

Schwyzertüütsch, bambara i języki bezkontekstowe

Noam Chomsky w swoim znanym artykule (Chomsky, 1959) zaproponował hierarchię gramatyk formalnych rozumianych jako zbiory reguł przepisywania ciągów znaków terminalnych i nieterminalnych na inne ciągi znaków terminalnych i nieterminalnych. Dziś powszechnie znana i stosowana jest pod nazwą hierarchii Chomsky'ego.

Co prawda autor, jak sam twierdzi, opisuje w ten sposób jedynie języki formalne, jednak na marginesie rozważań sformułował problem, który dotyczy języków naturalnych: do jakiej klasy gramatyk formalnych należeć będą gramatyki adekwatnie opisujące języki naturalne? Sam nie udzielił na to pytanie odpowiedzi, ale wzbudził nim długą dyskusję filozoficzną i językoznawczą na temat złożoności konstrukcji w różnych językach naturalnych.

Już na samym początku Chomsky odrzucił dwie klasy swojej hierarchii, jako nieadekwatne do opisu gramatyk języków naturalnych: gramatyki typu 0 (gramatyki nieograniczone, generujące języki rekurencyjnie przeliczalne) oraz gramatyki typu 3 (gramatyki regularne). Pierwszą z klas uznał za zbyt szeroką — informacja, że języki naturalne można opisywać za pomocą gramatyk nieograniczonych, które potrafią generować dowolny język formalny, przy jednoczesnym założeniu, że adekwatnego opisu języków naturalnych można dokonać metodami czysto formalnymi (a takie założenie niewątpliwie Chomsky czyni), jest właściwie bezużyteczna. Drugą z klas — gramatyki regularne — Chomsky odrzuca jako mechanizm zbyt słaby, na co wskazywał również w napisanej wcześniej, ale opublikowanej dopiero 1957 roku książce *Syntactic Structures* (Chomsky, 1957). Zatem postawiony przez Chomsky'ego problem dotyczy tylko dwóch klas: gramatyk kontekstowych i gramatyk bezkontekstowych.

Istnieją co najmniej dwa powody, dla których pytanie, czy gramatyki bezkontekstowe są wystarczającym mechanizmem do opisanie wszystkich języków naturalnych, jest interesujące. Pierwszy powód jest czysto praktyczny: w przypadku ogólnym problem rozpoznawania (i parsowania) języków bezkontekstowych jest wielomianowy, czyli obliczalny w sensie praktycznym. W przypadku języków kontekstowych jest to jednak problem obliczeniowo trudny. W obliczu wciąż rosnących potrzeb automatycznego przetwarzania tekstów i nagrań w językach naturalnych dobrze byłoby wiedzieć, czy nasze możliwości obliczeniowe wystarczają do poradzenia sobie z dowolnymi językami naturalnymi.

Drugim powodem są ogólne pytania teoretyczne o związki pomiędzy tradycyjnym lingwistycznym opisem języków, a ich opisem formalnym. Wiedząc jakie własności czysto formalne mają języki naturalne, można próbować orzekać o adekwatności różnych ich opisów gramatycznych.

W wystąpieniu chciałbym zaprezentować wybrane argumenty na poparcie tezy o niebezkontekstowości języków naturalnych, które były przytaczane w trakcie dyskusji. Najbardziej znany z tych argumentów pochodzi od

Stuarta Shiebera (Shieber, 1985). Autor rozważa przykłady zdań ze szwajcarskiej odmiany języka niemieckiego, których konstrukcja nie pozwala na reprezentowanie ich za pomocą gramatyki bezkontekstowej. Zaprezentuję też inne tego typu argumenty — dotyczące m.in. składni języka niderlandzkiego (Manaster-Ramer, 1987) oraz morfologii malijskiego języka bambara (Culy, 1985). Chciałbym się również przyjrzeć filozoficznym konsekwencjom, jakie niosą ze sobą zarówno te argumenty, jak i wysuwane przeciwko nim kontrargumenty — przekładają się one bowiem na to, jak przez poszczególnych autorów rozumiany jest język naturalny oraz językoznawstwo jako dziedzina służąca jego badaniu.

Literatura

- Chomsky N., *Three models for the description of language*. w: *IRE Transactions on Information Theory*, s. 113–123, 1956.
- Chomsky N., *Syntactic Structures*. Mouton & Co., Publishers 1957.
- Chomsky N., *On certain formal properties of grammars*. *Information and Control*, t. 2 s. 137–167 1959.
- Culy C., *The complexity of the vocabulary of Bambara*. *Linguistics and Philosophy*, t. 8 s. 345–351 1985.
- Joshi A., *How much context-sensitivity is necessary for characterizing structural descriptions*. w: *Natural language processing: Theoretical, computational and psychological perspectives*, pod red. D. Dowty, L. Karttunen i Zwicky, New York 1985.
- Kac M.B., Manaster-Ramer A. i Rounds W.C., *Simultaneous-distributive coordination and context-freeness*. *Computational Linguistics*, t. 13 (1–2) s. 25–30 1987.
- Manaster-Ramer A., *Dutch as a formal language*. *Linguistics and Philosophy*, t. 10 s. 221–246 1987.
- Mecner P., *Elementy gramatyki umysłu*. Kraków 2005.
- Pullum G. i Gazdar G., *Natural languages and context-free languages*. *Linguistics and Philosophy*, t. 4 s. 471–504 1982.
- Shieber S., *Evidence against the context-freeness of natural language*. *Linguistics and Philosophy*, t. 8 s. 333–343 1985.
- Weir D.J., *A geometric hierarchy beyond context-free languages*. *Journal of Computer and System Sciences*, t. 104 (2) s. 235–261 1992.