

Zygmunt SALONI

Warszawa

O KONWENCJONALNYCH ZNAKACH POMOCNICZYCH W TEKŚCIE POLSKIM

Problem dyskutowany w niniejszym artykule wyłonił się w trakcie prac nad *Słownikiem gramatycznym języka polskiego*¹ (dalej: *SGJP*). Autorzy musieli zdecydować, jaką postać mają słowa, które uznawali za wykładniki polskich form wyrazowych. Rozstrzygnięcia, które przedstawiam niżej, nie były oczywiste.

Polski tekst pisany stanowi ciąg znaków piśmiennych (nie uwzględniamy wyróżnień typograficznych). Wśród tych znaków możemy wyróżnić dwie ważne klasy. Pierwsza – to litery alfabetu polskiego. Druga – to separatory, czyli znaki ograniczające słowa polskie z obu stron: znaki interpunkcyjne i odstęp międzywyrazowy (w tekstach elektronicznych kodowany m.in. za pomocą znaku spacji).

Fizycznie istniejący tekst ma naturę mocno złożoną, heterogeniczną (ograniczamy się tu do właściwego tekstu, pomijając elementy nietekstowe, które mogą obok niego wystąpić, zwłaszcza ilustracje). Mogą się w nim pojawić elementy inne niż polskie słowa. Takich elementów nie opisujemy.

Jednak w polskim tekście, a w szczególności w polskich słowach, mogą wystąpić inne znaki, które nie są ani znakami interpunkcyjnymi, ani literami polskimi. Są to przede wszystkim znaki należące do innych systemów znakowych.

Elementy kodów jawnie obcych muszą być uznane za cytaty w innych językach, są one skreślane przy jego opracowywaniu. Tak zostaną potraktowane napisy takie, jak słowa greckie, rosyjskie czy hebrajskie zapisane w oryginalnych alfabetach, a także wszelkie wstawki zawierające cyfry arabskie (przede wszystkim cyfrowe zapisy liczb) i inne symbole specjalne (np. @, & czy %). Skutkiem tej decyzji jako słowa polskie zostają zdyskwalifikowane jednostki, które intuicyjnie są tak traktowane i które wielokrotnie wystąpiły w zbadanych polskich tekstach, np. *F-16*, *A4*. Jednak ich struktura jest wyraźnie różna od typowych słów polskich:

¹ Z. Saloni, W. Gruszczyński, M. Woliński, R. Wołosz, *Słownik gramatyczny języka polskiego* Wiedza Powszechna, Warszawa 2007.

dopuszczenie ich spowodowałoby konieczność rozszerzenia znaków elementarnych alfabetu polskiego o jednostki o innym statusie niż jednostki podstawowe – litery. Na tej samej podstawie jako słowa polskie zdyskwalifikujemy na przykład adresy internetowe, w których skład wchodzi jako element znaczący znak obcy – @ (średniowieczna ligatura *ad*, która przetrwała w języku angielskim, nazywana w normach międzynarodowych i standardzie Unicode *commercial at*).

Kryterium wystarczającym do usunięcia jednostki ze zbioru jest bowiem występowanie w niej symboli nie należących do polskiego alfabetu² ani do zbioru znaków nieliterowych, których występowanie dopuszczamy w słowach polskich. Są bowiem znaki, które musimy dopuścić w polskim tekście. Oprócz wskazanych wyżej separatorów, które nie nastroczają trudności, ponieważ nie wchodzi w skład słów, są jeszcze inne znaki. Nazwiemy je **konwencjonalnymi znakami pomocniczymi**. Jest ich cztery.

1. Najmniej kłopotów sprawia **apostrof**, który jest wyraźnie jednofunkcyjny; występuje wewnątrz słów polskich, z reguły pochodzenia obcego: rozdziela je ortograficznie na dwie części, stanowi nieformalną wskazówkę, że dane słowo ma być zrealizowane fonetycznie w sposób nietypowy. Apostrof pisze się najczęściej w formach przypadków zależnych nazw obcych, jak np. *Kennedy’ego*, *Beaumar-chais’go*, albo nawet rzeczowników pospolitych, jak np. *college’u* – wprowadzenie ich do słownika gramatycznego języka polskiego było rozwiązaniem naturalnym. Ponadto są leksemy zapożyczone z języków obcych wraz z apostrofem, jak *dell’arte*. One także znalazły się w SGJP.

2. Natomiast drugi znak pomocniczy używany w piśmie polskim jest wielofunkcyjny. To **/: kreska ukośna**, zwana też potocznie ukośnikiem lub z angielska slashem. W pewnych użyciach może być ona traktowana jako nowy, niekanoniczny znak przestankowy, choć często jest z obu stron dosunięta do tekstu literowego, np. *jego/jej*, *i/lub*. Takich napisów zawierających kreskę ukośną nie będziemy traktować jako wyrazów polskich, słowami polskimi mogą być natomiast napisy występujące po obu jej stronach, kreska występuje więc jako separator. Jednak ten sam znak ma jeszcze inną funkcję: jest nieusuwalnym elementem składowym skrótów, które pojawiają się w tekstach polskich na tyle często, że mogą być uznane za należące do języka polskiego. Takie są skróty typów statków – z pochodzenia angielskie, np. *M/S* ‘motor ship, statek motorowy’, a także nie

² Do alfabetu polskiego zaliczamy wszystkie litery występujące we właściwych słowach polskich (według wykazu podawanego w materiałach dla szkoły podstawowej) oraz trzy litery podstawowego alfabetu łacińskiego, które mogą być użyte w słowach zapożyczonych z języków obcych: *q*, *v* i *x*. Jako wyjątki dopuścimy też w SGJP warianty liter polskich, różniące się od liter podstawowych znakiem diakrytycznym, np. *é* czy *ü*.

uznawane za poprawne skróty pewnych pomocniczych wyrażań polskich np. *d/s* ‘do spraw’ – takie skróty włączamy do SGJP na zasadzie wyjątków na podstawie arbitralnych decyzji indywidualnych.

Jako znak pomocniczy kreska ukośna występuje wewnątrz słowa, dosunięta z obu stron do sąsiedniej litery. Jest to wyraźna różnica w stosunku do użycia tego znaku w jeszcze jednej funkcji: nawiasu (zarówno otwierającego, jak i zamykającego) – wtedy z jednej strony sąsiaduje z innym separatorem, najczęściej spacją. W ten sposób tradycyjnie symbolizowano nawias w tekstach pisanych po polsku na mechanicznej maszynie do pisania, która nie zawierała znaków nawiasów. Zaskakująco dużo takich użyć kreski ukośnej znalazło się w naszym materiale tekstów zakodowanych w postaci elektronicznej.

3. Jako znak pomocniczy traktowany jest też w polskich przepisach ortograficznych **łącznik**: -³.

Tymczasem funkcja jego jest skomplikowana. Występuje wewnątrz napisów złożonych z liter, dzieląc je na dwie części (czy raczej łącząc je w jedną całość). Uznajemy go na ogół za separator form wyrazowych, zwłaszcza wtedy, kiedy każdy z członów złożenia daje się łatwo zinterpretować jako forma leksemu należącego do języka polskiego. Takie są nazwy własne typu *Bielsko-Biała* i dwuczłonowe nazwiska, np. *Pawlikowska-Jasnorzewska*. Także słowa tekstowe typu *polsko-radziecka* i *marksistowsko-leninowski* uznajemy za złożone z dwu form wyrazowych (pierwsze człony złożań *polsko+* i *marksistowsko+* traktujemy jako formy przymiotników i zaliczamy odpowiednio do leksemów **polski** i **marksistowski** – znak plusa wskazuje, że słowo nie występuje samodzielnie). Co więcej, w funkcji separatora łącznik może wystąpić także na końcu słowa: w wypadkach, gdy współrzędnie zestawione są pierwsze człony złożań, np. *pierwszo- i drugoklasiści*, *chłopi mało- i średniorolni*, *pokoje cztero- i pięcioosobowe*. Są one pochodne od przymiotników i liczebników (jako formy liczebników traktujemy zatem również słowa *cztero+* i *pięcio+*).

Jednak w istniejących słownikach są zanotowane napisy z łącznikiem wewnętrznym, dla których znacznie naturalniejsza jest interpretacja całościowa, jak *pop-art* czy *ping-pong*: traktujemy je jako polskie słowa. Inne podobne napisy powtarzają się też dostatecznie często w tekstach, np.: *W-wa* ‘Warszawa’ albo *b-cia* ‘bracia’ – zdecydowaliśmy się potraktować je tak samo: jako polskie słowa. Jakie elementy z łącznikiem wewnętrznym zostają włączone do SGJP (wśród napisów zawierających łącznik są też takie, które wywołują dyskusje, czy uznać je za jednolite słowa czy złożenia dwu słów, jak np. *figi-migi*, *figle-migle*), zdecydowaliśmy

³ Bywa on mylony z pauzą, czyli myślnikiem: – lub —. Kwestią tą nie będziemy się tutaj jednak zajmować.

arbitralnie na podstawie naszej intuicji językowej podbudowanej danymi z korpusów⁴.

4. Wiele współczesnych wydawnictw poprawnościowych traktuje **kropkę** jednoznacznie tylko jako znak przestankowy⁵. Jednak dla każdego, kto pracował nad zakodowanym tekstem polskim⁶, a nawet dla tego, kto starał się – z jakiegokolwiek strony – dokładniej zrozumieć problem, jest jasne, że fizyczna kropka – znak graficzny – ma w tekście polskim (i tekstach języków stosujących podobny zapis) dwie jaskrawo różne funkcje: znaku przestankowego wskazującego na określoną strukturę tekstu, dzielącego go na mniejsze całości, oraz znaku graficznego służącego do sygnalizowania, że część wyrazu, najczęściej końcowa, została pominięta. Kropka w funkcji pierwszej będzie oczywiście traktowana jako separator, nie należący do poprzedzającego go słowa; kropka w funkcji drugiej powinna już być uznana za elementarny znak pomocniczy wchodzący w skład słowa polskiego, ale z bardzo silnymi ograniczeniami kontekstowymi.

W pierwszej chwili bylibyśmy skłonni podejrzewać, że kropka jako znak pomocniczy występuje tylko na końcu słów polskich⁷. Tak jednak nie jest: może wystąpić także wewnątrz słowa. Istnieją bowiem utarte skróty wyrażen złożonych z kilku słów, w których kropkę pisze się bezpośrednio przed literą, np.: *m.in.*, *p.n.e.*, *n.p.m.* Dopuszczymy więc w tekście polskim pozycje kropki między literami polskimi. Skróty zawierające kropkę, podobnie jak inne ustabilizowane skróty polskie, wprowadzamy do SGJP arbitralnie – na podstawie naszego doświadczenia językowego.

Powyższa analiza pokazała, że konwencjonalne znaki nieliterowe muszą być uwzględniane w opisach tekstu polskiego, sporządzanego zarówno w celach leksykograficznych, jak i dydaktycznych. W *Słowniku gramatycznym języka polskiego* bezpośrednio po alfabecie polskim podaliśmy zatem: *Znaki pomocnicze (ignorowane przy wyszukiwaniu): . ' – /*

⁴ Dodatkowy problem sprawiają napisy typu *XIX-wieczna*. Omówione tutaj konwencje prowadzą do uznania ich za złożenia, których druga część jest niewątpliwie słowem polskim, status zaś pierwszej części jest niejasny. W tym wypadku podejmowaliśmy decyzje arbitralnie i liczb rzymskich, w których symbolami elementarnymi są litery łacińskie (w konwencji polskiej – wielkie), nie uznajemy za słowa polskie (mogłyby być uznane za skróty).

⁵ Na przykład NSO PWN (1996) obie funkcje kropki omawia w rozdziale *Znaki interpunkcyjne* (s. XCVII–C). Rozgraniczenie widać jednak w mającej moc obowiązującą *Pisowni polskiej* (1957), która rozgranicza użycie kropki jako znaku przestankowego (w części *Interpunkcja*, s. 60–62) i jako znaku pomocniczego (w części *Inne kwestie graficzne*, s. 35–37).

⁶ Por. np. M. Rudolf, *Metody automatycznej analizy tekstów polskich*, Warszawa 2004.

⁷ Oczywiście przy analizie tekstu polskiego musi być uwzględniona reguła, że w wypadku wystąpienia na końcu zdania – przed kropką wymaganą interpunkcyjnie – skrótu kończącego się kropką, kropkę – znak graficzny stawiamy tylko jedną.

ON CONVENTIONAL AUXILIARY CHARACTERS IN POLISH TEXTS

S u m m a r y

The aim of this paper is to discuss the usage of four characters that can be met in Polish texts in functions different than letters and punctuation marks: apostrophe, hyphen, slash and dot (the two last characters are also used as punctuation marks). They occur in words introduced to *Grammatical Dictionary of Polish* by Zygmunt Saloni, Włodzimierz Gruszczyński, Marcin Woliński, and Robert Wołosz, published in 2007. The conventional auxiliary characters have been neglected in alphabetical order and are not taken into account in searches.