

## Ile jest języków?\*

### How many languages exist?

Jerzy Pogonowski<sup>†</sup>

Institute of Linguistics, Adam Mickiewicz University

ul. Międzychodzka 5, 60-371 Poznań, POLAND

[pogon@amu.edu.pl](mailto:pogon@amu.edu.pl)

#### Abstract

This is the text of a talk given on March 2, 2002 at the conference *Applications of Algebra, VI* in Zakopane. We share with the audience a few reflections concerning the problem of how many languages there are. Logical, mathematical, linguistic and methodological presuppositions of this question are taken into account.

#### 1 Jaki sens ma powyższe pytanie?

W lingwistyce matematycznej przez *język* rozumie się dowolny zbiór skończonych ciągów elementów ustalonego (skończonego lub przeliczalnego) alfabetu. Jeśli więc alfabet  $A$  jest skończony lub przeliczalny, to wszystkich skończonych ciągów jego elementów jest przeliczalnie wiele. W konsekwencji, wszystkich języków nad alfabetem  $A$  jest  $2^{\aleph_0}$ , czyli kontinuum. Ponieważ zbiorów rekurencyjnych jest przeliczalnie wiele, więc *prawie wszystkie* języki (w powyższym sensie) to zbiory, dla których nie istnieje efektywna metoda pozwalająca ustalić co jest, a co nie jest elementem tak rozumianego języka. W innym jeszcze, nieco uproszczonym, sformułowaniu: ponieważ istnieje jedynie przeliczalnie wiele gramatyk formalnych, więc nieprzeliczalna liczba języków nie ma efektywnego opisu syntaktycznego.

Uznany standardem metodologicznym w opisie języków rozpatrywanych w logice jest jednoznaczne określenie ich składni (np. opis kategoryalny) oraz semantyki (np. klasyczne ujęcie tarskian-skie). Gdy język jest ustalony, logika interesować mogą operacje konsekwencji w tym języku, algebraika zaś struktury, o których w tymże języku mowa.

Są to wszystko powszechnie znane oczywistości i nie to jest tematem niniejszych rozważań. Będą nas mianowicie interesować *języki naturalne* — istniejące, możliwe oraz brane *in abstracto*, jako pewne *typy* systemów semiotycznych, z uwzględnieniem proponowanych przez lingwistów (a także logików, matematyków i filozofów) ich teoretycznych opisów.

Na postawione w tytule pytanie dać można różne *prawidłowe* odpowiedzi. Ich postać zależy od perspektywy semiotycznej — od tych aspektów procesu komunikacji symbolicznej, które uważamy

\*Tekst referatu wygłoszonego 2 marca 2002 roku podczas konferencji *Zastosowania Algebry VI* organizowanej w Zakopanem-Jaszczurówce przez Instytut Matematyki i Informatyki Wyższej Szkoły Pedagogicznej w Częstochowie. Serdecznie dziękuję organizatorom — przede wszystkim Pani Profesor Joannie Grygiel oraz Panu Profesorowi Piotrowi Wojtylakowi — za życzliwą zachętę do wygłoszenia tego referatu. Odczyt miał charakter popularyzatorski, jego celem było pokazanie logikom i algebraikom garści ciekawostek dotyczących języków świata.

<sup>†</sup>Autor jest pracownikiem Zakładu Logiki Stosowanej w Instytucie Językoznawstwa UAM. Instytut nie potrafił partycipować finansowo w udziale autora w tej konferencji.

(w danym ujęciu) za niemożliwe do pominięcia. Ewentualnym pożytkiem ze sformułowania tych odpowiedzi oraz ich porównania może być np. ustalenie hierarchii celów badań lingwistycznych i ukazanie — w wielu przypadkach — interdyscyplinarności tych badań. Nadto, interesujące wydaje się zwrócenie szczególnej uwagi na rolę rozumowań przez analogię w językoznawstwie. Przy tego rodzaju rozumowaniach dokonujemy — świadomie — znacznych uproszczeń w opisie tak złożonych systemów, jakimi są języki naturalne. Mamy przecie w tym wypadku do czynienia z *żywymi* systemami semiotycznymi, zmiennymi w czasie, uwarunkowanymi społecznie, których pierwszoplanowe funkcje komunikacyjne mają charakter *pragmatyczny*. Od właśnie wymienionych czynników najczęściej abstrahujemy w logiczno-matematycznych opisach języka. Matematyka i logika nie ma jednak zakazu wstępu na te obszary badawcze. Co więcej, nie ma podstaw, aby sądzić, że matematyczny opis języków naturalnych jest gorszy od opisów „czysto lingwistycznych”. Może jest nawet na odwrót?

## 2 Ustalenia lingwistów

Odpowiedzi lingwistów na pytanie, ile jest języków świata są różne. Sytuacja ta spowodowana jest faktem, że jednoznaczna definicja pojęcia *język naturalny* jest trudna do sformułowania. Nadto, w układanie spisów języków (i odróżnianie ich np. od dialektów) ingerują czynniki pozamerytoryczne — np. polityczne.

Po przyjęciu rozsądnych kryteriów odróżniania poszczególnych języków (etnolektów) można ustalić liczbę istniejących (będących obecnie w użyciu) języków świata na 3500–6500.

Dane na temat języków świata, choć wydawać się mogą olbrzymie ilościowo, są jednak bardzo niekompletne. Dla około 3% (ok. 350) języków istnieją stosunkowo pełne opisy: słowniki jedno- i dwujęzyczne, podręczniki i gramatyki. Częstkowe opisy mamy dla ok. 1300 języków (nie więcej niż 20%), mniej niż 30% (około 2000) języków ma jakąkolwiek dokumentację materiałową. O co najmniej 70% języków wiadomo niewiele ponad to, że w ogóle istnieją.

Około 98% ludności świata posługuje się nieco ponad 2% (około 150) językami. Tak więc, 98% spośród wszystkich języków świata jest używane przez jedynie 2% całej ludności świata. Te dane procentowe mają rzecz jasna charakter rozmyty.

Obecnie 10 języków ma populację liczącą ponad 100 mln użytkowników, zaś około 50 języków — populacje liczące od 10 do 99 mln.

Szacuje się, że w ciągu 175000 lat historii gatunku *homo sapiens* maksymalna liczba języków będących jednocześnie w użyciu wynosiła 20000.

Według prognoz, w ciągu najbliższych dwudziestu lat zniknie ponad połowa będących dziś w użyciu języków.

## 3 Gramatyzacja informacji

Wypowiedzi językowe przekazują różne rodzaje informacji: leksykalną, gramatyczną, emotywną, konatywną, ... W opisie struktury języka szczególna rola przypada informacji *zgramatyzowanej* (lingwiści używają też często terminu *morfo-syntaktyka* języka). W uproszczeniu, informacja gramatyczna ma (wedle wypowiedzi lingwistów) następujące cechy rozpoznawcze:

- jest informacją *obligatoryjnie* wyrażaną w wypowiedziach;
- jest informacją wyrażaną w sposób *regularny*, poprzez kilka – kilkanaście formalnych operacji na znakach;
- z reguły nie występuje samodzielnie (to, nieco mylące sformułowanie oddaje przekonanie lingwistów, że informacja gramatyczna jest niejako dodatkiem do — prymarnej — informacji leksykalnej zawartej w komunikacie; również w pragmatycznych opisach języka do cechy tej podejść trzeba nieco inaczej).

Pierwsze dwie z wymienionych cech wymagają komentarza. W każdym języku pewne aspekty struktury sytuacji, o których mowa w wypowiedziach muszą być formalnie zaznaczone. Np. w polskim osobowe formy czasownika muszą być określone względem czasu gramatycznego oraz liczby, w angielskim lub rumuńskim obligatoryjnie wyrażana jest kategoria określoności, w suahili elementy nominalne niosą wykładniki klas informujące do jakich kategorii ontologicznych odnosi się rzeczownik, w japońskim konieczne jest zaznaczenie tematu, itd. Każdy język ma sobie specyficzny zestaw takich obligatoryjnych rodzajów informacji. Nadto, ta sama kategoria gramatyczna (np. czas) może być różnorodnie reprezentowana w poszczególnych językach (np. uprzedniość – nieuprzedniość w japońskim a przeszłość – teraźniejszość – przyszłość w polskim).

Wyliczenie *wszystkich* rodzajów tak obligatoryjnie wyrażanej informacji w językach świata nie jest — ze względów praktycznych (zob. poprzedni paragraf) — możliwe. Ograniczmy się do podania przykładów kategorii gramatycznych, często wymienianych w tekstach lingwistycznych:

- rodzaj gramatyczny
- klasy nominalne
- przypadek
- osobo-liczba
- sposób istnienia
- klasyfikatory
- określoność
- dostępność poznawcza
- abstrakcyjność
- czas gramatyczny
- aspekt
- modalność
- posesywność
- stopień natężenia cechy
- rodzaj zbioru
- ruch
- rozmiar
- sposób powstania
- stopień afektywności
- wartościowanie.

Regularność wyrażania informacji gramatycznej polega na tym, że każdy z języków wykorzystuje określone techniki do jej zaznaczania. Do najczęściej używanych należą:

- szyk

- afiksacja
- elementy pomocnicze (auxiliaries)
- intonacja
- reduplikacja
- modyfikacja rdzenia
- złożenia
- jakość (artykulacyjna, akustyczna) głosek
- tony.

Jak wiadomo, tzw. teza determinizmu językowego (Sapira-Whorfa) głosi — w uproszczonej postaci — iż język wyznacza obraz świata. Bywa ona powodem do wielu, nieuzasadnionych spekulacji, prowokowanych przez niewłaściwe jej rozumienie (zob. np. dywagacje o „bezczasowości” kultury Hopi, lub „rewelacje” śnieżne Eskimosów).

Wydaje się, że to, co najistotniejsze w tezie determinizmu (i relatywizmu) językowego streścić można krótko: poszczególne języki gramatyzują różne informacje. Może właściwa byłoby w tym miejscu przywołanie analogii z używaniem różnego rodzaju układów współrzędnych (raczej niż różnego rodzaju systemów geometrii).

Stwierdzono, iż każdy przedstawiciel gatunku *homo sapiens* jest w stanie nauczyć się dowolnego języka naturalnego — to, jakiego się nauczy, zależy od otoczenia. Nadto, w każdym języku naturalnym oddać można każde znaczenie (z użyciem odpowiednich środków metajęzykowych oraz peryfrazy) wyrażalne w każdym innym języku. Może to być zachętą do prób konstrukcji *uniwersalnych* „leibnizjańskich” systemów semiotycznych, których języki naturalne byłyby szczególnymi przypadkami (konkretyzacjami). Istotnie, do prób takich zaliczyć można chyba propozycje dot. „języka myśleńskiego” (zob. [Fodor 2001]).

Dodajmy na koniec tego paragrafu kilka ciekawostek pokazujących zróżnicowanie języków świata, jeśli idzie o gramatyzację informacji. Fakty te są „ciekawostkami” z punktu widzenia języków SAE, podobnie jak gramatyka SAE wydawać może się „dziwaczna” z punktu widzenia innych języków.

1. W wielu językach świata występuje konstrukcja zwana *ergatywną*. Jej specyfika polega — w grubym uproszczeniu — na tym, że agens zdania nieprzechodniego i patiens zdania przechodniego mają ten sam wykładnik gramatyczny, różny od wykładnika agensa zdania przechodniego. Zob. np.:

Guugu Yimidhirr

Jerzy-ngun	nganhi	nhaandhi
Jerzy	mnie	widzieć-PAST
Jerzy mnie widział.		
Ngayu	Jerzy	nhaandi
Ja	Jerzy	widzieć-PAST
Widziałem Jerzego.		
Jerzy		dhaadaa
Jerzy	iść-BLISKA PRZYSZŁOŚĆ	
Jerzy zaraz pójdzie.		

2. Mohawk

Agens oraz patiens zaznaczane są w czasowniku, a nie w elementach nominalnych:

Iekša:’a raks:á wahonwá:ienhte’  
girl boy hit  
The girl hit the boy.  
Iekša:’a raks:á wahshakó:ienhte’  
girl boy hit  
The boy hit the girl.

### 3. Mam (zachodnia Guatemala)

Nie ma czasownika *leżeć*. Są natomiast następujące formy:

mutsl he is lying on his stomach  
pak’l he is lying on his back  
tʂaltʂ he is lying on his side  
qinl he is lying outstretched  
leql he is lying sprawled (probably intoxicated)  
kutʂ he is lying alone in a house (probably sick).

Podobnie rzecz się ma — wedle [Aitchison 1996] — z czasownikiem *stać*. Zauważmy, że np. pierwsze trzy z podanych wyżej form leżenia moglibyśmy próbować oddać przez nowotwory językowe: *brzuszkuje*, *pleckuje*, *boczkuje*. Gdyby formy takie przyjęły się w języku, to być może miałyby to konsekwencje dla innych paradygmatów gramatycznych. Oczywiście nie należy uwagi tej traktować poważnie. W tej samej tonacji dodajmy, że przedostatnia z wymienionych form leżenia, a dokładniej przyczyna takiego leżenia może być w polskim oddana przez całe szerokie spektrum form (i to z jednym rdzeniem leksykalnym!).

### 4. Lisu (Birna)

lâma nya ânà kyñ-a  
tigers TOPIC dogs bite

Zdanie to można przetłumaczyć na dwa sposoby: *It’s tigers that bite dogs* lub też *It’s dogs that bite tigers*. W tym przypadku role agensa i patiensa nie są zaznaczone — właściwe znaczenie wynikać musi z kontekstu. W języku lisa zaznacza się patiensa w niektórych zdaniach przez partykułę *lae*. Por. niejednoznaczne polskie: *oskarżenie arcybiskupa okazało się słuszne*.

### 5. Nootka

Następujący przykład może skłaniać do uznania, że w rozpatrywanym języku zatarciu ulega różnica między elementami nominalnymi i werbalnymi:

mamu·k·ma qu·?as·?i  
work PRES man - the  
The man is working.  
qu·?as·ma mamu·k·?i  
man PRES work - the  
The worker is a man.

Dokładniejsza analiza gramatyczna kategorii określoności w tym języku (którą tu pominiemy) zmusza jednak do odrzucenia tej — atrakcyjnej dla poszukiwaczy sensacji — hipotezy.

6. W wielu językach świata spotyka się *klasyfikatory* — wyrażenia odnoszące się np. do kategorii ontologicznej denotacji elementów nominalnych, do ich kształtu, wielkości, dostępności poznawczej, itd. W szczególności, klasyfikatory numeryczne dołączane są obligatoryjnie w konstrukcjach kwantyfikatorskich. Klasyfikatory występować mogą także jako część elementów werbalnych.

## 4 Diachronia

Matematyczne opisy języka naturalnego (języków naturalnych) z reguły dotyczą jego (ich) aspektów synchronicznych — rozważa się język jako twór statyczny i zarazem „semiotycznie domknięty”, całkowicie określony. Tego typu podejście ma uzasadnienie metodologiczne oraz historyczne. Opisy wypracowane dla języków formalnych (siłą rzeczy, synchroniczne!) służą jako aproksymacje języka naturalnego. Nadto, matematyczne opisy języka zaczęły być rozwijane, gdy w językoznawstwie przeważał (synchroniczny!) paradygmat strukturalistyczny (amerykański i europejski).

W ostatnich latach pojawiło się sporo ciekawych prac dotyczących diachronicznych aspektów gramatyzacji informacji w językach świata. Pozwalają one spojrzeć na informację gramatyczną z innej perspektywy; w szczególności, próbują podać (częstkowe przynajmniej) wyjaśnienia, dlaczego dany typ informacji ulega gramatyzacji.

Gramatyzacja informacji dokonuje się poprzez *zmiany językowe*. Jest to proces dość powolny — mierzony w setkach lat. Logika opisującego język naturalny najbardziej chyba interesują zmiany *składniowe* (choć oczywiście przy opisie tych zmian niezbędne jest odwoływanie się do zmian w systemie dźwiękowym, leksykalnym i morfologicznym języka).

Kariere zrobiło *dictum* Bocheńskiego: *Syntax reflects ontology*. Każe ono uważać, że informacja zgramatyzowana oddaje strukturę świata. Nie pozostaje to w sprzeczności z tezami relatywizmu i determinizmu językowego, o ile rozumiane są one w sposób rozsądny (zob. wyżej). Skąd jednak bierze się składnia? I, ogólniej, skąd bierze się informacja zgramatyzowana?

Na temat początków składni można jedynie spekulować. „Horyzont zdarzeń gramatycznych” nie może zostać przekroczony (rozsądnie jest chyba uważać, że w miarę wiarygodne ustalenia dotyczą ostatnich 10000 lat rozwoju gramatycznego języków). Lingwiści zakładają tzw. *uniformitarian principle* — (dogmatyczne) założenie, że prawa zmian językowych (a co za tym idzie, prawa dotyczące gramatyzacji informacji) były zawsze takie same. Ustalanie struktury języka na pewnym stadium jego rozwoju jest więc pewną projekcją biorącą pod uwagę zależności (zmiany) między innymi tego typu stadiami. Popularne wśród typologów jest także hasło *today's syntax is tomorrow's morphology*. Innym dogmatem (dość dobrze potwierdzonym) jest przekonanie o *jednokierunkowości* zmian językowych. Nie należy stwierdzenia tego mylić z ustaleniami dotyczącymi cykliczności pewnych procesów (np. przejść typu: izolujący → aglutynacyjny → fleksyjny → izolujący → ...), ani z obserwacjami dotyczącymi powstawania takich form jak np. *duzen, tutoyer, tykać*.

Lingwistyka historyczna stosuje wiele technik, nie mamy tu miejsca, aby je bardziej szczegółowo omawiać. Metoda filologiczna polega na porównywaniu tekstów (np. dane z okresu 1000 lat). Metoda rekonstrukcji wewnętrznej wykorzystuje ustalenia dotyczące praw zmian językowych (np. fonetycznych). Pomocne bywają inne jeszcze metody (np. badanie języków nowopowstałych — *pidgin, creole*, ustalenia antropologiczne, geograficzne, genetyczne, itd.).

Opisując powody zmian językowych brać należy pod uwagę m.in.: potrzebę jednoznaczności komunikatu, frekwencję jednostek, kolizje znaczeniowe, zasadę najmniejszego wysiłku, itd. Bardzo wiele czynników jest tu istotnych — od klityzacji po sacjację.

Do najważniejszych rodzajów zmian językowych zaliczają lingwiści: metaforyzację, inferencję (konwencjonalizację implikatury), uogólnianie (rozszerzanie znaczenia).

Podajmy, dla ilustracji, garstkę przykładów pokazujących zmiany językowe i związane z nimi procesy gramatyzacji informacji.

1. Tworzenie się kategorii gramatycznych omówione zostało przystępnie np. w pracy [Kuryłowicz 1968]. Rekonstruuje się tam np. tworzenie się liczby gramatycznej: od wyróżniania obiektów na zasadzie styczności (liczba mnoga eliptyczna), przez wyróżnianie obiektów skojarzonych przez styczność i podobieństwo (liczba mnoga kolektywna) do wyróżniania obiektów na podstawie podobieństwa, przynależności do pewnej klasy (liczba mnoga „matematyczna”).

2. Angielskie *will* uzyskiwało status wykładnika informacji gramatycznej przez ok. 1000 lat (od leksykalnego wyrażania chęci, woli do gramatycznego wyrażania czasu). Wyrażenie *going to* (a póź-

niej — *gonna*) funkcjonuje obecnie jako wykładnik bezpośredniej przyszłości; mamy tu do czynienia z deleksykalizacją znaczenia ruchu i kierunkowości.

3. W Tok Pisin wyrażenie *long* (od angielskiego *along*) jest używane jako „uniwersalny” przyimek: na angielski tłumaczone jest przez *in, at, on, to, from, with, by, about, because of, during, for*, zależnie od czasownika występującego w zdaniu.

4. Deleksykalizacja znaczenia widoczna jest np. w angielskich *while, ahead, beside*.

5. Leksykalizacja członu hipotaktycznego widoczna jest np. w polskim *bog daj (by) zdechł → bodaj zdechł*. Dodanie do *bodaj* sufiksu *-że* stwarza formę wyrażającą prawdopodobieństwo.

6. Polskie formy *robiłem, robiłeś*, itd. pochodzą od form *robił jeśm, robił jeś*, itd.

7. W dawnej polszczyźnie używało się jednocześnie form: *temu wydarzeniu minął rok* oraz *dziecku minął rok*. Współcześnie pierwsza z tych form uzyskuje postać *od tego wydarzenia minął rok*, druga trwa nadal.

Gdy ujmujemy język w perspektywie diachronicznej poszczególne znaczenia (leksykalne i gramatyczne) nie są stowarzyszone z pojedynczymi wyrażeniami, ale raczej z przyporządkowaniami stadiom językowym (momentom lub przedziałom czasowym) wyrażen. Jeśli więc opisywać poszczególne stadia języka jako stosowne struktury relacyjne, to język ujmowany diachronicznie staje się czymś w rodzaju produktu kartezjańskiego tych struktur. Nasuwa się oczywiście dalsza analogia — tworzenie ultraprodktu (rolę odpowiedniego ultrafiltru odgrywać mogłyby struktury czasowe związane ze stabilnością informacji).

## 5 Typy języków

Jak podobne do siebie są poszczególne języki? Czy pewne języki różnią się od siebie tak znacząco, że zasadne staje się mówienie o typach języków? Czym w takim przypadku są owe *typy* języków? Oto niektóre z (prostszych) pytań zadawanych w *typologii strukturalnej języków świata*.

Wedle niektórych lingwistów, typy języków to po prostu zbiory złożone z języków o ustalonym zestawie cech. W literaturze przedmiotu znajdujemy także propozycje, aby typy języków traktować jako swego rodzaju ich wewnętrzne własności (*intrinsic properties*), które odkrywać można na drodze badań empirycznych. Dla piszącego te słowa najbardziej atrakcyjna wydaje się propozycja Theo Vennemanna, aby typy języków traktować jako obiekty wyznaczone przez przyjmowaną koncepcję lingwistyczną [Vennemann 1982]. Pewne uogólnienie ujęcia Vennemanna prezentowaliśmy w innym miejscu (zob. np. [Pogonowski 1996]). Postaramy się poniżej, najkrócej jak to możliwe, przypomnieć ten pomysł.

Punktem wyjścia jest zbiór  $G$  obiektów (np. języków, etnolektów, dialektów, stadiów rozwoju języka, itd.), którym przypisać chcemy typy. Traktujemy obiekty z  $G$  jako struktury relacyjne, o których mówić można w stosownym języku pierwszego rzędu  $L$ , będącym językiem (logicznej rekonstrukcji) branej pod uwagę teorii lingwistycznej  $T$  (teoria ta może dotyczyć np. morfologii, składni, zmian językowych, itd.). Budując typologię obiektów z  $G$  lingwiści tworzą pewną teorię  $T^+$ , której stałe pozalogiczne są interpretowane jako pewne relacje na zbiorze  $G$  (np. przekład, zależności diachroniczne, relacje polegające na porównywaniu systemów fonologicznych, morfologicznych, leksykalnych, syntaktycznych, itd.). Przez  $F(n)$  oznaczmy zbiór wszystkich formuł języka teorii  $T^+$  o  $n$  zmiennych wolnych, zaś przez  $TP(n)$  algebrę Boole’a, której elementami są klasy abstrakcji relacji inferencyjnej równoważności formuł z  $F(n)$  względem teorii  $T^+$ . Elementy algebry  $TP(n)$  nazywamy *n-typami*. Pewne zbiory *n*-typów (spełniające proste warunki, których nie ma potrzeby tu szczegółowo omawiać) są *n*-typologiami. Konstrukcja ta stanowi uogólnienie oryginalnej propozycji Vennemanna, w której typami były jedynie formuły teorii  $T^+$  z jedną zmienną wolną. Uogólnienie to pozwala m.in. precyzyjnie scharakteryzować pojęcia *inkorporacji* oraz *polisynntezy*, które nie mogły być opisane w formalizmie Vennemanna.

Zilustrujmy naszkicowaną konstrukcję przykładem. W przyjętej aparaturze pojęciowej dość prosto zrekonstruować można np. — z różnym stopniem mętności podawaną w podręcznikach — kla-

syczną typologię morfologiczną wychodzącą od propozycji Schlegla i Humboldta. Niech mianowicie  $T$  będzie teorią morfologiczną używającą (standardowo rozumianych) parametrów: derywacji, fleksji, syntezy, semazji i fuzji (parametry te są trychotomiczne). Załóżmy, że wśród aksjomatów rozpatrywanej teorii morfologicznej występuje 29 uniwersale Greenberga. W uproszczeniu oznacza to, że jeśli w języku występuje fleksja, to występuje też derywacja. Nazwiemy język *afiksacyjnym*, jeśli nie jest on analityczny (tzn. rdzenie podlegają afiksacji) oraz każda modyfikacja syntetyczna jest afiksacją (a nie np. mutacją tonu lub akcentu, reduplikacją, itp.).

Tradycyjną typologię morfologiczną reprezentują cztery 1-typy:

<b>An</b>	1-typ języków analitycznych
<b>St-I</b>	1-typ języków izolujących tematy
<b>In</b>	1-typ języków fleksyjnych
<b>Ag</b>	1-typ języków aglutynacyjnych

Typy te wyznaczone są przez następujące własności wyrażalne w  $T^+$ :

- **An**: analityczne [rdzenie nie podlegają afiksacji] (np. chiński);
- **St-I**: fleksja stosuje się wyłącznie do tematów oraz każdy wyraz jest tematem (np. samoński);
- **In**: języki z fleksją, które są jednocześnie polisematyczne [nieanalityczne z wieloznaczeniowymi afiksami gramatycznymi] i fuzyjne [nieanalityczne z wewnętrznym sandhi] (np. łacina, hebrajski, niemiecki);
- **Ag**: języki syntetyczne, monosematyczne [nieanalityczne z jednoznaczeniowymi afiksami gramatycznymi], kolokacyjne [nieanalityczne bez wewnętrznego sandhi] i afiksacyjne (np. koreański, turecki, suahili).

Poszczególne języki mogą podpadać pod więcej niż jeden typ w tej typologii. Nadto, względy czysto logiczne (definicyjne) wykluczają większość z  $3^5 = 243$  wszystkich kombinacji wartości branych pod uwagę parametrów. Szczególnie ciekawe wydaje się to, że pewne z tych kombinacji nie są (?) reprezentowane przez żaden język świata. Być może zatem następująca hipoteza stanowić może punkt wyjścia do sformułowania stosownego uniwersale:

*W każdym języku, w którym występuje fleksja i derywacja, sandhi wewnętrzne występuje zawsze łącznie z polisematycznością.*

Trafność tej hipotezy ocenić mogą stosowne badania materiałowe.

Zauważmy jeszcze, że rozważanie  $n$ -typów dla  $n > 1$  pozwala na przypisywanie typów nie tylko poszczególnym obiektom z  $G$ , ale również np. relacjom między tymi obiektami, itd. Nadto, dopuszczenie formuł z kwantyfikatorami wiążącymi zmienne przebiegające zbiór  $G$  umożliwia charakterystykę zjawisk w rodzaju inkorporacji i polisynntezy.

Jako ciekawostkę, przytoczmy kilka przykładów ilustrujących formy polisyntetyczne w językach świata:

1. Eskimo (południowogrenlandzki)

qasuiiΓsaΓBiγcaΓsiΓññitluinaΓnaΓpuq

Es gelang einem überhaupt nicht, einen Rastplatz zu finden.

Całkiem mu się nie udawało znaleźć miejsca, aby odpocząć.



Afiks	Przybliżony odpowiednik polski
qasu	być zmęczonym
iiq	negacja (jakby przyorzecznikowa)
saq	afiks kauzatywny
BBik	miejsce dla ...
caq	afiks powtarzalności
si	znaleźć
ńñit	negacja (jakby przyzdaniowa)
luinaq	całkowicie
naq	tam oto jest
puq	3 Pers. Sing.

## 2. Chinook

	anialot
	dają go jej
a	czas
n	afiks osobowy 1 sing.
i	afiks osobowy 3 sing.
a	afiks osobowy 3 sing.
l	kierunek (do)
o	kierunek (od)
t	afiks oznaczający dawanie

## 3. Eskimoski (jupik)

	kimuxsigjaxkauakuk
	On powinien jechać saniami zaprzężonymi w psy.
kimuxsik	sanie zaprzężone w psy
ua	duratywność
jaxka	konieczność, przeznaczenie do czegoś
kuk	afiks osobowy 3 sing.

## 4. Ajnu (sachaliński)

	aneisantusuixci
	Oni chcą, abyś ty mnie poślubiła.
an	
e	afiks osobowy
i	afiks osobowy 2 sing.
san	poślubić
tusui	chcieć
xci	plur.

## 6 Uniwersalia i inżynieria semantyczna

Jakie cechy strukturalne występują we *wszystkich* językach świata? Czy jakieś kategorie gramatyczne *muszą* być wspólne dla wszystkich języków? Czy występowanie w danym języku pewnych kategorii gramatycznych *wymusza* (lub *wyklucza*) obecność w nim innych kategorii? Te i podobne (nieco bardziej subtelne) pytania stanowią punkt wyjścia do badań *uniwersaliów* językowych.

Wielokrotnie próbowano układać listy absolutnych uniwersaliów językowych (zob. np. [Hockett 1958], [Lyons 1977]). Proponowane cechy miały charakter bardzo ogólny, dotyczyły właściwie systemów semiotycznych w ogólności. Wspomnijmy też, że nie całkiem spełniły się nadzieje związane z tzw. *glottochronologią* oraz *leksykostatystyką* — ustaleniem listy podstawowych słów, występujących we wszystkich językach. Nieco inną perspektywę metodologiczną przyjmuje się w pracach

Wierzbickiej (szukanie tzw. *semantic primitives*, atomów znaczenia uniwersalnych dla wszystkich języków).

Znacząca w badaniu uniwersaliów była praca [Greenberg 1966], poświęcona przede wszystkim syntaksem elementom. Przytoczmy, dla ilustracji, niektóre z propozycji Greenberga:

*Universal 3.* Languages with dominant VSO order are always prepositional.

*Universal 4.* With overwhelmingly greater than chance frequency, languages with normal SOV order are postpositional.

*Universal 5.* If a language has dominant SOV order and the genitive follows the governing noun, then the adjective likewise follows the noun.

*Universal 17.* With overwhelmingly more than chance frequency, languages with dominant order VSO have the adjective after the noun.

Dodajmy jeszcze prozycję z pracy [Aitchison 1996]:

„A possible list of ‘narrow’ absolute universals — features found in all languages — runs as follows.

All languages:

- (1) have consonants and vowels.
- (2) combine sounds into larger units.
- (3) have nouns — words for people and objects.
- (4) have verbs — words for actions.
- (5) can combine words.
- (6) can say what did what to who.
- (7) can negate utterances.
- (8) can ask questions.
- (9) involve structure-dependence.
- (10) involve recursion.

But even these few could be criticized as misleading.”

Wydaje się, że szczególnie mocne charakterystyki języków otrzymać można poprzez ustalanie, jakie cechy *nie mogą* wystąpić w żadnym języku naturalnym. Innymi słowy, chodzi o podanie nie tych cech, które po prostu nie są rejestrowane w żadnym języku (jest ich oczywiście nieskończenie wiele), lecz tych, które są wykluczone — ze względu na pozostałe własności (fonologiczne, leksykalne, morfo-syntaktyczne, składniowe, pragmatyczne, itd.).

Pomocne w ustalaniu takich cech może być stosowanie *inżynierii semantycznej*: przekornego konstruowania potencjalnych języków o zadanych własnościach (np. z wybranymi kategorizacjami rzeczywistości pozajęzykowej odzwierciedlanymi w informacji zgramatyzalizowanej, ustalonymi metodami wyrażania informacji gramatycznej, itp.). Przykład takiej zabawy podany jest w broszurze opisującej wydumany pewnej styczniowej nocy 1982 roku język *ningueño* [Majewicz, Pogonowski 1983a]. Rozważmy kilka dalszych przykładów.

1. W żadnym języku negacja nie jest oddawana przez inwersję szyku elementów. Właściwie trudno wytłumaczyć, dlaczego tak jest. Natomiast w każdym języku negacja występuje i nie wydaje się możliwy system semiotyczny (opisujący rzeczywistość gatunku *homo sapiens*) bez możliwości zaprzeczania wypowiedzi.

2. O ile wiadomo, nie istnieją języki, które w każdym (!) paradymacie fleksyjnym miałyby formę supletywną. Można zastanawiać się, jakie prawa strukturalne wymuszają tę sytuację.

3. Czy mogą istnieć języki pozbawione wyrażen metaforycznych lub idiomatycznych? Jakie są granice nasycenia języka metaforą i idiomami (*wszystkie* wyrażenia nie mogą być metaforami; podobnie, *wszystkie* wyrażenia nie mogą być idiomami).

4. W wielu językach świata nie występuje gramatyczna kategoria (osobo)-liczby. W każdym języku są jednak (o różnym stopniu komplikacji) systemy zaimkowe. Wyobrażenie sobie języka bez

elementów deiktycznych jest możliwe — tematem spekulacji może być nawet ew. pierwotność tego stadium językowego.

## 7 I co dalej?

Czy języki naturalne są jakiegoś rodzaju strukturami algebraicznymi? Czym wtedy (w interpretacji lingwistycznej) są morfizmy określone na takich strukturach?

Jakiego rodzaju strukturą algebraiczną jest *Język Naturalny* — przy założeniu, że obserwowalne różnice między językami etnicznymi są, z należytej perspektywy teoretycznej, do pominięcia?

Odpowiedzi na to pytanie są jednocześnie odpowiedziami na pytanie zadane w tytule tego odczytu. Staje się to oczywiste, gdy uznamy, że liczyć języki chcemy z *dokładnością do izomorfizmu*.

Spójrzmy, jakie konsekwencje mogą mieć konkretne, numeryczne charakterystyki liczby języków:

1. Języków naturalnych jest tyle, ile wyliczają ich lingwiści — np. 5643. Oznacza to, że każdy język naturalny jest strukturą (relacyjną) o indywidualnej sygnaturze. Trochę dziwaczne, prawda? Nadto, grozi bezrobociem lingwistom teoretycznym. . .

2. Języków naturalnych (możliwych) jest  $2^{\aleph_0}$ . Taką odpowiedź uzasadnia np. traktowanie języka jako zbioru zdań (skończonych ciągów morfemów). Czy *Uniwersalny Indeks Języków* może mieć kontinuum pozycji? Kto to przeczyta?!

3. Języków naturalnych (możliwych) jest  $\aleph_0$ . Zgadzamy się na nieskończoność liczby języków, ale niech będzie ona „rozumnie” nieskończona. W szczególności, możemy uznać, że mamy tylko ograniczone możliwości tworzenia teorii lingwistycznych (i to teorii o pewnych wewnętrznych ograniczeniach). Bardzo mętna to odpowiedź (ale spotyka się takie w refleksjach lingwistycznych).

4. Liczba języków naturalnych (możliwych) zależy od akceptowanej teorii lingwistycznej. W szczególności, można szacować liczbę 1-typów języków, otrzymywanych w typologiach bazujących na rozważanej teorii lingwistycznej. Odpowiedź ponętna dla niektórych metodologów.

5. Jest 1 Język Naturalny. Odpowiedź taka zakłada, że potrafimy skonstruować uniwersalną teorię lingwistyczną, której *jedynym* zamierzonym modelem jest Język Naturalny. Znane metatwierdzenia logiczne (twierdzenie Gödla, twierdzenie Löwenheima-Skolema) ukazują pewne ograniczenia, jeśli idzie o istnienie takiego (jedynego z dokładnością do izomorfizmu) modelu. Jednak odpowiedź taka jest obiektem pożądania zarówno niektórych lingwistów jak i filozofów.

## Bibliografia

- Aitchison, J. 1996. *The seeds of speech: language origin and evolution*. Cambridge: Cambridge University Press.
- Allan, K. 1977. Classifiers. *Language* vol. 53, nr 2, 285–311.
- Butterworth, B., Comrie, B., Dahl, Ö. (Eds.). 1984. *Explanations for Language Universals*. Berlin – New York – Amsterdam: Mouton Publishers.
- Bybee, J., Perkins, R., Pagliuca, W. 1994. *The Evolution of Grammar. Tense, Aspect and Modality in the Languages of the world*. Chicago and London: The University of Chicago Press.
- Bybee, J., Fleischman, S. (Eds.). 1995. *Modality in Grammar and Discourse*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Fodor, J.A. 2001. *Eksperci od wiązków. Język myśleński i jego semantyka*. Warszawa: Fundacja Aletheia.
- Givón, T. 1984. *Syntax. A Functional-Typological Introduction. Volume I*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Greenberg, J. 1966. *Language Universals*. Janua Linguarum, Series Minor 59. The Hague: Mouton.
- Hockett, C.F. 1958. *A Course in Modern Linguistics*. New York: The Macmillan Company.
- Hopper, P.J. (Ed.). 1982. *Tense–Aspect: between Semantics & Pragmatics*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Kahrel, P., van der Berg, R. (Eds.). 1994. *Typological Studies in Negation*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Kuryłowicz, J. 1968. *O rozwoju kategorii gramatycznych*. Kraków.
- Lehmann, W.P. (Ed.). 1986. *Language Typology 1985*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Lyons, J. 1977. *Semantics*. Cambridge University Press.
- Maciejewski, W. 1999. *Świat języków*. Wielka Encyklopedia Geografii Świata, tom 14. Poznań: Wydawnictwo Kurpisz.

- Majewicz, A.F. 1989. *Języki świata i ich klasyfikowanie*. Warszawa: Państwowe Wydawnictwo Naukowe.
- Majewicz, A.F., Pogonowski, J. 1983. On categorial marking in natural language. *Lingua Posnaniensis* **XVI**, 55–68.
- Majewicz, A.F., Pogonowski, J. 1983a. *Kilka obserwacji dotyczących struktury języka ɲingueño i konsekwencje dla teorii lingwistycznej*. Poznań: Working Papers of the Institute of Linguistics, Adam Mickiewicz University.
- Pinnow, K.J. 1964. *Die nordamerikanischer Indianersprachen*. Wien: Otto Harrassowitz.
- Pisarkowa, K. 1984. *Historia składni języka polskiego*. Ossolineum.
- Pogonowski, J. 1996. Przyczynek do logicznej rekonstrukcji typologii strukturalnej. In: J. Pogonowski, T. Zgółka (Eds.) *Przyczynki do metodologii lingwistyki*. Poznań: Wydawnictwo Naukowe UAM, 93–108.
- Seiler, H. 1986. *Apprehension. Language, Object and Order. Part III: The Universal Dimension of Apprehension*. Tübingen: Gunter Narr Verlag.
- Shopen, T. (Ed.). 1992. *Language Typology and Syntactic Description. Volume I: Clause Structure*. Cambridge University Press.
- Vennemann, T. 1982. What is linguistic typology? *Studies in Linguistics* **2**, 3–44.
- Whaley, L.J. 1997. *Introduction to Typology. The Unity and Diversity of Language*. Thousand Oaks – London – New Delhi: SAGE Publications.