University of Chicago Press.
- Bickerton D. (1998). Catastrophic evolution: The case for a single step from protolanguage to full human language. W: J. Hurford, M. Studdert-Kennedy, C. Knight (red.), *Approaches to the evolution of language: Social and cognitive bases* (s. 341-358). Cambridge: Cambridge University Press.
- Bickerton D. (2003). Symbol and structure: A comprehensive framework of language evolution. W: M.H. Christiansen, S. Kirby (red.), *Language evolution* (s. 77-93). Oxford: Oxford University Press.
- Bickerton D. (2009). *Adam's tongue: How humans made language, how language made humans*. New York, NY: Macmillan.

- Brenowitz E.A., Margoliash D., Nordeen K.W. (1997). An introduction to birdsong and the avian song system. *Journal of Neurobiology*, 33, 495-500.
- Browman C.P., Goldstein L. (1989). Articulatory gestures as phonological units. *Phono-logy*, 6(2), 201-251.
- Brown S. (2000). The „musilanguage” model of music evolution. W: N.L. Wallin, B. Merker, S. Brown (red.), *The origins of music* (s. 271-300). Cambridge, MA: MIT Press.
- Burling R. (2005). *The talking ape: How language evolved*. Oxford: Oxford University Press.
- Bynon T. (1977). *Historical linguistics*. Cambridge: Cambridge University Press.
- Chomsky N. (1995). *The minimalist program*. Cambridge, MA: MIT Press.
- Clement A.F., Hillson S.W., Aiello L.C. (2012). Tooth wear, Neanderthal facial morphology and the anterior dental loading hypothesis. *Journal of Human Evolution*, 62(3), 367-376.
- Condillac É.B. (1756). *An essay on the origin of human knowledge (Essai sur l'origine des connaissances humaines)*. Przeł. T. Nugent. London: J. Nourse.
- 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), 199-208.
- Darwin C. (1859). *On the origin of species by means of natural selection*. London: John Murray.
- Darwin C. (1871). *The descent of man, and selection in relation to sex*. London: John Murray.
- Darwin C. (1872). *The expression of the emotions in man and animals*. London: John Murray.
- Deacon T.W. (1992). The neural circuitry underlying primate calls and human language. W: J. Wind, B. Chiarelli, B. Bichakjian, A. Nocentini, A. Jonker Winds (red.), *Language origin: A multidisciplinary approach* (s. 121-162). Dordrecht: Kluwer.
- Deacon T.W. (1997). *The symbolic species: The coevolution of language and the brain*. New York, NY: W.W. Norton.
- Dennett D. (2017). *From bacteria to Bach and back: The evolution of minds*. New York, NY: W.W. Norton.
- Dissanayake E. (2001). *Homo aestheticus: Where art comes from and why*. Seattle, WA: University of Washington Press.
- Donald M. (1991). *Origins of the modern mind: three stages in the evolution of culture and cognition*. Cambridge, MA: Harvard University Press.
- Donald M. (2001). *A mind so rare: The evolution of human consciousness*. New York, NY: W.W. Norton.
- Dunbar R. (1996). *Grooming, gossip and the evolution of language*. London: Faber and Faber.

- 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(1), 448-458.
- Falk D. (2004). Prelinguistic evolution in early hominins: Whence motherese? *Behavioral and Brain Sciences*, 27(4), 491-503.
- Fay N., Arbib M., Garrod S. (2013). How to bootstrap a human communication system. *Cognitive Science*, 37(7), 1356-1367.
- Fay N., Lister C.J., Ellison T.M., Goldin-Meadow S. (2014). Creating a communication system from scratch: Gesture beats vocalization hands down. *Frontiers in Psychology*, 5, 354.
- Filippi P. (2016). Emotional and interactional prosody across animal communication systems: A comparative approach to the emergence of language. *Frontiers in Psychology*, 7, 1393.
- Fitch W.T. (2000). The evolution of speech: A comparative review. *Trends in Cognitive Sciences*, 4(7), 258-267.
- Fitch W.T. (2002). The evolution of language comes of age. *Trends in Cognitive Sciences*, 6(7), 278-279.
- Fitch W.T. (2005). Protomusic and protolanguage as alternatives to protosign. *Behavioral and Brain Sciences*, 28(2), 132-133.
- Fitch W.T. (2006). The biology and evolution of music: A comparative perspective. *Cognition*, 100(1), 173-215.
- Fitch W.T. (2010). *The evolution of language*. Cambridge: Cambridge University Press.
- Fitch W.T., Hauser M., Chomsky N. (2005). The evolution of the language faculty: Clarifications and implications. *Cognition*, 97(2), 179-210.
- Fitch W.T., de Boer B., Mathur N., Ghazanfar A. (2016). Monkey vocal tracts are speech-ready. *Science Advances*, 2(12), e1600723.
- Fröhlich M., Kuchenbuch P., Müller G., Fruth B., Furuichi T., Wittig R., Pika S. (2016). Unpeeling the layers of language: Bonobos and chimpanzees engage in cooperative turn-taking sequences. *Scientific Reports*, 6, 25887.
- Galantucci B. (2009). Experimental semiotics: A new approach for studying communication as a form of joint action. *Topics in Cognitive Science*, 1(2), 393-410.
- Gardner R.A., Gardner B.T. (1969). Teaching sign language to a chimpanzee. *Science*, 165, 664-672.
- Gärdenfors P. (2017). Demonstration and pantomime in the evolution of teaching. *Frontiers in Psychology*, 8, 415.
- Gärdenfors P., Högberg A. (2017). The archaeology of teaching and the evolution of *Homo docens*. *Current Anthropology*, 58(2), 188-208.

- Ghazanfar A., Rendall D. (2008). Evolution of human vocal production. *Current Biology*, 18(11), 457-460.
- Goffman E. (1967). *Interaction ritual: Essays on face-to-face behavior*. New York, NY: Doubleday.
- Goldin-Meadow S. (2011). What modern-day gesture can tell us about language evolution. W: M. Tallerman, K.R. Gibson (red.), *The Oxford handbook of language evolution* (s. 545-557). Oxford: Oxford University Press.
- Grice H.P. (1975). Logic and conversation. W: P. Cole, J. Morgan (red.), *Syntax and semantics* (t. 3: *Speech acts*, s. 41-58). London: Academic Press.
- Hamilton W.D. (1964a). The genetical evolution of social behavior I. *Journal of Theoretical Biology*, 7(1), 1-16.
- Hamilton W.D. (1964b). The genetical evolution of social behavior I. *Journal of Theoretical Biology*, 7(1), 17-52.
- Hauser M., 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.
- Hayes K., Hayes C. (1952). Imitation in a home-raised chimpanzee. *Journal of Comparative and Physiological Psychology*, 45, 450-459.
- Heine B., Kuteva T. (2007). *The genesis of grammar: A reconstruction*. Oxford: Oxford University Press.
- Hewes G.W. (1973). Primate communication and the gestural origin of language. *Current Anthropology*, 14(1-2), 5-24.
- Hockett C. (1960). The origin of speech. *Scientific American*, 203, 88-111.
- Hrdy S.B. (1981). *The woman that never evolved*. Cambridge, MA: Harvard University Press.
- Hyde J.S., Linn M. (1988). Gender differences in verbal ability: A meta-analysis. *Psychological Bulletin*, 104, 53-69.
- Itard J. (1801/1802). *A historical account of the discovery and education of a savage man: Or, the first developments, physical and moral, of the young savage caught in the woods near Aveyron in the year 1798 (Mémoire et rapport sur Victor de l'Aveyron)*. London: Richard Phillips.
- Jackendoff R. (2002). *Foundations of language: Brain, meaning, grammar, evolution*. New York, NY: Oxford University Press.
- Jespersen O. (1922). *Language: Its nature, development and origin*. London: Allen and Unwin.
- Johansson S. (2005). *Origins of language: Constraints on hypotheses*. Amsterdam-Philadelphia, PA: John Benjamins.

- Jürgens U. (1994). The role of the periaqueductal grey in vocal behavior. *Behavioral Brain Research*, 62(2), 107-117.
- 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* (s. 179-237). Cambridge, MA: MIT Press.
- Kellogg W.N., Kellogg L.A. (1933). *The ape and the child: A comparative study of the environmental influence upon early behavior*. New York, NY: Hafner.
- Kendon A. (1991). Some considerations for a theory of language origins. *Man*, 26, 199-221.
- Kendon A. (2004). *Gesture: Visible action as utterance*. Cambridge: Cambridge University Press.
- 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*, 3(1), 79-108.
- Kendon A. (2014). The „poly-modalic” nature of utterances and its relevance for inquiring into language origins. W: D. Dor, C. Knight, J. Lewis (red.), *The social origins of language* (s. 67-76). Oxford: Oxford University Press.
- Kirby S. (2017). Culture and biology in the origins of linguistic structure. *Psychonomic Bulletin and Review*, 24(1), 118-137.
- Kirby S., Cornish H., Smith K. (2008). Cumulative cultural evolution in the laboratory: An experimental approach to the origins of structure in human language. *Proceedings of the National Academy of Sciences*, 105(31), 10681-10686.
- Krebs J.R., Dawkins R. (1984). Animal signals: Mind-reading and manipulation. W: J.R. Krebs, N.B. Davies (red.), *Behavioural ecology: An evolutionary approach* (s. 380-402). Oxford: Blackwell.
- Kuypers H. (1958). Corticobulbar connections to the pons and lower brainstem in man. *Brain*, 81, 364-388.
- Lange B., Euler H., Zaretsky E. (2016). Sex differences in language competence of 3- to 6-year-old children. *Applied Psycholinguistics*, 37(6), 1417-1438.
- Langmore N.E., Davies N.B., Hatchwell B.J., Hartley I.R. (1996). Female song attracts males in the alpine accentor *Prunella collaris*. *Proceedings of the Royal Society of London*, B263(1367), 141-146.
- Laromiguière P. (1826). *Leçons de philosophie sur les principes de l'intelligence, ou sur les causes et sur les origines des idées*. Paris: Brunot Labbe.
- Levinson S.C. (2006). On the human interaction engine. W: N. Enfield, S.C. Levinson (red.), *Roots of human sociality* (s. 39-69). Oxford: Berg Publishers.
- Lewin R., Foley R. (2004). *Principles of human evolution*. Oxford: Blackwell.

- MacNeilage P.F. (2008). *The origin of speech*. Oxford: Oxford University Press.
- de Mandeville B. (1729). *The fable of the bees, part II*. London: J. Roberts.
- McNeill D. (1992). *What gestures reveal about thought*. Chicago, IL: University of Chicago Press.
- McNeill D. (2005). *Gesture and thought*. Chicago, IL: University of Chicago Press.
- McNeill D. (2012). *How language began: Gesture and speech in human evolution*. Cambridge: Cambridge University Press.
- Mendez M.F. (2001). Generalized auditory agnosia with spared music recognition in a left-hander: Analysis of a case with a right temporal stroke. *Cortex*, 37(1), 139-150.
- Miller G. (2000). Evolution of human music through sexual selection. W: N.L. Wallin, B. Merker, S. Brown (red.), *The origins of music* (s. 329-360). Cambridge, MA: MIT Press.
- Mithen S. (2005). *The singing Neanderthals: The origins of music, language, mind and body*. London: Weindelfed and Nicholson.
- Mocerino R.P. (2016). Gesture, interjection and onomatopoeia in Edward Burnett Tylor's theory of the origin and development of language. *Theoria et Historia Scientiarum*, 13, 71-84.
- Motamedi-Mousavi Y., Schouwstra M., Culbertson J., Smith K., Kirby S. (2017). The cultural evolution of complex linguistic constructions in artificial sign languages. W: *CodSci 2017: Proceedings of the 39th Annual Conference of the Cognitive Science Society, London, UK, 26-29 July 2017, Computational foundations of cognition* (s. 2760-2765), online: https://cogsci.mindmodeling.org/2017/cogsci17_proceedings.pdf [dostep: 24 marca 2020].
- Nettl B. (2000). An ethnomusicologist contemplates universals in musical sound and musical culture. W: N.L. Wallin, B. Merker, S. Brown (red.), *The origins of music* (s. 463-472). Cambridge, MA: MIT Press.
- Notman H., Rendall D. (2005). Contextual variation in chimpanzee pant hoots and its implications for referential communication. *Animal Behavior*, 70, 177-190.
- Patterson F.G. (1981). Ape language. *Science*, 211(4477), 86-87.
- Peirce C.S. (1868). On a new list of categories. *Proceedings of the American Academy of Arts and Sciences*, 7(1868), 287-298.
- Peretz I., Kolinsky R., Tramo M., Labrecque R., Hublet C., Demeurisse G., Belleville S. (1994). Functional dissociations following bilateral lesions of auditory cortex. *Brain*, 117(6), 1283-1301.
- Perlman M. (2017). Debunking two myths against vocal origins of language. *Interaction Studies*, 18(3), 376-401.

- Perlman M., Dale R.D., Lupyan G. (2015). Iconicity can ground the creation of vocal symbols, *Royal Society Open Science*, 2, 150-152.
- Perlman M., Clark N., Johansson Falck M. (2015). Iconic prosody in story reading. *Cognitive Science*, 39(6), 1348-1368.
- Piccirilli M., D'Alessandro P., Sciarma T., Cantoni C., Dioguardi M.S., Giuglietti M., Ibba A., Tiacci C. (1994). Attention problems in epilepsy: Possible significance of the epileptogenic focus. *Epilepsia*, 35(5), 1091-1096.
- 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-Philadelphia, PA: John Benjamins.
- Pika S., Liebal K., Call J., Tomasello M. (2005). The gestural communication of apes. *Gesture*, 5(1-2), 41-56.
- Ravignani, A., Delgado, T., Kirby, S. (2016). Musical evolution in the lab exhibits rhythmic universals. *Nature Human Behaviour*, 1(1), 1-7.
- Rizzolatti G., Arbib M.A. (1998). Language within our grasp. *Trends in Neurosciences*, 21(5), 188-194.
- Rousseau J.J. (1781/1998). *Essay on the origin of languages and writings related to music*. Przeł. J.T. Scott. Hanover, NH: University Press of New England.
- Sandler W., Meir I., Padden C., Aronoff M. (2005). The emergence of grammar: Systematic structure in a new language. *Proceedings of the National Academy of Sciences*, 102(7), 2661-2665.
- Sandler W., Aronoff M., Padden C., Meir I. (2014). Language emergence Al-Sayyid Bedouin sign language. W: J. Sidnell, P. Kockelman, N. Enfield (red.), *The Cambridge handbook of linguistic anthropology* (s. 246-278). Cambridge: Cambridge University Press.
- Schouwstra M., de Swart H. (2014). The semantic origins of word order. *Cognition*, 131, 431-436.
- Scott-Phillips T.C. (2014). *Speaking our minds: Why human communication is different, and how language evolved to make it special*. Basingstoke: Palgrave Macmillan.
- Slocombe K. (2011). Have we underestimated great ape vocal capacities? W: M. Tallerman, K. Gibson (red.), *The Oxford handbook of language evolution* (s. 90-95). Oxford: Oxford University Press.
- Slocombe K.E., Zuberbühler K. (2007). Chimpanzees modify recruitment screams as a function of audience composition. *Proceedings of the National Academy of Sciences*, 104(43), 17228-17233.
- Smith A. (2008). Protolanguage reconstructed. *Interaction Studies*, 9(1), 100-116.

- Smith C.J. (1996). Musical form and fundamental structure: An investigation of Schenker's Formenlehre. *Music Analysis*, 15, 191-297.
- Sonesson G. (2007). From the meaning of embodiment to the embodiment of meaning. W: T. Ziemke, J. Zlatev, R.M. Frank (red.), *Body, language and mind* (t. 1: *Embodiment*, s. 85-128). Berlin: Mouton de Gruyter.
- Stokoe W.C. (1991). Semantic phonology. *Sign Language Studies*, 71, 107-114.
- Takahashi N., Kawamura M., Shinotou H., Hirayama K., Kaga K., Shindo M. (1992). Pure word deafness due to left hemisphere damage. *Cortex*, 28(2), 295-303.
- Thorpe W.H. (1958). The learning of song patterns by birds, with especial reference to the song of the chaffinch *Fringilla coelebs*. *Ibis*, 100, 535-570.
- Tomasello M. (2008). *Origins of human communication*. Cambridge, MA: MIT Press.
- Tomasello M. (2009). *Why we cooperate*. Cambridge, MA: MIT Press.
- de la Torre S., Snowdon C.T. (2002). Environmental correlates of vocal communication of wild pygmy marmosets, *Cebuella pygmaea*. *Animal Behavior*, 63(5), 847-856.
- Tylor E.B. (1870). *Researches into the early history of mankind and the development of civilization*. London: John Murray.
- Tylor E.B. (1871). *Primitive culture: Researches into the development of mythology, philosophy, religion, language, art, and customs*. London: John Murray.
- Vigliocco G., Perniss P., Vinson D. (2014). Language as a multimodal phenomenon: Implications for language learning, processing and evolution. *Philosophical Transactions of the Royal Society, B: Biological Sciences*, 369(1651), 20130292.
- Vygotsky L. (1934/1964). *Thinking and speech*. Przeł. A. Kozulin. Cambridge, MA: MIT Press.
- Waciewicz S., Żywiczyński P. (2018). Language origins: Fitness consequences, platform of trust, cooperation and turn-taking. *Interaction Studies*, 19(1-2), 197-182.
- Whitney W.D. (1872). *Oriental and linguistic studies: The Veda, the Avesta, the science of language*. New York: Scribner and Armstrong.
- Winters J., Kirby S., Smith K. (2015). Languages adapt to their contextual niche. *Language and Cognition*, 7(3), 415-449.
- Wray A. (1998). Protolanguage as a holistic system for social interaction. *Language and Communication*, 18(1), 47-67.
- Yaqub B.A., Gascon G.G., Noshah M.A., Whitaker H. (1988). Pure word deafness (acquired verbal auditory agnosia) in an Arabic speaking patient. *Brain*, 111(2), 457-466.
- 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-Philadelphia, PA: John Benjamins.

- Zlatev J. (2014). Human uniqueness, bodily mimesis and the evolution of language. *Humana Mente. Journal of Philosophical Studies*, 27, 197-219.
- Zlatev J., Persson T., Gärdenfors P. (2005). Triadic bodily mimesis is the difference. *Behavioral and Brain Sciences*, 28(5), 720-721.
- Zlatev J., Wacewicz S., Żywiczyński P., van de Weijer J. (2017). Communicating events through pantomime with and without vocalization. *Interaction Studies*, 18(3), 455-475.
- Żywiczyński P. (2018). *Language origins: From mythology to science*. Berlin: Peter Lang.
- Żywiczyński P., Gontier N., Wacewicz S. (2017). The evolution of (proto-)language: Focus on mechanisms. *Language Sciences*, 63, 1-11.
- Żywiczyński P., Wacewicz S. (2015). *Ewolucja języka. W stronę hipotez gesturalnych*. Toruń: Wydawnictwo Naukowe UMK.
- Żywiczyński P., Wacewicz S., Sibierska M. (2018). Defining pantomime for language evolution research. *Topoi*, 37(2), 307-318.