

Posłowie redaktora

Kiedy 10 lipca 1978 r. odbywała się obrona pracy doktorskiej Szpakowicza, wydawało się, że praca ta stanowi tylko pierwszy krok do formalnego opisu składniowego języka polskiego i wkrótce zostanie zaktualizowana nowszymi opracowaniami. Z różnych względów tak się jednak nie stało. Przez wiele lat nie miała ona żadnej konkurencji, co uzasadniało dwa wydania jej wersji książkowej ([13], [14]), obecnie dostępnej także w formie elektronicznej. Kiedy pojawiła się obszerna gramatyka formalna Świdzińskiego ([15], [16]), charakteryzowała się ona nie tylko dużą złożonością, ale i programowym ignorowaniem obliczeniowych aspektów gramatyki. W tej sytuacji praca Szpakowicza nadal posiada m.in. duże walory dydaktyczne, i to stanowi główny powód opracowania jej elektronicznego wydania w ponad dwadzieścia lat od pojawienia się oryginału.

Niniejszy tekst traktuję jako wersję wstępną posłowia. Jeśli pozwoli na to czas, będę starał się ją uzupełnić o dodatkowe informacje i w szczególności uwzględnić ewentualne uwagi czytelników wersji elektronicznej.

1 Konwencje notacyjne

Oryginał pracy doktorskiej Szpakowicza został sporządzony na maszynie do pisania nie dysponującej nawiasami okrągłymi (zamiast nich stosowany był ukośnik), dostępny był też tylko jeden rodzaj cudzysłowu. W wydaniu elektronicznym pracy staraliśmy się stosować analogiczne konwencje, jak we wspomnianym wcześniej elektronicznym wydaniu książki. Oryginalna książka została wydana techniką „małej poligrafii” istotnie ograniczającą możliwość stosowania wyróżnień typograficznych, które były zastąpione różnego rodzaju podkreśleniami, a np. jeden znak cudzysłowu (") występował w kilku funkcjach. Wyróżnienia te staraliśmy się w wydaniu elektronicznym oddać w sposób jak najbardziej czytelny i zgodny z normami i zwyczajami typograficznymi. W szczególności podkreślenia potraktowaliśmy zgodnie z ich znaczeniem jako znaków adiustacyjnych: linię falistą oddajemy kursywą (z wyjątkiem opisanym niżej), zaś linię ciągłą wytłuszczeniem; bardziej zgodne z polską tradycją byłoby stosowanie spacjowania zamiast kursywy, ale niestety byłoby to trudniejsze technicznie.

W specjalny sposób zostały potraktowane cudzysłowy wykorzystywane do cytowania przykładowych zdań i ich fragmentów. Choć w typowych publikacjach lingwistycznych w takich sytuacjach nie stosuje się w ogóle cudzysłowów, a przykłady wyróżnia kursywą, tutaj przykłady ujmujemy w tzw. cudzysłowy francuskie

czyli „łapki”. Dzięki temu możliwe jest czytelne cytowanie nawet takich fragmentów zdań, jak pojedyncze znaki interpunkcyjne, np. «,» na s. 23.

Zdarza się dość często, że cytowane przykłady zawierają pewne elementy metajęzykowe, tzn. zawierają znaki lub słowa, które w istocie nie należą do cytowanej wypowiedzi. Wyrazistym przykładem są znaki, które nie są składnikami cytowanego tekstu, lecz służą do skróconego zapisu kilku wariantów przykładów lub stanowią jego rozszerzenie, np. «*wiem o [czymś]*» na s. 3. Jak widać, właściwy przykład złożony jest kursywą, a metajęzykowe nawiasy antykwą; aby jeszcze bardziej zwiększyć czytelność takich przykładów, występujące w oryginale ukośniki zastąpiliśmy nawiasami prostokątnymi.

Stosowanie antykwy (bezszerzyfowej) w przykładach rozszerzyliśmy również na inne sytuacje, kiedy cytowany napis nie może być mechanicznie traktowany jako fragment autentycznej wypowiedzi — stąd na s. 86 mamy «-śmy», ale «*gdzibyśmy*», na s. 5 mamy «-no» i «-to» (choć «*czytano*» i «*pito*»).

Już tylko dla zasady postanowiliśmy odróżniać rzeczywiste wystąpienie wielokropka (w praktyce nie występujące w zawartych w książce przykładach) od wielokropka metajęzykowego oznaczającego pominięcie fragmentu przykładu, np. «*wiem, co [. . .]*» na s. 23. Alternatywą dla tej konwencji mogło by być pisanie wielokropka tekstowego — zgodnie z przyjętym zwyczajem — bez odstępów od poprzedzającego go słowa, a wielokropka metajęzykowego — z odstępem; nie wydaje się jednak, aby tak subtelna konwencja była wygodna w praktyce.

W oryginale fragmenty niektórych przykładów są wyróżniane podkreśleniem. Choć w zasadzie w druku nie stosuje się podkreśleń, postanowiliśmy utrzymać tę konwencję. Konwencja ta pozwala nie tylko wyróżnić pojedynczy znak interpunkcyjny, jak w przykładzie «*najlepiej, najskuteczniej i najefektowniej*»^{kps2} na s. 60, ale także — jak czyni to Autor — np. zaznaczyć granice między wyróżnionymi składnikami, jak w przykładzie «*daję mu książkę*» na s. 3.

Pewnym problemem były dla nas takie fragmenty oryginału, gdzie ujęty w cudzysłowy napis jest w mniej lub bardziej oczywisty sposób reprezentacją pewnej abstrakcyjnej jednostki czy własności. Najczęściej traktujemy go wtedy tak samo, jak omawiane wyżej elementy metajęzykowe, np. czasownik posiłkowy «*będę*» na s. 5.

Pozostałych wypadków użycia cudzysłowu nie traktowaliśmy z absolutnym pietyzmem. Cudzysłowy pozostawialiśmy zawsze wtedy, gdy dane słowo było użyte w przenośnym znaczeniu, ale tam, gdzie wydawało nam się to stosowne, zastępowaliśmy użycie cudzysłowu kursywą.

Wspomniane wyżej konwencje notacyjne dla reguł gramatycznych są realizowane za pomocą programu *pretprint* autorstwa Marcina Wolińskiego (patrz [17], [18] i [19]), aktualnie dostępnego m.in. w archiwach Grupy Użytkowników Systemu $\text{T}_{\text{E}}\text{X}$ (patrz www.gust.org.pl). Stosowanie tego programu pozwoliło w automatyczny sposób zbudować skorowidz symboli.

Do zapisu reguł stosujemy pismo bezszerzyfowe. Wartości ustalone (zarówno cytowane w tekście, jak i występujące w regułach) oraz nazwy warunków zapisywane są majuskułami zwykłej antykwy, symbole nieterminalne — majuskułami

antykwy wytluszczonej, a zmienne — minuskułami kursywy. Dodane w wydaniu elektronicznym symbole reguł drukowane są przy prawym marginesie na wysokości pierwszego wiersza odpowiedniej reguły. Symbole te zostały dodane do większości przykładów w postaci indeksów górnych.

Podane w pracy programy Prologowych stosują tzw. składnię marsylską, która całkowicie wyszła z użycia. Obecnie rozpowszechniona jest składnia nazywana edynburską — por. np. [6]. W składni tej przykłady ze s. 75 mają postać następującą:

```
suma(X, 0, X).
suma(X, s(Y), s(Z)) :- suma(X, Y, Z).
:-suma(s(0), s(0), X).

iloczyn(X, 0, 0).
iloczyn(X, s(Y), Z) :- iloczyn(X, Y, T), suma(T, X, Z).
```

Na zakończenie tego omówienia spraw notacyjnych chciałbym wspomnieć, że skrót *p.* oznacza zawsze *punkt*.

2 Terminologia

Warto zwrócić uwagę na to, że Szpakowicz stosuje termin *rodzaj męskożywotny* (por. np. s. 1) w specyficznym znaczeniu wprowadzonym (jak się wydaje, przeze mnie) na potrzeby systemu MARYSIA — patrz [1], (Łukaszewicz, Szpakowicz 1973) — które jednak nie okazało się wygodne. W użyciu jest obecnie terminologia, zgodnie z którą rodzaj męski dzieli się na rodzaj męskoosobowy, męskozwierzęcy i męskorzeczowy, zaś przez rodzaj męskożywotny tradycyjnie rozumie się rodzaj męskoosobowy i męskozwierzęcy łącznie.

Inna wykorzystana przez Szpakowicza moja propozycja terminologiczna z tego okresu, mianowicie *bezosobnik* w znaczeniu *forma bezosobowa czasownika zakończona na -no, -to* bardzo wcześnie została zaakceptowana m.in. przez Andrzeja Bogusławskiego i Zygmunta Saloniego; jest ona obecnie używana m.in. w podręcznikach akademickich takich jak [10].

Bardzo udaną propozycją terminologiczną Szpakowicza był *rodzaj przymnogi*; termin ten jest coraz częściej używany, choć nie zawsze w identycznym znaczeniu.

Pozostałe nietradycyjne stosowane w pracy terminy raczej się nie przyjęły, stąd dla wielu lingwistów są one nieznane i niezrozumiałe.

3 Poprawki

W oryginale reguły opisujące prostą frazę czasownikową miały postać następującą:

```
FCZ1W(OSOB, r, l, o, OZN, PRZY, TAK)
= CZPRZYSZ(r, l, o, TAK) . (fczpw1)
FCZ1W(OSOB, r, l, o, OZN, PRZY, NIE)
= PART(NIE) CZPRZYSZ(r, l, o, NIE) . (fczpw2s)
```

FCZ1W(*wd, r, l, o, tr, cz, neg*)
 = **KCZ**(*wd, p, r, l, o, tr, cz, neg*) –ALT(*wd, OSOB.BEZOS*) . (fczpw3)

Jak zauważył p. Adam Wachowski, powoduje to, że poprawne zdanie «*Ojciec nie będzie płakał.*» analizuje się tylko dlatego, że gramatyka dopuszcza pomijanie przecinków i w konsekwencji przytoczony przykład równoważny jest zdaniu szeregowemu «*Ojciec nie będzie, płakał.*». Aby umożliwić analizę tego zdania w sposób zgodny z intuicją, regułę drugą zastąpiliśmy przez

FCZ1W(*OSOB, r, l, o, OZN, PRZY, NIE*)
 = **PART**(*NIE*) **CZPRZYSZ**(*r, l, o, neg*) . (fczpw2)

Jest to jedyna wprowadzona wprowadzona poprawka merytoryczna, pozostałe zmiany dotyczą oczywistych błędów literowych.

4 Program analizy syntaktycznej

Przedstawiony w pracy program zachował się na nośniku (taśmie magnetycznej) wraz z niezbędnym do jego wykonania interpreterem języka Prolog. Pliki te od dłuższego czasu są dostępne w Internecie (pod adresem <ftp://ftp.mimuw.edu.pl/pub/users/jsbien/mainframe>), a ostatnio te historyczne interpretery Prologu zostały przeniesione na PC i po odpowiednim udokumentowaniu zostaną również udostępnione zainteresowanym.

5 Ortografia

Kiedy ś.p. Jan Tokarski, wówczas docent, przeglądał wstępną wersję jednego z moich pierwszych artykułów lingwistycznych, zwrócił mi uwagę, że pisownia słowa (*liczba*) *pojedyncza* przez *n* z *kreską* jest błędna. Stwierdził wówczas, że ta niczym nie uzasadniona anomalia wzięła się z błędu literowego w słowniku ortograficznym, który nie zauważony w porę został uznany za normę.

W pracy doktorskiej Szpakowicz konsekwentnie stosował pisownię przez *n* z *kreską*. Wszystkie te przypadki zostały dostosowane do oficjalnej pisowni słowa *pojedynczy* pomimo naszego niechętnego do niej stosunku.

6 Bibliografia uzupełniająca

- (1) Bień, J.S. Towards computer systems for conversing in Polish. In Antonio Zampolli and Nicoletta Calzolari, editors, *Computational and Mathematical Linguistics*, pages 139–159, Firenze, 1980. Proceedings of the International Conference on Computational Linguistics, Pisa, 27/VIII–1/IX 1973, Leo S. Olschki.
- (2) Colmerauer A.: *Metamorphosis grammars*. In L. Bolc (ed.): *Natural Language Communication with Computers. Lecture Notes in Computer Science 63*. Springer-Verlag 1978, s. 133-189.
- (3) M. van Emden, R. Kowalski, *The semantics of predicate logic as a programming language*. Journal of ACM 23, , 4, s. 733-744.

- (4) Z. Grodzki, Z problemów maszynowego przetwarzania tekstów polskich. Fleksja liczebników. *Poradnik Językowy* nr 1, 1972.
- (5) Kluźniak F., Szpakowicz S.: *Prolog*. WNT, Warszawa, 1983.
- (6) Kluźniak F., Szpakowicz S.: *Prolog for Programmers*. Academic Press 1985, 1987.
- (7) R. Kowalski, *Predicate logic as programming language*. Proc. IFIP Congress 1974, North-Holland 1974, pp 569-574.
- (8) Kowalski R.A.: *Logic for problem solving*. North-Holland 1979.
- (9) Kowalski R.A.: *Logika w rozwiązywaniu zadań* WNT 1989.
- (10) Z. Saloni, M. Świdziński. *Składnia współczesnego języka polskiego*. Wydanie czwarte, zmienione. Warszawa: Wydawnictwo Naukowe PWN SA 1998.
- (11) S. Szpakowicz, *Syntactic analysis of written Polish*. W: *Natural Language Communications with Computers*, Lecture Notes in Computer Science 63. Springer-Verlag 1978, pp 261-292.
- (12) S. Szpakowicz, *Syntactic analysis of written Polish*. W: *Natural Language Communications with Computers*, Lecture Notes in Computer Science 63. Springer-Verlag 1978, pp 261-292.
- (13) S. Szpakowicz, *Formalny opis składniowy zdań polskich*. Wydawnictwa UW: Warszawa 1983.
- (14) S. Szpakowicz, *Formalny opis składniowy zdań polskich*. Wydanie drugie. Wydawnictwa UW: Warszawa 1986.
- (15) M. Świdziński. Formalny opis polskich zdań o składniku zdaniowym. Praca habilitacyjna. Maszynopis powielony, Instytut Języka Polskiego UW, Warszawa 1987.
- (16) M. Świdziński. *Gramatyka formalna języka polskiego*. Wydawnictwa UW: Warszawa 1992.
- (17) Woliński, M. 1996. Zgrabne formatowanie tekstów programów komputerowych przy użyciu $\text{T}_{\text{E}}\text{X}$ -a. Biuletyn GUST nr 7, 1996, s. 38–40.
- (18) Woliński, M. 1998. Automatyczne formatowanie strukturalne tekstów w językach formalnych. Praca magisterska (opiekun J. S. Bień), Instytut Informatyki UW 1998.
- (19) Woliński, M. 1998a. PrePrint: a $\text{L}^{\text{A}}\text{T}_{\text{E}}\text{X}$ 2 ϵ package for prettyprinting texts in formal languages. In: TUG '98. Proceedings of the $\text{T}_{\text{E}}\text{X}$ Users Group Meeting, August 17–20, 1998, Toruń, Poland, pp 43–48.