

Janusz S. Bień

Warszawa

## Traktat Parkosza. Eksperymentalna edycja elektroniczna

### 1. Wstęp

Jak wiadomo, żyjący w pierwszej połowie XV wieku Jakub Parkosz (podaję tu nazwisko w formie stosowanej przez Mariana Kucałę, inne formy to m.in. *Par-koszowic*) – profesor i rektor Akademii Krakowskiej, kanonik krakowski – jest autorem pierwszego traktatu na temat pisowni polskiej (świadomie unikam słowa *ortografia*, które moim zdaniem sugeruje istnienie jakiejś skodyfikowanej normy). Jak pisał Kucała,

w miarę postępu badań i coraz lepszego poznawania średniowiecznego okresu polszczyzny traktatowi temu przyznaje się coraz większą wartość bezpośredniego świadectwa o fonologii ówczesnego języka polskiego. Ważny jest bowiem ten zabytek nie tyle dla dziejów ortografii, co dla historycznej fonologii polskiej. [Kucała 1985: 5]

Warto zatem zadbać o to, by w XXI wieku był on łatwo dostępny w maksymalnie wiernej formie.

Zasadniczym celem artykułu jest prezentacja założeń przygotowanej przeze mnie eksperymentalnej edycji elektronicznej, ma on jednak również ważny cel poboczny, mianowicie zainteresowanie Czytelników funkcjonowaniem konsorcjum Unicode, które ustala sposoby reprezentowania znaków w komputerach; dodanie znaków przydatnych dla dawnej polszczyzny do standardu jest trudne, ale nie niemożliwe – wymaga jednak nie tylko wiedzy specjalistycznej, ale i orientacji w organizacji działalności normalizacyjnej. Innym celem pobocznym jest zwrócenie uwagi Czytelników na narzędzia informatyczne, które zaczynają być stosowane również w humanistyce.

Kluczowy dla artykułu termin *znak* – a dokładniej *znak piśmienny* w sensie *piśmienny* ‘odnoszący się do pisma’<sup>1</sup> – używany jest w znaczeniu maksymalnie zbliżonym do znaczenia angielskiego terminu *character*. Słowo to występuje także w potocznym języku angielskim, ale chodzi tutaj o jego znaczenie w powszechnie obecnie używanym standardzie Unicode, który będzie tu jeszcze wielokrotnie wspominany. Niestety nie ma możliwości podania zwięzłej definicji terminu *character* ze względu na liczne niuanse i wyjątki – w związku z tym w artykule *Standard Unicode 4.0. Wybrane pojęcia i terminy* [Bień 2004] pół żartem, pół serio zaproponowałem przyjęcie, że znaki Unicodu to abstrakcyjne obiekty zdefiniowane przez jawne wyliczenie.

Swoje stanowisko pozwolę sobie poprzeć obszernymi cytatami z książki właściwie w całości poświęconej temu tematowi w moim swobodnym tłumaczeniu:

Pojęcie znaku jest jednym z najtrudniejszych podstawowych pojęć technologii informacyjnej, a jednocześnie podstawowym pojęciem przetwarzania tekstów, baz danych, World Wide Web („wszechświatowej pajęczyny”), anotowania tekstów za pomocą znaczników XML, internacjonalizacji oprogramowania i innych dziedzin. [Korpela 2006: ix]

Z samej natury rzeczy problemy kodowania znaków są trudne i bardzo często źle rozumiane. [Korpela 2006: xii]

Podjęmowane były różne próby opisanie, czym są znaki. Porównywano je nawet do idei platońskich. Istota problemu polega na tym, że w pojęciu znaku jest tak dużo aspektów negatywnych: jest on zdefiniowany przede wszystkim przez wskazanie, czym znak **nie jest**. Definiując znak, abstrahujemy od różnych własności i konkretnych cech, aż pozostanie bardzo mało – coś, co można nazwać ideą pewnego konkretnego znaku. [Korpela 2006: 9]

## 2. Rękopis

Jak wiadomo, traktat znajduje się w Bibliotece Jagiellońskiej. Jest to fragment kodeksu opisanego na karcie katalogowej jako *Opera varia* o sygnaturze Rkps. BJ 1961 II; zachowany egzemplarz traktatu to kopia bez tytułu. Jak pisał Kucąła:

Karty, zapisane dwustronnie, mają wymiary 21,2 x 15,5 cm, szpalty zapisane – 18,5 x 11,5 cm. Na niektórych marginesach są dopisane pojedyncze litery,

1 <https://sjp.pwn.pl/doroszewski/pismienny;5472352.html> [dostęp: 10 maja 2019].

ich połączenia, wyjątkowo – inne dopiski. Traktat nie jest pisany jedną ręką. Niewątpliwie jeden z kopistów przepisał połowę tekstu do strony 9 włącznie, drugi połowę od strony 10 do końca. [Kucała 1985: 6]

Rękopis w formie faksymile został udostępniony dopiero przez Kucałę [1985: 39–54], wcześniej w tej formie dostępne były tylko drobne fragmenty dwóch ostatnich stron traktatu [Maciejowski 1852: 95–98]. Tekst rękopisu był wcześniej publikowany dwukrotnie [Bandtkie, Raczyński 1830; Łoś 1907]<sup>2</sup>. Za podstawę edycji elektronicznej przyjąłem odczytanie Kucały [1985: 53–79], zakładając, że jako najnowsze jest ono najlepsze.

W 2003 roku kodeks został zmikrofilmowany, mikrofilm otrzymał numer inwentarza BJ 17471. Kopia odpowiedniego fragmentu została zeskanowana w Bibliotece Uniwersyteckiej w Warszawie (Biblioteka Jagiellońska nie dysponowała odpowiednim sprzętem) i skany te były wykorzystywane przeze mnie do opisanych dalej eksperymentów. Starałem się uzyskać od Biblioteki Jagiellońskiej pisemną zgodę na udostępnienie tych skanów publicznie, ale napotkałem bierny opór; w końcu uzyskałem pośrednio – z pomocą Sybilli Stanisławskiej-Kloc – zgodę ustną. Nieco więcej informacji o prawnych i organizacyjnych aspektach sprawy można znaleźć w moim artykule [Bień 2016c]; tutaj warto podkreślić, że prawa Kucały do odczytania rękopisu wygasły niedawno i jest ono obecnie w domenie publicznej, nie dotyczy to jednak komentarzy i tłumaczenia.

W 2018 roku kodeks został zeskanowany (lub sfotografowany) w kolorze i udostępniony w Jagiellońskiej Bibliotece Cyfrowej pod adresem <https://jbc.bj.uj.edu.pl/publication/476371> [dostęp: 10 maja 2019] jako plik PDF o objętości prawie 300 MB; niektóre strony przy skanowaniu zostały omyłkowo powtórzone, w konsekwencji traktat Parkosza zaczyna się na s. 7 i kończy na s. 22.

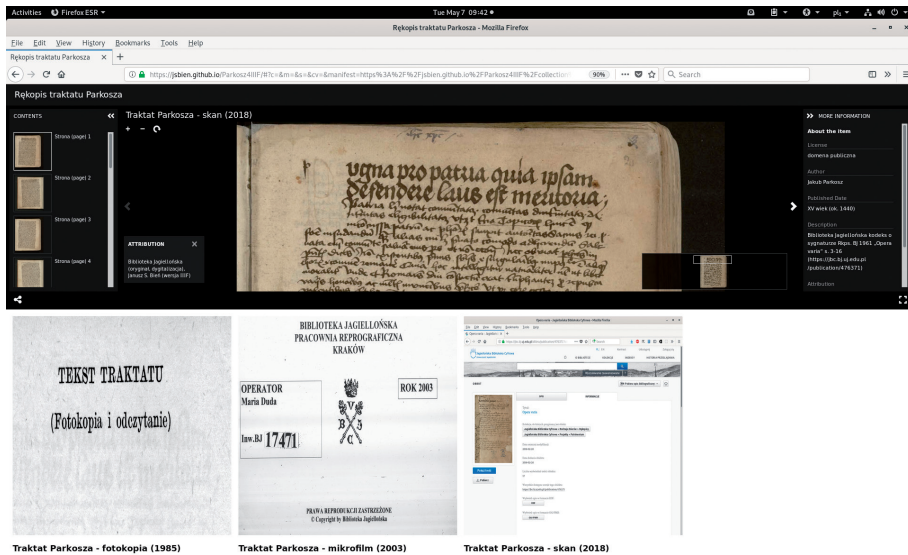
Wszystkie wymienione reprodukcje traktatu udostępniłem ostatnio w dość wygodnej formie pod adresem <https://jsbien.github.io/Parkosz4IIIF>, wykorzystując stosunkowo nowy standard IIIF<sup>3</sup> i przeglądarkę *Universal Viewer* (por. il. 1)<sup>4</sup>.

---

2 Zdygitalizowana w skandalicznie nieudolny sposób edycja Łosia znalazła się w już nieistniejącej Polskiej Bibliotece Internetowej i stała się podstawą co najmniej dwóch e-booków (opublikowanych przez Virtualo i Netpress Digital), które niestety są nadal dostępne w różnych serwisach i być może używane np. przez studentów.

3 *International Image Interoperability Framework* (<https://iiif.io/> [dostęp: 10 maja 2019]); skrót po angielsku czytany „triple i f”.

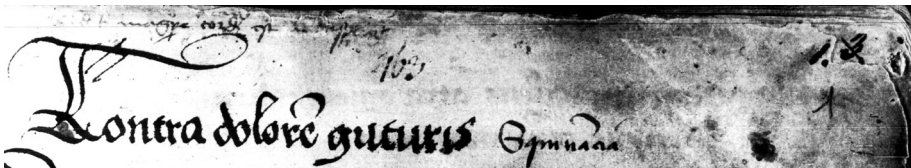
4 Było to możliwe dzięki wykorzystaniu oprogramowania <https://github.com/UniversalViewer/uv-app-starter> [dostęp: 10 maja 2019], którego autorem jest Edward Silverton. Patrz także <https://github.com/jsbien/Parkosz4IIIF> [dostęp: 10 maja 2019].



Ilustracja 1. Traktat w przeglądarce IIIF (*Universal Viewer*)

Między faksymile z książki Kucalę i mikrofilmem oraz skanem są pewne opisane dalej różnice.

Strony traktatu zostały ponumerowane ręcznie, przy czym uwzględniono kartę poprzedzającą traktat (por. il. 2). Numery zostały wpisane tylko na stronie *recto* kart.



Ilustracja 2. Paginacja karty poprzedzającej traktat

Ta paginacja została przez kogoś przekreślona, wprowadzono za to numerację kart, również uwzględniając kartę poprzedzającą (por. il. 3).

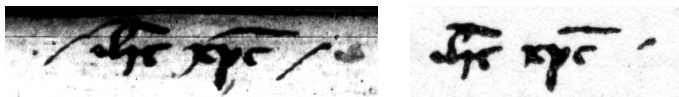


Ilustracja 3. Paginacja pierwszej (*sic!*) i piątej (*sic!*) strony traktatu

Kucała [1985] w swoim odczytaniu podaje przekreśloną numerację stron, stąd w książce na stronach 41–54 znajduje się faksymile stron oznaczonych odpowiednio od [3] do [16]. W niniejszym artykule do stron traktatu odwołuję się, stosując oznaczenia Kucały.

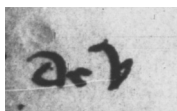
Niektóre różnice między faksymile w książce a mikrofilmem to prawdopodobnie efekty uboczne obróbki graficznej na etapie przygotowywania książki do druku – niestety nie udało mi się dotrzeć do osób w to zaangażowanych.

Różnice najwyraźniej widać w pierwszym wierszu rękopisu (patrz il. 4). Podejrzewam, mówiąc technicznie, zbyt agresywną binaryzację obrazu.

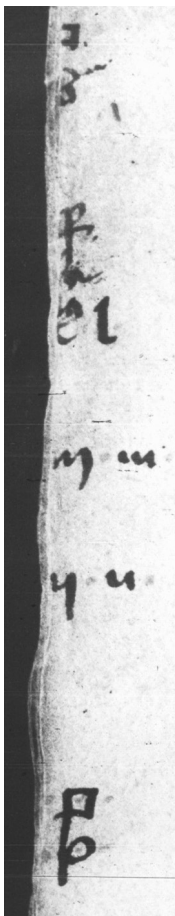


Ilustracja 4. *Iesus Christus*: książka i mikrofilm

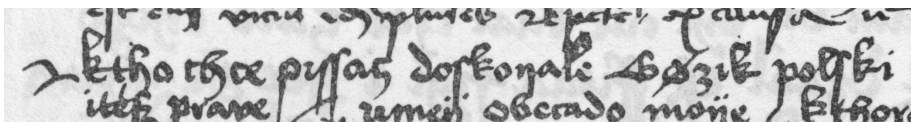
Dwukrotnie zostały w faksymile pominięte marginalia, przy czym tylko w jednym wypadku ich tekst został pominięty w odczytaniu Kucały (patrz il. 5). Marginalia uwzględnione w odczytaniu przedstawiam na il. 6.



Ilustracja 5. Marginalia na stronie [6], wiersz 4 od dołu, pominięte w faksymile i w odczytaniu



Ilustracja 6. Marginalia na stronie [8] pominięte tylko w faksymile



Ilustracja 7. Niezidentyfikowany znak na marginesie, strona [15], wiersz 3

We wszystkich edycjach i odczytaniach został pominięty trudny do zidentyfikowania znak występujący na początku wiersza 3 na stronie [15] (patrz il. 7). Może to jakaś forma znaku akapitu [por. Houston 2015: 15–34].

Numeracja wierszy na stronach stosowana przez Kucalę nie budzi wątpliwości, z wyjątkiem tabeli na przedostatniej stronie, pełnej międzywierszowych dopisków.

### 3. Repertuar znaków rękopisu

Identyfikacja znaków piśmiennych występujących w rękopisie jest podstawowym warunkiem jego odczytania, a ustalenie reprezentacji znaków nietypowych jest podstawowym warunkiem zapisania konkretnego odczytania.

#### 3.1. Abrewiatury

Powszechnie znany jest fakt występowania w traktacie specyficznych polskich liter, które nigdy nie weszły do użytku i w związku z tym nie doczekały się standardowej reprezentacji w druku, a tym bardziej w komputerze.

Mniej znany jest fakt występowania w rękopisie licznych abrewiatur, które istotnie utrudniają jego odczytanie. Kucała zauważał:

Pismo jest bardzo dobrze zachowane, wyraźne, tylko niektóre miejsca są trudne do odczytania. Główna trudność polega na wielkiej ilości skrótów, skrótów tak znacznych, że w niektórych wypadkach nie można ich jednoznacznie rozwiązać. [Kucala 1985: 6–7]

Podobnie pisał Jan Łoś:

oryginał jest bardzo trudny do odczytania, ponieważ użyto w nim mnóstwa skrótów, z którymi tylko bardzo wprawny i doświadczony paleograf poradzić sobie zdoła. [Łoś 1907: 19 <365>]

Warto zwrócić uwagę na stwierdzenie Kucaly, że nie zawsze skróty są jednoznaczne. Na przykład słowo zapisywane przez Kucalę jako *eciam* jest często pisane za pomocą skrótu (można to sprawdzić, korzystając z narzędzia korpusowego wspomnianego na s. 120), który można odczytywać również jako *etiam*. Mamy więc do czynienia nie tyle z odczytaniem, ile z interpretacją rękopisu. Dla wygody używam tych określeń wymiennie.

Dobrym przykładem abrewiacji jest pierwszy wiersz traktatu, przytoczony na wspomnianej już il. 4. Każde ze słów *Iesus Christus* jest zapisane za pomocą trzech liter, kreska nad nimi (lepiej widoczna na mikrofilmie) wskazuje na to, że są to skróty. We wszystkich edycjach skróty te są rozwiązane, powstaje jednak pytanie, czy i jak można – gdyby była taka potrzeba – zapisać w komputerze same skróty.

Kwestię tego, jakie konkretnie litery występują w tym napisie, konsultowałem na zamkniętej liście dyskusyjnej środowiska Medieval Unicode Font Initiative; niestety nie ma forum, na którym można by takie pytania zadawać



polonistom, a próby uzyskania rzeczowej odpowiedzi drogą kontaktów osobistych zakończyły się niepowodzeniem.

Moi zagraniczni internetowi konsultanci – mając do dyspozycji tylko ten fragment – nie byli do końca zgodni, jak należy go odczytać, wątpliwości dotyczyły przede wszystkim tego, czy są to litery greckie, czy łacińskie. Przytoczę tutaj tylko jeden z poglądów, zgodnie z którym są to litery *ihc xpc*. Kwestię być może można rozstrzygnąć definitywnie, analizując wnikliwie wystąpienia poszczególnych liter w całym rękopisie (a dokładniej we fragmencie pisanym przez tego samego skrybę).

Kreska nad literami wskazująca na wystąpienie abrewiatury nie ma – o ile mi wiadomo – żadnej nazwy w polskiej terminologii paleograficznej, np. Aleksander Gieysztor [2009: 143] mówi po prostu o kresce nadpisanej, Władysław Semkowicz [2007] uważa kreską za tak oczywistą, że praktycznie w ogóle o niej nie wspomina, choć oczywiście występuje ona w przykładach. Kreska o podobnej funkcji, choć innego kształtu, używana w piśmie staro-cerkiewno-słowiańskim nazywa się *tytlo*. Tytło dla pojedynczego znaku było dostępne w Unicodzie już od wersji 1.1 z 1993 roku (COMBINING CYRILLIC TITLO). W 2013 roku została złożona propozycja [Andreev, Shardt, Simmons 2013] dodania do Unicodu dwóch znaków: COMBINING CYRILLIC TITLO LEFT HALF i COMBINING CYRILLIC TITLO RIGHT HALF, które razem ze znakiem COMBINING CONJOINING MACRON wprowadzonym do standardu w wersji 5.1.0 w 2008 roku [Everson i in. 2007] umożliwiają tworzenie tytuła nad kilkoma literami (*conjoining* w tym kontekście oznacza, że sąsiadujące znaki makronu nad literami się stykają). Propozycja ta została przyjęta i wspomniane znaki są dostępne w standardzie od wersji 8.0 z 2015 roku. Warto podkreślić, że standard nie określa dokładnie wyglądu znaków, a ich nazwy mają charakter umownych identyfikatorów. Nie widzę więc zasadniczych przeszkód, dlaczego znaki te nie mogłyby być używane w tekstach łacińskich czy polskich – oczywiście po stworzeniu odpowiedniego fontu.

Inne przykłady znajdziemy dalej na il. 9. Porównując ten fragment z interpretacją Kucaly na il. 10, widzimy, że skrócone było słowo *parte* (w oryginale *pte*), słowo *ponuntur* (w oryginale chyba *poñūt*), a przede wszystkim *videlicet* (słowo samo już będące skrótem *videre licet*), skrócone do dwóch liter, z których pierwsza to niewątpliwie litera *v*. Skrót tego słowa występuje w słowniku Capelliego [1961: 284, 1912]<sup>5</sup> w kilku wariantach, przez co nie wyjaśnia on, jak czytać drugą literę skrótu. Skłaniam się do tego, żeby odczytywać ją

---

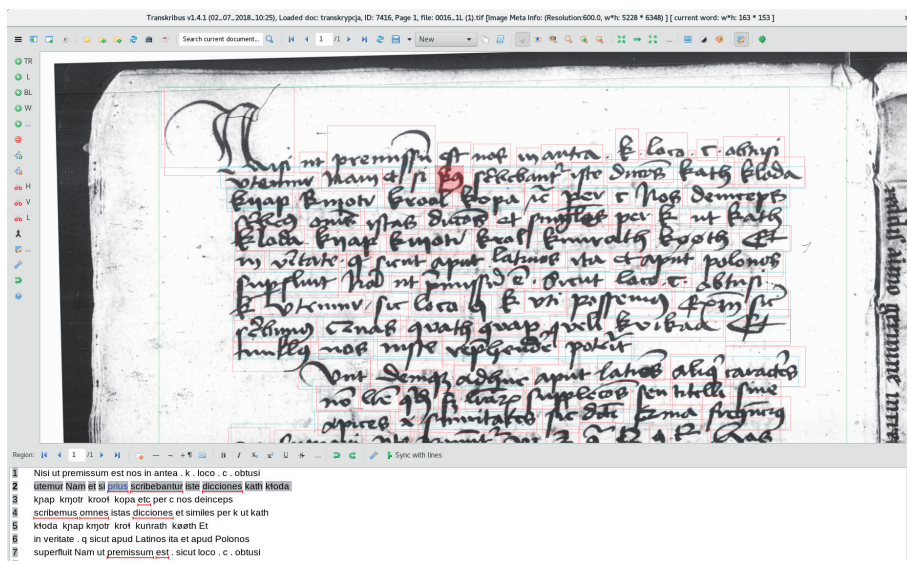
5 Warto odnotować, że dygitalizacja słownika była wykonana metodą *crowdsourcing*, czyli przez wolontariuszy pracujących zdalnie za pomocą odpowiednich narzędzi internetowych.



jako literę *et* (LATIN SMALL LETTER ET – patrz niżej), której niewątpliwe wystąpienie widzimy dalej na rysunkach 9 i 10, ale może to być również litera *z* [por. Cappelli 1961: 284, 1912]; niewystępująca w skracanym wyrazie litera *z* była używana z powodu podobieństwa do tzw. noty tyrońskiej *et* [Wikipedia contributors 2019d].

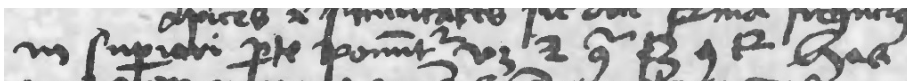
Jak dotąd nikt nie podjął się zidentyfikowania skrótów występujących w rękopisie i sporządzenia ich wykazu. Uczestnicy moich zajęć z grafemiki i paleografii cyfrowej wykonali pewne eksperymenty za pomocą systemu Transkribus<sup>6</sup>, które mogą stanowić punkt wyjścia do takiej pracy – por. il. 8; wyniki tych eksperymentów mogą być udostępnione na życzenie osobom, które założą konto w tym systemie (inne formy udostępnienia albo nie zawierają pełnej informacji, albo wymagają specjalnych narzędzi informatycznych).

Normalne użycia abrewiatur mogą i są we wszystkich opublikowanych odczytaniach po prostu rozwijane. Jest jednak także fragment traktatu, w którym kilka skrótów jest omawianych – jest on przedstawiony właśnie na już wspomnianej wyżej il. 9.



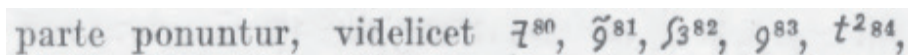
Ilustracja 8. Okno transkrypcyjne systemu Transkribus – skróty są wyróżnione podkreśleniem

6 <https://transkribus.eu/> [dostęp: 10 maja 2019] [patrz też np. Seaward, Kallio 2017; Mühlberger, Colutto, Kahle 2017].



Ilustracja 9. Strona 14, wiersz 14 rękopisu

W edycji Kucaly fragment ten ma postać przedstawioną na il. 10, a polskie skomentowane tłumaczenie [Kucala 1985: 107] widzimy na il. 11. Znaki dla występujących tam skrótów najwyraźniej do drukowanego tekstu zostały dopisane ręcznie – to przypuszczenie potwierdziła w 2017 roku w rozmowie telefonicznej redaktor merytoryczna książki Zenobia Mieczkowska<sup>7</sup>.



Ilustracja 10. Strona 14, wiersz 14 rękopisu (tylko środkowy fragment) w odczytaniu i zapisie Kucaly

Na koniec: Są jeszcze (w piśmie) u Łacinników pewne znaki nie będące wprawdzie literami, ale zastępnikami liter. Są to skróty albo wierchołki, zwane *summitates*, ponieważ częściej są umieszczane w części górnej, a mianowicie: 7 (*eciam* 'także'), ǫ (*contra* 'przeciw'), s (*scilicet* 'a mianowicie, rozumie się'), 9 (*con, com, cum* – sylaba początkowa), t<sup>2</sup> (*tur* – końcówka).

Ilustracja 11. Tłumaczenie (z komentarzami) Kucaly akapitu zawierającego wiersz 14 strony 14 rękopisu

Elektroniczna reprezentacja tekstu wymaga przypisania znakom konkretnych wartości liczbowych (współrzędnych kodowych). Potrzeba kodowania dawnych tekstów jest dość powszechna, więc najczęstsze znaki ze średnio-wiecznych rękopisów i druków zostały już dawno uwzględnione w standardzie Unicode. Konkretnie znak TIRONIAN SIGN ET pojawił się w 3.0 wersji standardu w 1999 roku, przejęty z międzynarodowej normy ISO 5426 [Wikipedia contributors 2019b], praktycznie chyba nieużywanej (może z wyjątkiem nielicznych bibliotek) z powodu trudności technicznych. Znaki LATIN SMALL LETTER CON i LATIN SMALL LETTER ET zostały wprowadzone

<sup>7</sup> Mówiąc dokładniej, znaki narysowane przez grafika były wklejane w odpowiednie miejsca, co było bardzo pracochłonne, bo takich „wklejek” było chyba ok. 2000 – jeśli dobrze zrozumiałem, brała w tym udział prawie cała redakcja.

w wersji 5.1.0 w 2008 roku [Everson i in. 2005]. Znaki, na które było mniejsze zapotrzebowanie lub które budziły jakieś wątpliwości, zostały uwzględnione w tzw. obszarze użytku prywatnego (ang. PUA – *Private Use Area*) w rekomendacjach Inicjatywy Unikodowych Fontów Średniowiecznych (MUFI – *Medieval Unicode Font Initiative*); znak COMBINING ABBREVIATION MARK SUPERSCRIPT UR ROUND R FORM pojawił się (pod nieco inną nazwą) już w wersji 1.0 z 2003 roku [Haugen 2003: 60]. Z pomocą takich dodatkowych znaków jak COMBINING SHORT STROKE OVERLAY i COMBINING TILDE możliwe jest zakodowanie tego fragmentu w satysfakcjonujący sposób (por. il. 12).

Font	Współrzędne kodowe				
	204A, 0335	A76F, 0303	017F, A76B	A76F	0074, 00A0, F153
Junicode	ɹ	ṙ	ƚ	9	t <sup>2</sup>
Cardo	7	ṙ	ƚ	9	t <sup>2</sup>
Palemonas	7	ṙ	ƚ	9	t <sup>2</sup>
Quivira	7	ṙ	ƚ	9	

Ilustracja 12. Omawiane skróty w kilku fontach i ich współrzędne kodowe

### 3.2. Kwestia znaków diakrytycznych

Kucała pisał:

Trzeba tu zwrócić uwagę, że Parkosz nie stosował w ogóle znaków diakrytycznych. Jedynym wyjątkiem jest *y* z dwiema kropkami (czasem z ukosnymi kreszczkami): *ẏ*, które przejął z pisowni dotychczasowej (podobnie jak *i* z kropką). Nad żadną literą spółgłoskową nie położył ani jednej kreski. Na nieporozumieniu polega więc twierdzenie, już dawniej spotykane, a ostatnio dość stanowczo sformułowane przez S. Jodłowskiego: „Znaki diakrytyczne spotykamy m.in. w traktacie [...] Jakuba Parkoszowica [...] warto tu odnotować stosowanie przez Parkoszowica dwu znaków diakrytycznych w postaci dwu poziomo ułożonych kreszek nad literą *y* oraz przecinka nad literami spółgłoskowymi, mającego zaznaczać miękkość spółgłoski”. Nieporozumienie wynikło stąd, że J. Łoś w swoim wydaniu traktatu w 1907 r. [...] dał w transliteracji *l* z przecinkiem u góry [...]. Za Łosiem to powtarzano, m.in. Z. Klemensiewicz przy omawianiu ortografii Parkosza. [Kucała 1985: 20]

Stwierdzenie Stanisława Jodłowskiego [1979: 22–23] jest niewątpliwie błędne, i to z kilku powodów. Kreski nad *y* nigdy nie były umieszczane

poziomo, choć – jak wspomina wyżej Kucała – czasami zamiast kropek były użyte kreski ukośne. Przecinek – nie nad literami, ale nad literą *l* czy raczej obok niej – był wykorzystany przez Łosia zapewne ze względów technicznych do transkrypcji rękopiśmiennej litery *l* z pętelką – nie chodziło przy tym o miękkość we współczesnym rozumieniu tego słowa, ale raczej o zróżnicowanie *l* i *ḷ*. Swoją drogą jest dziwne, że Łoś tego w żaden sposób nie skomentował. Dziwne też, że nie próbował wiernie oddać tej litery; wydaje się, że dla innych liter wprowadzonych przez Parkosza na potrzeby edycji [Łoś 1907] przygotowano specjalne czcionki. Tak czy inaczej był to postęp w stosunku do edycji Samuela Bandtkiego [1830], który tych dwóch form *l* w ogóle nie odróżniał.

Klemensiewicz [1961: 100–104] w wykazie liter Parkosza zamieszcza *l* z akcentem (nie jest to ewidentnie przecinek, raczej matematyczny prim), ale o znakach diakrytycznych pisze wprost, że Parkosz ich nie stosował – najwyraźniej taką formę *l* uważa za samodzielną literę. Ciekawe, że ta niepoprawna forma tej litery nie została poprawiona lub przynajmniej skomentowana w późniejszych wydaniach (kilka wydań ukazało się już po śmierci autora, ostatnie w 2015 roku).

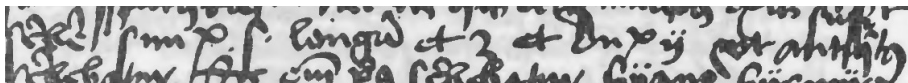
Nie próbowałem ustalić, w jakich jeszcze publikacjach występował ten błąd, natknąłem się natomiast na następujące stwierdzenie w *Średniowieczu* Teresy Michałowskiej [1995: 334], spopularyzowane niestety przez Wikipedię w haśle dotyczącym Parkosza [Wikipedia 2018a]:

Parkoszowie nie wprowadził znaków diakrytycznych, pozostawiając tylko wcześniejsze „y” oraz „i”.

Mamy tutaj ewidentny błąd drukarski: *y* zamiast *ÿ*. Ta wielokrotnie nagrodzona książka miała dotąd osiem wydań (ostatnie w 2008 roku), ale w żadnym z nich nie poprawiono tego błędu.

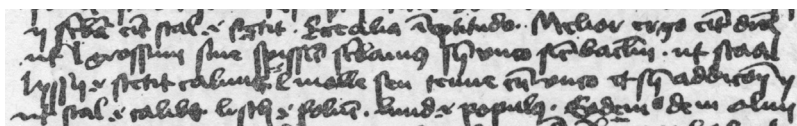
Ciekawe, że stosowanie przez Kucałę znaku *ÿ* jest niezgodne z *Zasadami wydawania tekstów staropolskich*, w których czytamy: „Wszystkie warianty *y* występujące w rękopisach [...] zastępujemy dzisiejszym *y*, chyba że analiza grafiki wykaże funkcjonalne zróżnicowanie poszczególnych odmian [...]” [Górski 1955: 28]. Inne możliwe odczytania to sekwencja liter *i* oraz *j* – tak czytał Bandtkie [1830] – oraz ligatura tych liter. W edycji Łosia położenie kropek wskazuje, moim zdaniem, że jest to ligatura. Rozstrzygający dla sprawy wydaje się jednak następujący fragment traktatu: „Między *n* twardym a *y* podwójnym będzie różnica, ponieważ ogonek *n* pociąga się na prawo, ogonek *y* zgina się w lewo. Ponadto nad *y* pisze się podwójną kropkę dla odróżnienia od *n*”

[Kucała 1985: 95]<sup>8</sup>. W tej sytuacji nie ma chyba wątpliwości, że znak – nazywany niekiedy w polskiej literaturze *kropkowanym y* [np. Mika 2015] – należy reprezentować, wzorując się na Kucałe, jako odpowiednio LATIN SMALL LETTER Y WITH DIAERESIS i LATIN CAPITAL LETTER Y WITH DIAERESIS. Warto dodać, że odróżnienie w rękopisie tych trzech przypadków tylko na podstawie kształtu jest chyba niemożliwe, zwłaszcza że przynajmniej niektóre użycia *y* wywodzą się z sekwencji *i* oraz *j* [por. Wikipedia 2019b].



Ilustracja 13. Strona [13], wiersz 19 – odczytanie Kucały: scribere simplex / longum et z, et duplex *y*, ut antiquitus

Co do wariantów litery *l*, to na il. 14 widzimy fragment rękopisu, jego odczytania przez Łosia i Kucałę oraz tłumaczenie. W obu edycjach *l* twarde (czyli *l̄*) jest oddawane przez zwykłe *l*, u Kucały *l* miękkie przez ręcznie dopisywaną literę *y*.



aliam ineptitudinem. Melior ergo erit differentia, ut *l* grossum sive spissum scribamus sine unco, sicut baculum ut: *staal lyssy* id est stetit calvus; *l* molle seu tenue cum unco et sine addicione *y*, ut *stal* id est calibs, *listh* id est folium, *luud* id est populus.

Ecce aliam ineptitudinem. Melior ergo erit differentia, | ut *l* grossum sive spissum scribamus sine unco, sicut baculum, ut *staal* | *lyssy* id est stetit calvus, *l* molle seu tenue cum unco et sine addicione *y*, | ut *stal* id est calibs, *listh* id est folium, *luud* id est populus.

kład niedorzeczności. Lepiej więc będzie wyrażona różnica, gdy napiszemy *l* twarde bez łuczka, jako łaskę np. *staal lyssy* 'stał lysi', a *l* miękkie z łuczkiem, a bez dodawania *y*, np. *stal* 'stal', *listh* (*list*) 'liść', *luud* 'lud'.

Ilustracja 14. Rękopis, strona [8], wiersze 40–43, edycja Łosia, strona 29 [405], odczytanie i tłumaczenie Kucały strony 67 i 97

<sup>8</sup> W rękopisie *y* rzeczywiście jest podobne często do *n*, por. il. 23.

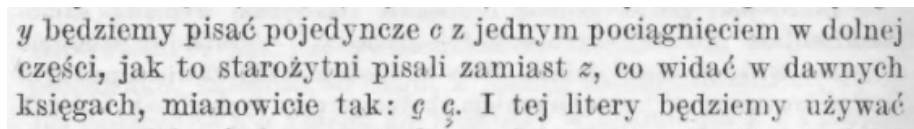


Niestety znaku o takim kształcie nie ma w standardzie, choć są podobne w bloku Międzynarodowego Alfabetu Fonetycznego (IPA). Moim zdaniem najbliższa kształtem litera to LATIN SMALL LETTER L WITH BELT wprowadzona w 1993 roku w wersji Unicode 1.1.0 właśnie na potrzeby transkrypcji fonetycznej (obecnie dostępna bez problemu w popularnych fontach), ale uzupełniona w 2014 roku w wersji 7.0 o LATIN CAPITAL LETTER L WITH BELT na potrzeby języków Alibamu i Kosati [Jensen, Pentzlin 2012].

Nasuwa się pytanie, czy decyzje Łosia i Kucwały dotyczące stosowania w transkrypcji litery *l* dla *l* twardego były optymalne – sądzę, że chodziło raczej o zminimalizowanie technicznych problemów składu. Ja podjąłem decyzję o reprezentowaniu tego znaku w odmienny sposób. Początkowo stosowałem literę LATIN SMALL LETTER L WITH BAR, wprowadzoną do standardu w wersji 1.1.0 w 1993 roku i uzupełnioną w wersji 4.1.0 z 2005 roku o LATIN CAPITAL LETTER L WITH BAR na potrzeby indiańskiego dialektu saanicz [Wikipedia 2018b]. Obecnie uważam, że lepiej do tego celu nadaje się LATIN SMALL LETTER L WITH DOT BELOW i LATIN CAPITAL LETTER L WITH DOT BELOW. Oba znaki pojawiły się już w wersji 1.1.0 z 1993 roku, prawdopodobnie na potrzeby Międzynarodowego Alfabetu Transliteracji Sanskrytu.

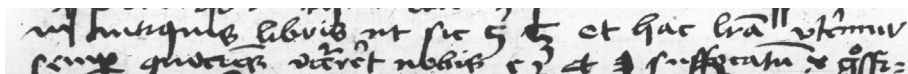
### 3.3. Znaki wspomniane przez Parkosza

Inny trudny technicznie fragment w tłumaczeniu Kucwały ma postać pokazaną na il. 15, a odpowiedni wers rękopisu jest przedstawiony na il. 16. Pierwszy omawiany znak u Kucwały, a także u Łosia, to niewątpliwie LATIN SMALL LETTER C WITH CEDILLA, razem z LATIN CAPITAL LETTER C WITH CEDILLA stosowany w wielu językach (najwcześniej pojawił się w starohiszpańskim) i dostępny w standardzie już od wersji 1.1.0 z 1993 roku; wiadomo, że litera ta pochodzi od litery *z* [Wikipedia contributors 2019a], co potwierdza również Parkosz, por. wspomniana już il. 15. Drugi znak, który w edycji Kucwały był tworzony za pomocą ręcznego dopisku, to moim zdaniem LATIN SMALL LETTER VISIGOTHIC Z wprowadzony do standardu razem z LATIN CAPITAL LETTER VISIGOTHIC Z w wersji 5.1.0 z 2008 roku [Everson i in. 2005].



y będziemy pisać pojedyncze *c* z jednym pociągnięciem w dolnej części, jak to starożytni pisali zamiast *z*, co widać w dawnych księgach, mianowicie tak: *c̣ c̣*. I tej litery będziemy używać

Ilustracja 15. Fragment strony [10] w tłumaczeniu Kucwały (s. 100)



Ilustracja 16. Strona [10], wiersz 23 rękopisu

### 3.4. Samogłoski nosowe

Na oznaczenie samogłoski nosowej Parkosz stosuje literę *o*. Według Lisowskiego [2010] w tej funkcji występowała ona w rękopisach już od XII wieku. Jednak Michałowska wprowadzenie tej litery przypisuje Parkoszowi:

Zasób łacińskich liter odpowiadających samogłoskom powiększył o nieodzowny w polskim systemie graficznym znak nosówki: „*o*”. [Michałowska 1995: 334]

To błędne, a przynajmniej mylące, stwierdzenie Michałowskiej zostało niestety spopularyzowane przez Wikipedię we wspomnianym już haśle [Wikipedia 2019a].

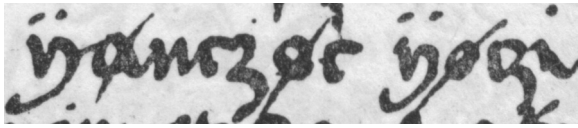
Kucała używa przekreślonego *o* w transkrypcji, przypisując mu głoskę [ǫ] [patrz Kucała 1985: 17]. Ciekawe, że użycie to jest niezgodne z *Zasadami wydawania tekstów staropolskich*, w których czytamy: „Różne postaci *o* przekreślonego [...] sprowadzamy do znormalizowanej postaci” [Górski 1955: 28] – patrz il. 17.

Nie zachowujemy następujących wariantów graficznych oryginałów: Różne postaci *o* przekreślonego (ϕ ϕ ϕ ϕ ϕ) sprowadzamy do znormalizowanej postaci *o*. Różne postaci *a* przekreślonego (ϕ ϕ ϕ)

Ilustracja 17. *O* przekreślone w *Zasadach wydawania tekstów staropolskich*

W standardzie Unicode tym literom odpowiadają znaki LATIN SMALL LETTER O WITH STROKE i LATIN CAPITAL LETTER O WITH STROKE, dostępne już od wersji 1.1 z 1993 roku ze względu na stosowanie ich w językach duńskim, norweskim i farserskim, a także w międzynarodowym alfabecie fonetycznym IPA. Uważam, że za Kucałą, a wbrew *Zasadom...*, należy używać właśnie tych znaków.





ÿanczoc ÿoci

Ilustracja 18. Samogłoski nosowe w rękopisie i edycji Kucala

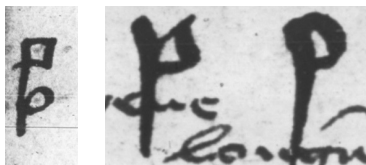
W rękopisie występuje również przekreślona litera *a*, którą Kucala oddaje literą *ą* – por. il. 18. Ja zdecydowałem się reprezentować ją wiernie przez LATIN SMALL LETTER A WITH STROKE, wprowadzoną w wersji 5.0.0 z 2006 roku i LATIN CAPITAL LETTER A WITH STROKE, wprowadzoną w wersji 4.1.0 w 2005 roku – najpierw została wprowadzona litera duża na potrzeby języka saanicz [Elliot, Brand, Harvey 2004], a potem dla kompletności litera mała [Freytag, Whistler 2005].

### 3.5. Warianty liter *b* i *p*

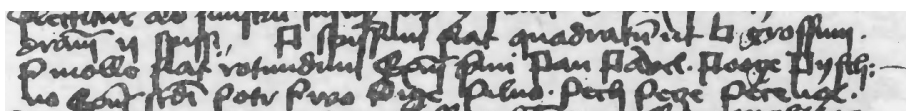
Najbardziej znaną innowacją Parkosza jest zapożyczenie z ówczesnej notacji muzycznej znaków *b durum* (znak nazywany także *b quadratum*) i *b molle* (znak nazywany także *b rotundum*) [por. np. Apel 1969: 5] na oznaczenie odpowiednio twardego i miękkiego *b* – patrz il. 19. Znak *b molle* przetrwał do naszych czasów w nieznacznie tylko zmienionej formie jako bemol. Znak *b durum* w notacji muzycznej z czasem zmienił się w krzyżyk. W standardzie Unicode kształt *b molle* całkiem dobrze oddaje LATIN SMALL LETTER B WITH HOOK wprowadzony w wersji 1.1.0 w 1993 roku na potrzeby notacji fonetycznej, ale uzupełniony o LATIN CAPITAL LETTER B WITH HOOK ze względu na zaprojektowany nieco wcześniej alfabet pannigeryjski [Wikipedia contributors 2019c]. Niestety nie ma w standardzie żadnego znaku, który choć trochę przypominałby *b durum*. Dlatego postanowiłem czysto umownie stosować na jego oznaczenie LATIN SMALL LETTER B WITH STROKE – ze względu na liczne zastosowania znak wprowadzony już w wersji 1.1.0 w 1993 roku, uzupełniony w wersji 5.0.0 w 2006 roku o LATIN CAPITAL LETTER B WITH STROKE na potrzeby m.in. języka katu używanego w Wietnamie [Priest 2004b: 7].



Ilustracja 19. Przykłady *b durum* i *b molle* w rękopisie: strona [15], wiersz 28, strona [8], wiersz 2, strona [7], wiersz 15



Ilustracja 20. Litery *p quadratum* i *p molle* w rękopisie: marginalia na stronach [8] i [15], wiersz 32



{*p p*}. *p* spissum fiat quadratum, ut *b* grossum. | *p* molle fiat rotundum. Exemplum primi: *Pan Pavel. Pooge Pysch|no.* Exemplum secundi: *Potr Pivo Pige Pilno. Pech Pecze Pecena*<sup>46</sup>. |

<sup>46</sup> Zaczyna się tutaj pojawiać *p* miękkie pisane nieco inaczej niż zwykle *p*, mianowicie *p̃p*.

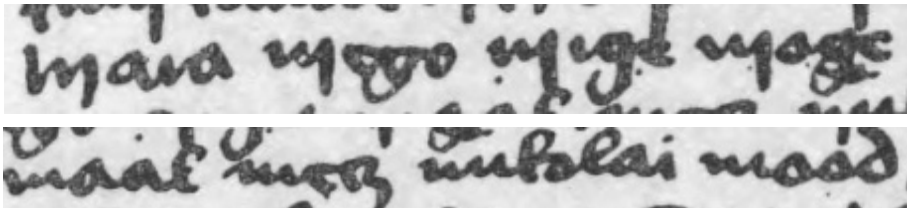
Ilustracja 21. Fragment rękopisu z *p quadratum* i *p molle* (strona [8], wiersz 18), jego odczytanie i tłumaczenie według Kucali

Przez analogię do *b durum* i *b molle* Parkosz wprowadził także kanciaste *p durum* i zaokrąglone *p molle*. Niestety w standardzie Unicode nie ma znaków, które choć w przybliżeniu przypominałyby kształtem *p durum*. Stąd *p durum* reprezentuję konwencjonalnie przez LATIN SMALL LETTER P WITH STROKE i LATIN CAPITAL LETTER P WITH STROKE. Pierwszy znak pojawił się w wersji 4.1.0 standardu w 2005 roku na potrzeby amerykańskiej notacji fonetycznej [Constable 2003: 8], drugi w wersji 5.0.0 w roku 2006 na potrzeby języka kolumbijskich plemion Tanimuca i Letuama [Priest 2004a: 5]. Co do *p molle*, to chyba dość dobrze jego kształt oddają staroangielskie litery LATIN LETTER WYNN (Unicode 1.1.0, 1993) i LATIN CAPITAL LETTER WYNN (Unicode 3.0.0, 1999), ale ponieważ oznaczają zupełnie inną głoskę, ich użycie wydaje się niewłaściwe. Między innymi z tego względu zdecydowałem się na użycie liter LATIN SMALL LETTER P WITH HOOK i LATIN CAPITAL LETTER P WITH HOOK wprowadzonych do standardu już w wersji 1.1.0 (1993) na potrzeby języków afrykańskich.

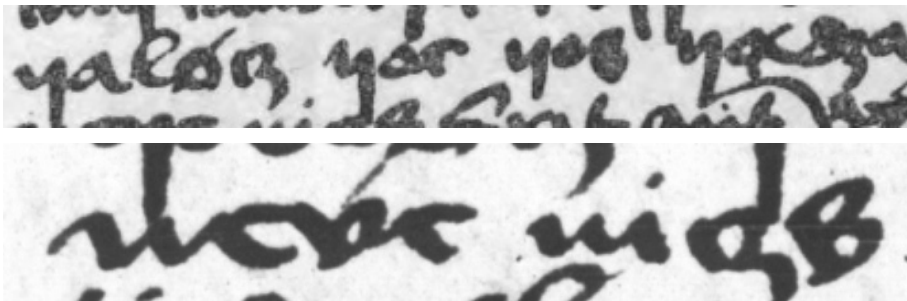
### 3.6. Warianty liter *m* i *n*

O literach tych Parkosz pisał: „Twarde *m* (niech będzie) z ogonkiem przy trzeciej nóżce, jak się zwykło pisać na końcu wyrazów [...]. *M* zaś miękkie – bez

ogonka [...]. Tak również *n* twarde (należy pisać) z ogonkiem, jak się zwykło pisać na końcu wyrazów [...]" [Kucała 1985: 95] – patrz il. 22 i 23.



Ilustracja 22. Przykłady *m* z ogonkiem i bez ogonka w rękopisie: strona [8], wiersze 10 i 11



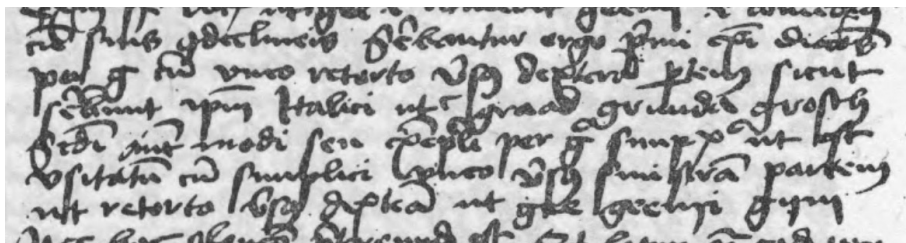
Ilustracja 23. Przykłady *n* z ogonkiem i bez ogonka w rękopisie: strona [8], wiersze 14 i 15

W rękopisie „ogonek” to po prostu wydłużenie (zarówno w przypadku polskich twardych *m n*, jak i w przypadku końcowych *m i n* w tekście łacińskim) – warto przypomnieć, że rękopis jest kopią, a nie oryginałem; wygląda na to, że kopiści nie zadbali o wierne oddanie kształtu ogonka. Jednak w edycjach Łosia i Kucały w literach tych występuje – zgodnie z cytowaną wcześniej deklaracją Parkosza – ogonek zakręcony w prawo. Niestety – z jednym wyjątkiem – brak w standardzie liter o tym kształcie. Wyjątek to LATIN SMALL LETTER N WITH RETROFLEX HOOK – stosuje go Grzegorz Seroczyński w swoich omówieniach traktatu [Seroczyński 2014–2017, 2014]. Jednak w sensie technicznym znak ten – wbrew swojej nazwie – nie jest literą, w szczególności nie ma swojego odpowiednika w górnej kaszcie. Mogłoby to powodować niepotrzebne komplikacje przy komputerowym przetwarzaniu tekstu; analogiczny problem dotyczy znaku LATIN SMALL LETTER M WITH RIGHT DESCENDER z rekomendacji MUFI. W konsekwencji zdecydowałem się na użycie znaków: LATIN SMALL LETTER M WITH HOOK (Unicode

1.1.0, 1993) wprowadzony na potrzeby międzynarodowego alfabetu fonetycznego (użyty także przez Seroczyńskiego w cytowanych omówieniach), LATIN CAPITAL LETTER M WITH HOOK (Unicode 5.1.0, 2008) dodany na potrzeby amerykańskiej notacji fonetycznej [Priest, Constable 2005: 9], LATIN SMALL LETTER N WITH LEFT HOOK (Unicode 1.1.0, 1993) wprowadzony na potrzeby międzynarodowego alfabetu fonetycznego i – przez analogię – LATIN CAPITAL LETTER N WITH LEFT HOOK (Unicode 1.1.0, 1993) wprowadzony na potrzeby języków afrykańskich; być może jednak lepsze byłyby znaki: LATIN SMALL LETTER N WITH LONG RIGHT LEG używany m.in. w indiańskim języku lakota i LATIN CAPITAL LETTER N WITH LONG RIGHT LEG wprowadzony specjalnie na potrzeby tego języka; alternatywą dla dwóch ostatnich znaków jest LATIN CAPITAL LETTER N WITH RIGHT DESCENDER i LATIN SMALL LETTER N WITH RIGHT DESCENDER z rekomendacji MUFI.

Wszystkie dotychczasowe wydania za Parkoszem *m* i *n* miękkie reprezentowały po prostu przez te litery bez żadnych modyfikacji ich kształtu. Uważam, że dla współczesnego czytelnika jest to bardzo mylące, dlatego reprezentuję je odpowiednio przez LATIN SMALL LETTER N WITH DOT BELOW i LATIN CAPITAL LETTER N WITH DOT BELOW oraz LATIN SMALL LETTER M WITH DOT BELOW i LATIN CAPITAL LETTER M WITH DOT BELOW. Znaki te występują w standardzie już od wersji 1.1.0 z 1993 roku, prawdopodobnie zapożyczone z jakiegoś istniejącego już wcześniej kodowania.

### 3.7. Inne spółgłoski



Ilustracja 24. Strona [11], wiersze 27 i następne rękopisu

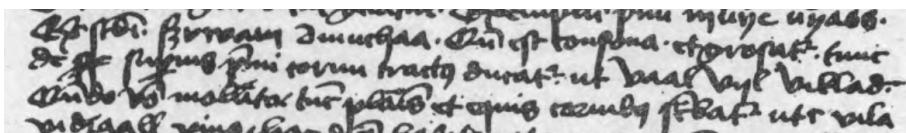
medimus | cum suis condeclineis. Scribantur ergo primi exem-  
pli dictiones | per *g* cum unco retorto versus dexteram partem,  
sicut | scribunt ipsum Italiciei, ut sic: *graad gruuda grofch,* |

Ilustracja 25. Strona [11], wiersze 27 i następne rękopisu – odczytanie Kucalę

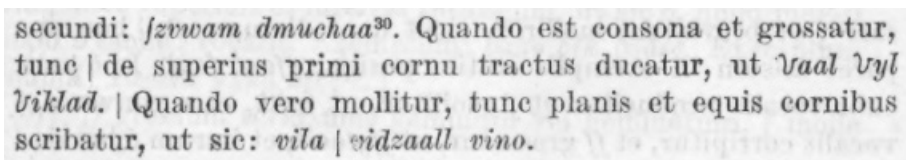
Parkosz stosuje dwa warianty litery *g*. Pierwszy to *g* pisane „z ogonkiem zakrzywionym w prawą stronę, tak jak je piszą Italczycy”<sup>9</sup>, drugie to „*g* proste, jak jest używane, z ogonkiem zwyczajnym, zagiętym w lewą stronę, tj. zwrócone (wybrzuszaniem) na prawo” [Kucała 1985: 102–103].

Pierwszy wariant proponuję reprezentować przez LATIN SMALL LETTER TURNED INSULAR *G* i LATIN CAPITAL LETTER TURNED INSULAR *G*, wprowadzone w 2008 roku do wersji 5.1.0 standardu ze względu na dawne teksty walijskie [Everson 2004]. Stosowanie przez Seroczyńskiego [2014–2017] znaku LATIN SMALL LETTER TURNED DELTA, pomimo pewnego podobieństwa wizualnego, uważam za niewłaściwe.

Zwykłe *g* oznacza u Parkosza spółgłoskę miękką, co dla współczesnego czytelnika jest mylące. Dlatego proponuję stosować zamiast niego LATIN SMALL LETTER *G* WITH DOT ABOVE i LATIN CAPITAL LETTER *G* WITH DOT ABOVE, dostępne w standardzie od wersji 1.1.0 z 1993 roku, wprowadzone m.in. ze względu na język maltański.



Ilustracja 26. Strona [7], wiersze 5 i następne rękopisu



Ilustracja 27. Strona [7], wiersze 5 i następne rękopisu – odczytanie Kucały

9 Chętnie zobaczyłbym jakieś potwierdzenie, że rzeczywiście tak pisali „Italczycy” – być może dostarczyłoby to argumentu za wprowadzeniem tego znaku do standardu Unicode. Jest natomiast znane wystąpienie podobnego znaku w XIII-wiecznym hiszpańskim rękopisie *El Cantar del Mio Cid* – patrz <https://folk.uib.no/hnooh/mufi/pipeline/pipeline-for-v4.html> [dostęp: 10 maja 2019].

3	ʎ	ɬ	Ł	ł	Ł	Ş	Ş
A76B	204A	026C	A7AD	019A	023D	A763	A762
ḃ	Ḃ	ḅ	Ḅ	Ḟ	Ḥ	ḟ	Ḟ
0253	0181	0180	0243	1D7D	2C63	01A5	01A4
ṁ	Ṁ	ṁ	Ṁ	ṁ	Ṁ	ṁ	Ṁ
0253	2C6E	0272	019D	028B	01B2	A77F	A77E

Ilustracja 28. Omawiane znaki w foncie Quivira i ich współrzędne kodowe

Parkosz – zgodnie z ówczesnym zwyczajem – nie używa litery *u*, litera *v* oznacza więc w zależności od kontekstu samogłoskę lub spółgłoskę. W edycji Kucąły jest ona odpowiednio oddawana przez *u* lub *v*, z wyjątkiem tabeli na przedostatniej stronie, w której – może przez przeoczenie – znajdujemy zapisy *breue*.

Litera *v* jako spółgłoska występuje w dwóch postaciach, o których pisze Kucąła: „Kiedy jest spółgłoską i twardnieje, wówczas (w pisowni) z wyższego pierwszego rożka wyprowadza się przeciągnięcie (linii) [...]. Kiedy się zaś miękczka, wtedy się pisze z równymi rożkami [...]” [Kucąła 1985: 93].

Proponuję pierwszy wariant oddawać literami LATIN SMALL LETTER V WITH HOOK i LATIN CAPITAL LETTER V WITH HOOK. Oba znaki są dostępne w standardzie już od wersji 1.1.0 z 1993 roku, wprowadzone częściowo ze względu na ich zastosowanie w transkrypcji fonetycznej, a częściowo ze względu na użycie w językach afrykańskich.

O literze *F* Parkosz pisze: „*F* piszmy jako podwójne, czyli podwojone, *f* miękkie jako pojedyncze” [Kucąła 1985: 95]. W standardzie ligatura dwóch *f* jest dostępna, ale tylko w dolnej kaszcie dla zgodności z wcześniejszymi kodowaniami – nie ma powodu używać jej w transliteracji. Natomiast mylące dla współczesnego czytelnika miękkie *f* pojedyncze proponuję oddawać znakami LATIN SMALL LETTER F WITH DOT ABOVE i LATIN CAPITAL LETTER F WITH DOT ABOVE, dostępnymi w standardzie od wersji 1.1.0 z 1993 roku.

Kucąła pisał: „majuskułową literę oznaczającą [i] i [j] podobną do późniejszego *J* oddaje się (zarówno w wyrazach łacińskich jak polskich) przez *i*, *I*” [Kucąła 1985: 34]. Decyzja ta wydaje mi się nieoczywista, ale jej nie zmieniałem.

Badacz zauważał również: „Długie *f* występujące z reguły na początku i w środku wyrazów zastąpiono *s* krótkim” [Kucąła 1985: 34]. Tej decyzji również nie zmieniałem.



### 3.8. Interpunkcja

Kucała pisał: „Zmodernizowano duże i małe litery, **interpunkcję** i podział na akapity” [Kucała 1985: 34; wyróż. – J.S.B.]. Modernizacja interpunkcji polegała głównie na dodaniu współczesnych znaków interpunkcyjnych dla ułatwienia rozumienia tekstu. Postępowanie takie było zgodne z zasadami przyjętymi dla transkrypcji [Górski 1955: 41], ale nie dla transliteracji [Górski 1955: 29].

Ciekawe, że Kucała pisał również: „Wyrazy i teksty polskie oddaje się w dokładnej transliteracji, pisząc tylko znak przestankowy (kropkę) u dołu wiersza, a nie w środku jak w rękopisie” [Kucała 1985: 34]. W rzeczywistości kropki te zostały wydrukowane w środku wiersza – patrz il. 29. Powstaje więc pytanie, jak je reprezentować w edycji elektronicznej. Najbardziej właściwe wydaje się użycie znaku MIDDLE DOT dostępnego w standardzie od wersji Unicode 1.1.0 z 1993 roku [por. Haugen 2015: 147].



Ilustracja 29. Faksymile, strona [14], wiersz 17 i odczytanie Kucały strony 77: interpunkcja (brak kropek nad ostatnim y to błąd drukarski)

W końcowej części rękopisu występuje znak interpunkcyjny w postaci ukośnej kreski – w książce został oddany przez zwykły ukośnik, któremu w standardzie Unicode odpowiada SOLIDUS. W rekomendacji MUF1 takie średniowieczne znaki mają bardziej precyzyjne odpowiedniki, dla traktatu Parkosza najbardziej właściwy wydaje się znak SMALL VIRGULA i właśnie on został użyty w elektronicznej edycji.

### 3.9. Znaki nieczytelne i wątpliwe

Standard Unicode dysponuje znakiem REPLACEMENT CHARACTER, którym oznaczano niedające się odczytać znaki m.in. w korpusie IMPACT [Bień 2014]. Ze względów technicznych w pominiętych marginaliach z il. 5 użyłem znaku INTERROBANG [por. Houston 2015: 37–51]; choć są one wyraźne, nie chciałem wprowadzać własnego odczytania do tekstu Kucały.

## 4. Kodowanie znaków w standardzie Unicode

Powyższe omówienie problemów kodowania traktatu warto uzupełnić o kwestię kodowania tekstów na temat traktatu. W rozdziałach poświęconych ortografii i fonologii [Kucała 1985: 16–25] występują znaki tzw. sławistycznego alfabetu fonetycznego – elektroniczna edycja komentarza do traktatu powinna



ten fakt uwzględnić. Mówiąc żartobliwie, geneza alfabetu sławistycznego ginie w pomroce dziejów – choć nie ma on chyba więcej niż 100 lat, pomimo ponawianych prób nie udało mi się ustalić, kto go stworzył i kiedy zaczął być używany, wielokrotnie powielane są natomiast wykazy znaków, dostępne również w internecie. Jednym z nich jest aneks do polskiej wersji Wikisłownika pod tytułem *Język polski – wymowa – głoski*<sup>10</sup>. Ta i podobne strony internetowe niejawnie określają sposób reprezentacji znaków alfabetu w standardzie Unicode, ponieważ w tym właśnie standardzie są one zapisane. Wykorzystywane są m.in. znaki LATIN SMALL LETTER EZH i GREEK SMALL LETTER CHI. Użycie litery LATIN SMALL LETTER EZH jest w pełni uzasadnione praktycznie i teoretycznie. Użycie litery GREEK SMALL LETTER CHI w witrynie jest w pełni uzasadnione z praktycznego punktu widzenia, bo niemal każdy użytkownik dysponuje fontem z tym znakiem. Z teoretycznego punktu widzenia lepszy byłby jednak znak LATIN SMALL LETTER CHI wprowadzony do standardu w wersji 7.0 z czerwca 2014 roku [por. NSAI 2012]. Dzięki temu zapis fonetyczny słowa składałby się w całości z pisma łacińskiego, bez „rozrywania” go znakami greckimi. Niestety aktualnie znak ten zawierają bardzo nieliczne fonty<sup>11</sup>.

Poniżej podsumuję informacje dotyczące standardu Unicode, uzupełniając je dodatkowo o kwestie bardziej perspektywiczne.

Standard Unicode jest bardzo obszerny, sama część opisowa najnowszej wersji (12.1.0 z 7 maja 2019 roku<sup>12</sup>) liczy 1018 stron, do tego dochodzą dodatki oraz tabele pokazujące 137 929 znaków zdefiniowanych w standardzie. Jest on kompromisem między założeniami teoretycznymi a potrzebami producentów oprogramowania oraz – raczej pośrednio – użytkowników tego oprogramowania. Czasami skutkiem kompromisu są ustalenia nieeleganckie lub nie do końca spójne.

W celu zakodowania – jak w przypadku traktatu Parkosza – znaku niewystępującego w standardzie można użyć następujących metod:

A. Transliteracja, czyli wykorzystanie już dostępnych znaków; oczywiście czytelnik musi być świadomy zastosowanej konwencji. Zaleta: tekst powinien praktycznie zawsze jakoś się wyświetlić, choć w przypadku rzadko używanych znaków może pojawić się potrzeba doinstalowania dodatkowych fontów. Tę właśnie metodę uznałem za najbardziej właściwą na obecnym etapie i opisałem w poprzedniej części artykułu.

10 <http://bit.ly/2TH3mga-AS> [dostęp: 10 maja 2019].

11 Patrz np. <https://www.fileformat.info/info/unicode/char/ab53/fontsupport.htm> [dostęp: 10 maja 2019]; może to zmienić się w przyszłości.

12 <http://www.unicode.org/versions/Unicode12.1.0/> [dostęp: 10 maja 2019].

## B. Wykorzystanie znaku z obszaru użytku prywatnego (PUA):

1. Wykorzystanie znaku już zdefiniowanego w ramach nieoficjalnej społeczności, takiej jak Medieval Unicode Font Initiative; ustalenia te są honorowane przez niewielką grupę twórców fontów, a ich fonty są łatwo dostępne, np. w standardowych dystrybucjach systemu Linux. Tak więc wyświetlenie tekstu też nie powinno sprawiać większego problemu, zwłaszcza że użytkownik może mieć już odpowiedni font zainstalowany. Możliwość ta została wykorzystana przeze mnie w kilku uzasadnionych przypadkach.

2. Wykorzystanie znaku zdefiniowanego całkowicie samodzielnie – to może w przyszłości okazać się koniecznością, bo działalność Medieval Unicode Font Initiative praktycznie zamarła, a inicjatywa LINCUA (*LINGuistic Corporate Use Area*), którą w 2012 roku zaproponował Andreas Stötzner, też wygląda na martwą<sup>13</sup>; szansą na jej ożywienie była propozycja, którą Stötzner złożył w lipcu 2018 roku, dotycząca włączenia się w ewentualne prace nad fontami dla dawnej polszczyzny<sup>14</sup>, która jednak nikogo nie zainteresowała. Wadą tego podejścia jest to, że w celu umożliwienia wyświetlenia tekstu niezbędne jest nie tylko stworzenie odpowiedniego fontu, ale i dostarczenie go użytkownikowi.

C. Podjęcie kroków w celu dodania znaku do standardu. Cytowane przeze mnie zaakceptowane propozycje dodania nowych znaków dość dobrze pokazują, jakiego typu argumenty przemawiają do organu konsorcjum Unicode podejmującego decyzje.

Warto zwrócić uwagę, że rozwój standardu Unicode odbywa się w zasadzie jawnie, choć niestety nie wszystkie dokumenty są dostępne publicznie. Formalnie rzecz biorąc, propozycję modyfikacji standardu, np. dodania nowego znaku, może zgłosić każdy. Stanowi to drastyczny kontrast w stosunku do biurokratyzowanej działalności normalizacyjnej, w której z nielicznymi wyjątkami propozycje mogą zgłaszać tylko tzw. komitety techniczne, a całe procedowanie jest poufne pod pretekstem ochrony własności intelektualnej. Konsorcjum Unicode powstało właśnie jako reakcja na nieudolność oficjalnej normalizacji w rozwiązywaniu naglących problemów stojących przed producentami oprogramowania; po burzliwych sporach osiągnięto kompromis polegający na systematycznym harmonizowaniu zestawu znaków standardu Unicode z tzw. Uniwersalnym Zestawem Znaków (ang. *Universal Character Set* – UCS) stanowiącym normę międzynarodową (ale nie europejską ani polską) o symbolu

13 Informacja o niej jest obecnie dostępna tylko w archiwum internetu, patrz <http://bit.ly/2XVTzRL-LINCUA> [dostęp: 10 maja 2019].

14 Patrz np. <http://bit.ly/2CjNpBS-GUST> [dostęp: 10 maja 2019] i [http://bit.ly/2HvfsBL\\_DARIAH-PL](http://bit.ly/2HvfsBL_DARIAH-PL) [dostęp: 10 maja 2019].

ISO/IEC 10646. Propozycje nowych znaków z reguły są zgłaszane jednocześnie do konsorcjum Unicode i odpowiedniego zespołu Międzynarodowej Organizacji Normalizacyjnej ISO, stąd w opisach bibliograficznych występują czasami oznaczenie ISO/IEC JTC1/SC2/WG2 – grupa robocza nr 2 podkomitetu nr 2 komitetu nr 1 wspólnego dla ISO i Międzynarodowej Komisji Elektrotechnicznej (IEC). Z praktycznego punktu widzenia norma ta jest jednak bez znaczenia, bo jest pod wieloma względami uboższa od standardu<sup>15</sup>.

Dla pełności obrazu warto dodać, że nie wszystkie znaki są w standardzie reprezentowane jednoznacznie, np. litera *ć* może być zapisana jako znak LATIN SMALL LETTER C WITH ACUTE lub jako sekwencja znaków LATIN SMALL LETTER C WITH ACUTE i znaku COMBINING ACUTE ACCENT. Zazwyczaj użytkownik nie jest tego świadomy, ponieważ oprogramowanie w razie potrzeby dokonuje odpowiedniej konwersji. Z teoretycznego punktu widzenia wskazane wydaje się jednak wprowadzenie pojęcia abstrahującego od konkretnej reprezentacji znaku. Takie obiekty proponuję nazywać tekstelami [por. np. Bień 2016b].

Chciałbym tutaj dodatkowo zwrócić uwagę na jedną mało znaną kwestię, a mianowicie na teoretyczną możliwość wykorzystania tzw. sekwencji wariacyjnych (ang. *variation sequences*, tłumaczenie jako np. *sekwencja wariantowa* wydaje mi się mylące). Odpowiedzi na często zadawane pytania<sup>16</sup> zawierają następujące pytanie i odpowiedź (swobodne tłumaczenie moje):

**Pytanie:** Do repertuaru historycznego systemu pisma chcę dodać pewien wariant istniejącego znaku. Należy złożyć propozycję dodania do standardu nowego znaku czy nowej sekwencji wariacyjnej?

**Odpowiedź:** Sekwencja wariacyjna pozwala zakodować warianty znaku bez kodowania każdego wariantu jako osobnego znaku. Jest to szczególnie pożyteczne wtedy, kiedy rozróżnienie tych wariantów nie zawsze jest potrzebne.

Ponieważ znak podstawowy jest częścią sekwencji wariacyjnej, można w tekście wyszukać wszystkie wystąpienia tego znaku niezależnie od tego, w jakim wariantcie on występuje, co byłoby o wiele trudniejsze, gdyby warianty były kodowane jako odrębne znaki. Jeśli zastąpienie wariantu przez znak podstawowy nie zniekształca zasadniczo treści tekstu, to sekwencja wariacyjna

---

15 Pojęcia normy i standardu są istotnie różne, ale nie jest to właściwe miejsce na szczegółową analizę.

16 <https://www.unicode.org/faq/vs.html#18> [dostęp: 10 maja 2019], także [bit.ly/2CjesVA-vsUnicode](http://bit.ly/2CjesVA-vsUnicode) [dostęp: 10 maja 2019].

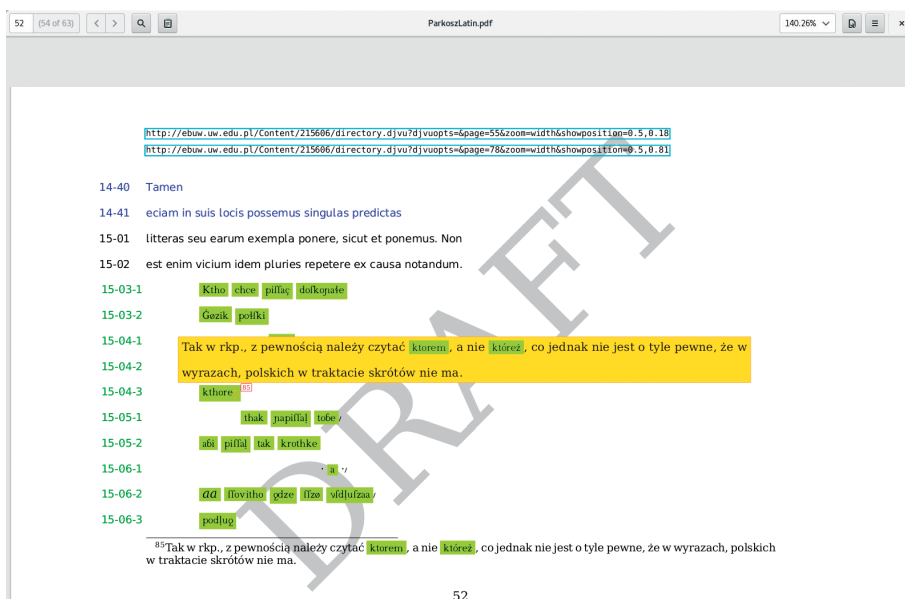
jest właściwym sposobem reprezentowania tego wariantu, i należy zaproponować dodanie do standardu właśnie sekwencji wariacyjnej.

Dla historycznych systemów pisma sekwencje wariacyjne są pożytecznym narzędziem, ponieważ pozwalają zakodować błędne lub rzadkie (ang. *nonce*) czcionki (ang. *glyphs*) w sposób ukazujący ich związek ze znakiem podstawowym. Sekwencje mogą również odzwierciedlać poglądy różnych badaczy, którzy mogą na różne sposoby postrzegać relację czcionki do znaku podstawowego. W dodatku nowe sekwencje mogą być dodawane sukcesywnie w miarę odkrywania nowych czcionek i nowych relacji do znaku podstawowego.

Niestety mechanizm sekwencji wariacyjnych jest w standardzie stosowany obecnie wyłącznie do dość egzotycznych dla nas języków i pism, zastosowanie go do historycznych tekstów polskich byłoby przedsięwzięciem pionierskim.

## 5. Forma edycji elektronicznej

Traktat jest pisany w sposób ciągły, praktycznie bez podziału na akapity. Dla zwiększenia czytelności tekstu Kucała w swojej edycji wprowadził – jak wspominałem wyżej – własny podział na akapity, który został uwzględniony w edycji elektronicznej. Dodatkowo uwzględniono również sugestie przedstawione w artykule *Jakub Parkoszowicz's Polish Mnemonic Verse about Polish Orthography from the 15th Century* [Wójcik, Wydra 2008].



Ilustracja 30. Fragment edycji elektronicznej

Zgodnie z zasadą *The best tool is the tool you know best* (autor tej zasady jest mi niestety nieznanym) edycję elektroniczną przygotowałem za pomocą systemu TeX, a konkretnie „silnika” LuaTeX i systemu autorskiego (tak tłumaczę ang. *document preparation system*) LaTeX2e. Wynik jest dostępny w publicznym repozytorium w postaci pliku PDF<sup>17</sup>, a także za pośrednictwem DOI: 10.13140/RG.2.2.12938.31685/1 w *ResearchGate*. Dokument zawiera przypisy powtórzone w formie tzw. *Optional Content Groups* (nazywane również warstwami), które wyświetlają się po najechaniu kursorem na odpowiedni fragment tekstu – pokazuje to il. 30; niektóre proste przeglądarki plików PDF mogą jednak tej funkcji nie obsługiwać.

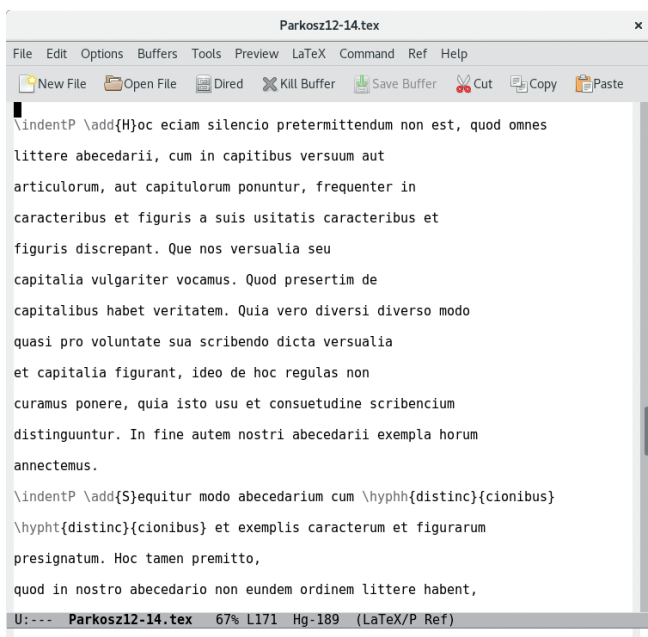
Edycja jest przeznaczona przede wszystkim do oglądania na ekranie, dlatego proporcje strony są zbliżone do proporcji typowego ekranu, a różne wyróżnienia oznaczane są za pomocą kolorów (np. słowa polskie są prezentowane na zielonym tle). Zostały też wprowadzone inne udogodnienia, takie jak hiperlinki do odpowiednich stron reprodukcji oryginału i jego odczytania w książce *Kucały* dostępnej w formacie DjVu w e-BUW (bibliotece cyfrowej Uniwersytetu Warszawskiego). Pierwotnie wystarczyło na nie kliknąć, obecnie trzeba je skopiować i w programie *djview4* wkleić do pola *Open Location*, co oczywiście jest kłopotliwe. Z tego powodu – gdy niedawno pojawiła się taka możliwość – linki do reprodukcji rękopisu w DjVu zastąpiłem linkami do skanów w formacie IIIF, które działają – przynajmniej w teorii – w każdej przeglądarce WWW<sup>18</sup>.

Wersji tej nie należy jednak traktować jako jedynej czy ostatecznej, z pewnością możliwe są różne ulepszenia, a także inne formy prezentacji.

---

17 Sekcja *Downloads* repozytorium <https://bitbucket.org/jsbien/parkosz-traktat> [dostęp: 10 maja 2019].

18 Umożliwił to Demian Katz, który odpowiednio zmodyfikował wspomniane wcześniej oprogramowanie *uv-app-starter*.



Ilustracja 31. Fragment źródeł edycji elektronicznej

Jednym z narzucających się ulepszeń jest skład tekstu z wykorzystaniem fontu uwzględniającego znaki alfabetu Parkosza – pośrednią korzyścią z ustalenia zasad transliteracji jest wskazanie pozycji, na których te znaki mają się w foncie znajdować (dodatkowo można znakom przydzielić również współrzędne kodowe w obszarze użytku prywatnego). Pierwszą próbę stworzenia takiego fontu, wzorowanego na ręcznie dopisywanych literach w książce Parkosza, wykonał mój student w 2013 roku<sup>19</sup> (por. il. 32), później samodzielnie próbowałem stworzyć font wzorowany na czcionkach z edycji Łosia<sup>20</sup> – por. il. 33. Nie jestem jednak zadowolony z rezultatu – *feci quod potui, faciant meliora potentes...*

*I a b c d e f g h i g k*  
*ł η η o ρ q r s t u v w*  
*y z ø g l n j b p v m ff*

Ilustracja 32. Repertuar fontu Parkosz (autor Maciej Strzelczyk)

19 <https://bitbucket.org/jsbien/parkosz-font-old> [dostęp: 10 maja 2019].

20 <https://bitbucket.org/jsbien/parkosz-font/> [dostęp: 10 maja 2019].

a c d e f i k o q r s t  
u x y z ç ø g ÿ f l p l  
b l m n v w q ff

Ilustracja 33. Repertuar fontu Parkosz1907

Podstawową formą edycji elektronicznej są jednak pliki źródłowe znajdujące się w sekcji *Source* wspomnianego repozytorium, a nie wspomniany plik PDF<sup>21</sup>. Jest ono zarządzane za pomocą systemu Mercurial, można więc obejrzeć nie tylko aktualne wersje plików, lecz także historię zmian. Przykładowy fragment pliku źródłowego jest podany na il. 31 (w trakcie składu wiersze numerowane są automatycznie); ze względu na brak ograniczeń licencyjnych pliki te mogą być wykorzystywane przez zainteresowanych w dowolny sposób.

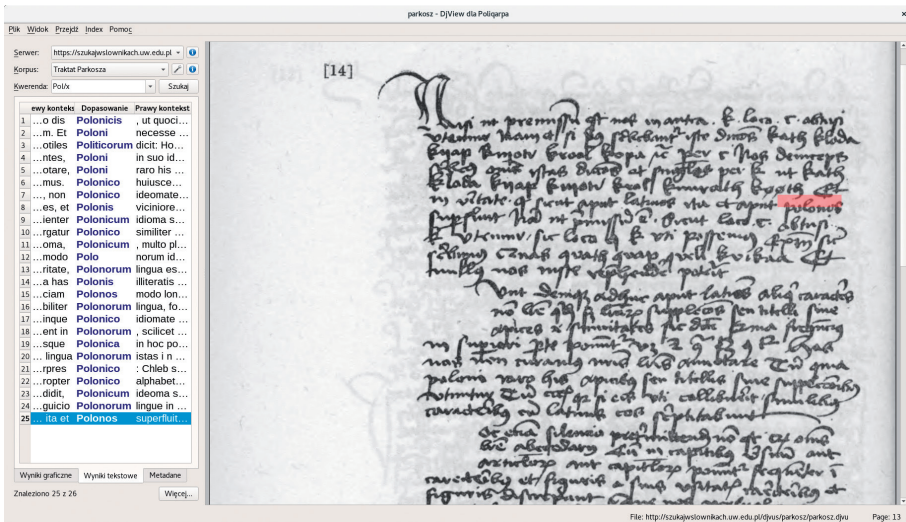
## 6. Alternatywne formy edycji elektronicznych

Na witrynie Fundacji Języka Polskiego<sup>22</sup> po pewnych wahaniach udostępniłem prowizoryczny „korpus” stworzony na podstawie tekstu traktatu, w przekonaniu, że pomimo jego wad może on okazać się dla kogoś przydatny. Dzięki temu możliwe jest przeszukiwanie tekstu za pomocą narzędzi korpusowych opisanych m.in. w tekście *Efficient search in hidden text of large DjVu documents* [Bień 2011] – por. il. 34. Możliwe jest również wyświetlenie odczytania fragmentu rękopisu wskazanego kursorem.

21 Dla bezpieczeństwa także w repozytorium [https://github.com/jsbien/Parkosz-traktat\\_git](https://github.com/jsbien/Parkosz-traktat_git) [dostęp: 10 maja 2019].

22 <https://szukajwslownikach.uw.edu.pl/parkosz/> [dostęp: 10 maja 2019].





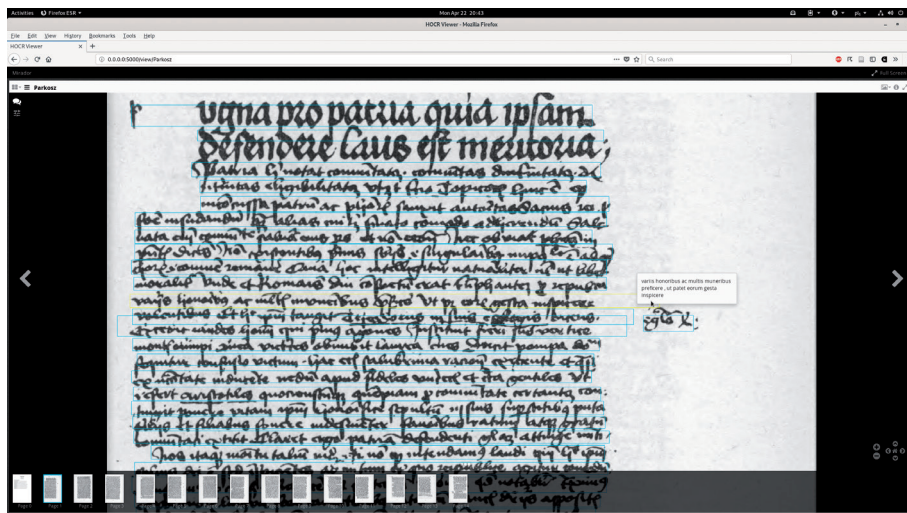
Ilustracja 34. Wyszukanie słów zawierających fragment *Pol*

Korpus ten został stworzony w następujący sposób (który może być zastosowany również do innych tekstów). Skan odczytania tekstu przez Kucalę został poddany automatycznemu rozpoznawaniu znaków za pomocą programu FineReader ze wskazaniem łaciny jako języka dokumentu; wynik był wysokiej jakości, choć program – zapewne korzystając ze słownika – czasami zmieniał pisownię słów (jeśli dobrze pamiętam, *eciam* podmieniał na nowsze *etiam* itp.); w zakresie wstawek polskojęzycznych wyniki – zgodnie z oczekiwaniami – były praktycznie bezużyteczne, zwłaszcza dla znaków wprowadzonych przez Parkosza. Wynikowy tekst za pomocą wspomnianego już wcześniej systemu Transkribus został przeze mnie zsynchronizowany ze skanem rękopisu z dokładnością do wierszy, bez korekty wstawek polskojęzycznych – proponowana tutaj transliteracja nie była jeszcze opracowana. Z Transkribusa wynik został wyeksportowany jako PDF zawierający skan z ukrytym tekstem transkrypcji. Plik ten został skonwertowany na plik DjVu z zachowaniem zarówno skanu, jak i tekstu ukrytego; mógł on już zostać użyty jako dane wyszukiwarki korpusowej – obecnie z pewnych względów technicznych można korzystać z niej wyłącznie za pomocą programu *djview4poliqarp*<sup>23</sup>.

Przywrócenie możliwości korzystania za pomocą przeglądarki WWW z funkcji dostępnych wcześniej dzięki formatowi DjVu wydaje się możliwe przy wykorzystaniu perspektywicznego – jak się wydaje – standardu IIIF,

23 <https://bitbucket.org/mrudolf/djview-poliqarp> [dostęp: 10 maja 2019].

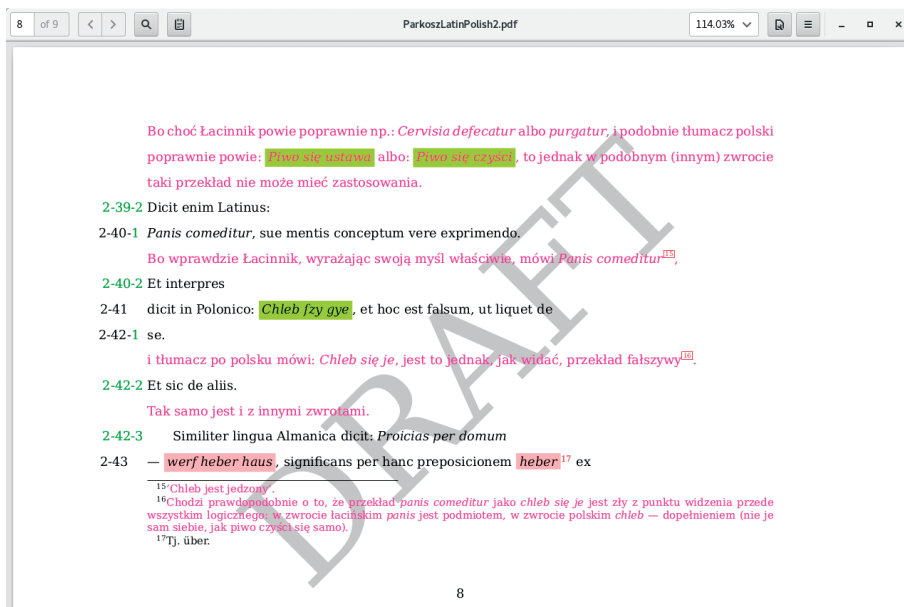
wymaga jednak pewnej pracy – il. 35 pokazuje wynik bardzo wstępnego eksperymentu<sup>24</sup>.



Ilustracja 35. Traktat z transkrypcją w przeglądarce IIF (Mirador)

Dla osób, które – tak jak ja – nie znają łaciny, ciekawe byłoby zestawienie tekstu łacińskiego (w odczytaniu lub w oryginale) z jego tłumaczeniem. Podjąłem pewne działania również w tym kierunku, a ich wynik widać na il. 36 (pliki źródłowe znajdują się w repozytorium).

<sup>24</sup> Wykorzystano oprogramowanie hocrviewer-mirador, którego autorem jest Johannes Baiter.



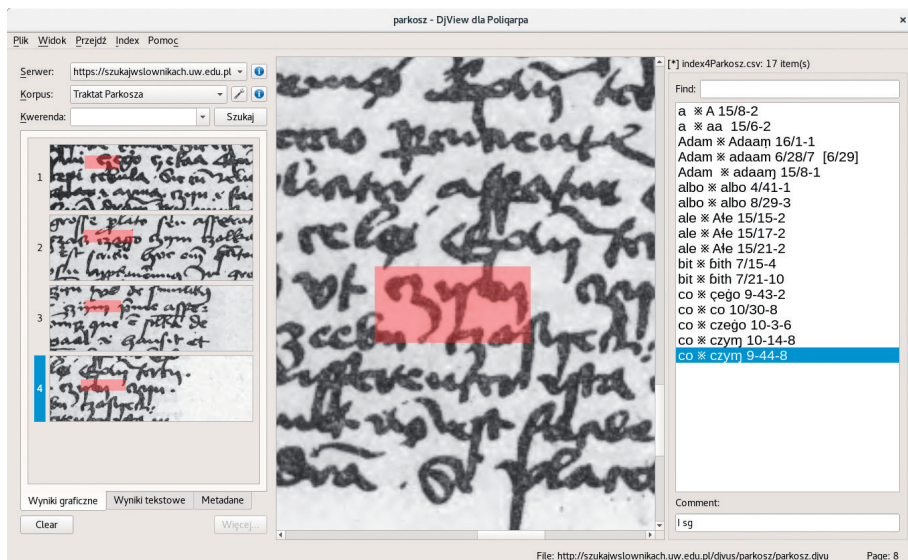
Ilustracja 36. Teksty równoległe: łaciński i polski

Zadanie realizowane w przyjęty sposób okazało się żmudne i nie miałem dostatecznej motywacji, żeby je dokończyć; być może należy użyć innych narzędzi, ale chętni mogą je też po prostu kontynuować.

Warto podkreślić, że w artykule tym koncentrowałem się na właściwym tekście traktatu. Dla takich elementów krytycznej edycji Kucaly jak *Wykaz wyrazów* [Kucala 1985: 111–126] właściwe wydają się narzędzia przygotowane dla elektronicznych lub zdigitalizowanych słowników. Ze swoimi studentami przeprowadziłem eksperymenty zmierzające do przekształcenia wykazu w słownik elektroniczny obsługiwany przez – program GoldenDict (wersja 1.5 lub wyższa<sup>25</sup>) – patrz il. 37.

25 Np. <https://github.com/goldendict/goldendict/wiki/Early-Access-Builds-for-Windows> [dostęp: 10 maja 2019].





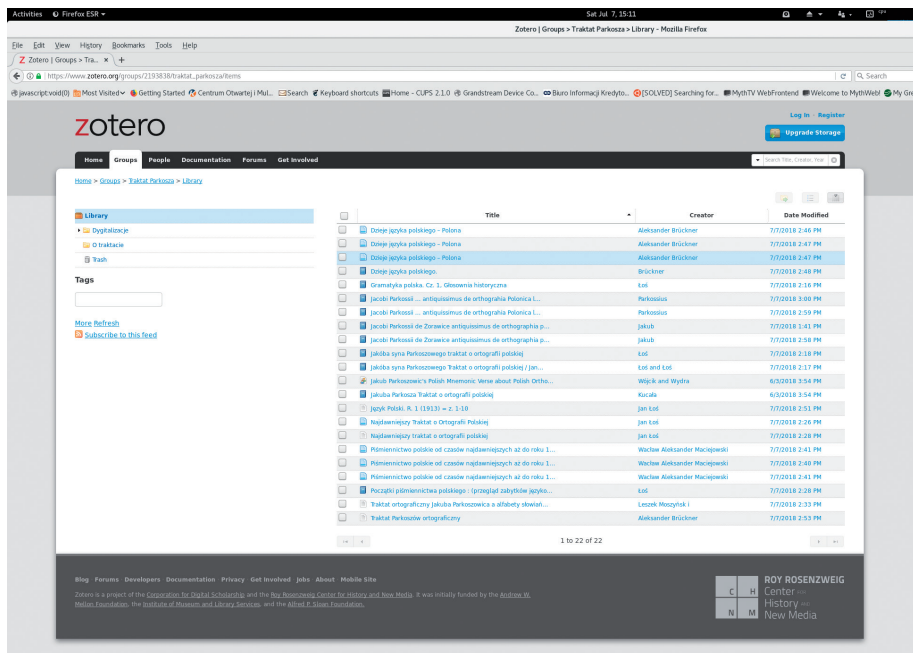
Ilustracja 38. Indeks wyrazów – program djview4poliqarp

Słownik i indeks dostępne są we wspomnianym repozytorium. Więcej informacji na temat korzystania z tego typu indeksów i ich tworzenia można znaleźć w mojej prezentacji *Elektroniczne indeksy leksykograficzne*<sup>26</sup> i w artykule *Elektroniczne indeksy fiszek słownikowych* [Bień 2018].

Na jeszcze inne potraktowanie zasługuje bibliografia. Miałaby ona istotnie większą wartość, gdyby została uzupełniona o informacje, które z przytoczonych pozycji zostały zdigitalizowane i gdzie się znajdują ich wersje cyfrowe. Dla niektórych publikacji informacje te są już dostępne jako efekt uboczny zadań realizowanych przez moich studentów (przedmiot *Repozytoria i biblioteki cyfrowe*). Udostępniłem je publicznie w tzw. kolekcji *Traktat Parkosza* systemu Zotero<sup>27</sup>; do kolekcji dodałem również wszystkie publikacje cytowane w niniejszym artykule – patrz il. 39. Niestety ten ciekawy system znam tylko bardzo powierzchownie. Dane uzyskane przez moich studentów udostępniłem w najprostszy technicznie sposób, czego efektem są różne niekonsekwencje. Być może znajdują się osoby – znające system lepiej ode mnie – które nie tylko usuną te usterki, ale również uzupełnią brakujące pozycje (system z założenia jest przystosowany do współpracy wielu osób nad kolekcjami).

26 Instytut Podstaw Informatyki PAN, seminarium *Przetwarzanie języka naturalnego*, 1 października 2018 roku; slajdy i nagranie wideo, patrz np. <https://www.slideshare.net/jsbien/jsb-i-linde181001ipi-117452985> [dostęp: 10 maja 2019].

27 [https://www.zotero.org/groups/2193838/traktat\\_parkosza/](https://www.zotero.org/groups/2193838/traktat_parkosza/) [dostęp: 10 maja 2019].



Ilustracja 39. Zotero – fragment kolekcji *Traktat Parkosza*

Warto odnotować, że jeszcze inną formą edycji elektronicznej rękopisu mogłyby być ćwiczenia paleograficzne. Moi studenci wykonali takie eksperymenty z programem Image Markup Tool (jego autorem jest Martin Holmes), który do tworzenia ćwiczeń paleograficznych – w 2010 roku lub wcześniej – przystosowała Marjorie Burghart<sup>28</sup> (por. il. 40). Narzędzie to zostało wykorzystane m.in. do stworzenia *Interactive Album of Mediaeval Palaeography*, o którym wspomina Aneta Pieniądz w swoim artykule [Pieniądz 2013] (podany tam adres internetowy jest już nieaktualny, witrynę zilustrowaną zrzutem ekranu trzeba obecnie szukać pod innym adresem<sup>29</sup>). Nawiasem mówiąc, jest ciekawe, że utworzona w 2013 roku witryna *In nomine Domini. Nauka paleografii łacińskiej*<sup>30</sup> używa identycznego interfejsu użytkownika jak witryna Burghart. Wydaje się, że obecnie do takich celów lepiej jest wykorzystać odpowiednie funkcje systemu Transkribus<sup>31</sup>.

28 [https://tapor.uvic.ca/~mholmes/image\\_markup/paleography.php](https://tapor.uvic.ca/~mholmes/image_markup/paleography.php) [dostęp: 10 maja 2019].

29 [http://paleographie.huma-num.fr/exo.php?id=gothique\\_livresque\\_Bible&!;en](http://paleographie.huma-num.fr/exo.php?id=gothique_livresque_Bible&!;en) [dostęp: 10 maja 2019].

30 <http://paleografia.home.amu.edu.pl> [dostęp: 10 maja 2019].

31 Por. <https://learn.transkribus.eu/> [dostęp: 10 maja 2019].



Ilustracja 40. Tekst Parkosza jako ćwiczenie paleograficzne

Osobnym zagadnieniem jest zwiększenie czytelności rękopisu. Jeden aspekt sprawy to zastosowanie technik, które są wykorzystywane rutynowo do wstępnego przygotowywania skanów do automatycznego rozpoznawania znaków: prostowanie, „odszumianie” czy „odplamianie” (and. *despeckling*, czyli usuwanie plam i skaz papieru), wreszcie binaryzacja przekształcająca obraz zawierający kolory lub odcienie szarości w obraz czarno-biały. Z różnych względów nie próbowałem wykonać żadnych kroków w tym kierunku.

Ciekawszym i trudniejszym aspektem sprawy jest rozdzielenie nachodzących na siebie wierszy. Wydaje mi się tylko kwestią czasu, kiedy będzie możliwe wykonanie tej operacji całkowicie lub w dużym stopniu automatycznie. Drogą do tego mogłoby być w szczególności odtworzenie ruchów ręki skryby. Nad takimi algorytmami pracuje się m.in. na potrzeby grafologii kryminalistycznej<sup>32</sup>, por. system MovAlyzeR firmy NeuroScript<sup>33</sup>, który jest reklamowany m.in. sugestią „zademonstruj ruchy pióra w sądzie”. O tym, że związki grafologii i paleografii są coraz bardziej ściśle, świadczy m.in. projekt *Digital forensics for historical documents. Cracking cold cases with new technology*

32 Warto zwrócić uwagę, że w języku angielskim *graphology* ma odmienne znaczenie i odnosi się do działalności uważanej za pseudonaukę, natomiast dla badań naukowych proponuje się termin *graphonomics*, por. <https://www.pertsinakis.com/graphonomics/> [dostęp: 10 maja 2019].

33 <http://www.neuroscript.net/forensics.php> [dostęp: 10 maja 2019].



(Cyfrowa kryminalistyka dla historycznych dokumentów. Rozwiązywanie starych problemów nowymi metodami)<sup>34</sup>.

Dla pełności obrazu warto wspomnieć, że ideałem wydaje się zastąpienie grafiki rastrowej grafiką wektorową. Jeden z moich studentów – pracujący jako grafik komputerowy – przekształcił na zapis wektorowy wybrane litery rękopisu Parkosza, dzięki czemu mogą one być prezentowane w dowolnym powiększeniu bez utraty jakości; wynik ten jest również dostępny publicznie we wspomnianym repozytorium. Taka ręczna konwersja nie jest jednak rozwiązaniem perspektywicznym.

## 7. Uwagi końcowe

Od wielu lat bardzo rozpowszechnionym standardem reprezentowania tekstów do celów badawczych są rekomendacje *Text Encoding Initiative* [por. np. Burnard 2015]. Być może w przyszłości warto będzie w tej formie zapisać również traktat Parkosza, ale – według mojej oceny – w chwili obecnej dałoby to znikome korzyści ze względu na to, że nieliczne dostępne oprogramowanie nie ma moim zdaniem widocznej przewagi nad rozwiązaniami prezentowanymi w tym artykule, chyba że chodziłoby o istotnie odmienne cele. Nie wykluczam, że np. *Versioning Machine*<sup>35</sup> mogłoby być właściwym narzędziem do porównywania wszystkich trzech edycji traktatu, ale podobne możliwości daje chyba przeglądarka *Mirador* dla standardu *IIIF*.

Na temat optymalnej formy edycji elektronicznej toczą się obecnie ożywione dyskusje [por. np. *Digital Scholarly Editions as Interfaces. Abstracts and Programme* 2016]. Nie zamierzam brać w nich udziału, ale chciałbym odnotować, że bliskie są mi poglądy przedstawione przez Petera Robinsona w referacie *Why Interfaces Do Not and Should Not Matter for Scholarly Digital Editions* [Robinson 2016], w szczególności teza, że najistotniejsze są dokładność i poprawność danych.

Wybitny polski informatyk prof. Andrzej Salwicki głosił tezę, że w nauce obowiązuje „partyzancka zasada”, która brzmi krótko: „Wnioskodawca wykonuje!” (być może formułował ją nieco inaczej). Zasada ta obowiązuje od dawna również w szeroko rozumianej informatyce – wspomniany wcześniej przeze mnie system *Mercurial* jest jednym z wielu narzędzi pozwalających tę zasadę wygodnie realizować w praktyce. Osoba proponująca zmiany w repozytorium, których właściciel repozytorium nie chce lub nie może (np. z braku czasu)

34 <https://www.huygens.knaw.nl/digital-forensics-for-historical-documents/?lang=en> [dostęp: 10 maja 2019].

35 <http://v-machine.org/> [dostęp: 10 maja 2019].

uwzględnić, tworzy nowe identyczne repozytorium (tzw. *fork*), wprowadza swoje zmiany i zgłasza do właściciela oryginału tzw. *pull request*. Jeśli właściciel je zaakceptuje i zaimportuje zmiany, zasób jest nadal rozwijany w oryginalnym repozytorium. Jeśli nie, zasoby te są rozwijane równoległe, a użytkownicy mają możliwość wyboru.

Oczywiście tworzenie własnego repozytorium to środek radykalny, stosowany tylko wtedy, gdy zostaną wyczerpane inne możliwości. Zwykle korzysta się z dostarczonych przez oprogramowanie repozytorium mechanizmów komunikacji, pozwalających na zgłaszanie problemów i postulatów, komentowanie i śledzenie podejmowanych w ich sprawie decyzji. Często z repozytorium jest połączone wiki, przeznaczone do przechowywania uzupełniających informacji i nieformalnych dyskusji. Takie wiki utworzyłem również dla traktatu Parkosza<sup>36</sup> i umieściłem w nim kilka informacji (część z nich przenieśliem później do Zotero).

Repozytoria tego typu są coraz częściej stosowane również w projektach humanistycznych i przypuszczam, że z czasem edycje elektroniczne też będą powszechnie w ten sposób udostępniane [por. np. Spiro 2016; Reeve 2016].

**Omówione w artykule<sup>37</sup> repozytoria dotyczące traktatu Parkosza nie są już przeze mnie rozwijane, ale będę się starał, aby były dostępne w internecie przez rok od publikacji niniejszego artykułu, potem zostaną zlikwidowane. Zainteresowane osoby proszone są więc o ich skopiowanie lub zgłoszenie się do mnie w celu przejęcia administracji repozytoriami.**

## Bibliografia

- Andreev Aleksandr, Shardt Yuri, Simmons Nikita (2013), *Proposal to Encode Combining Half Marks Used for Cyrillic Supralineation*. Spraw. tech. L2/13-139. ISO/IEC JTC 1/SC 2/WG 2, <http://std.dkuug.dk/jtc1/sc2/wg2/docs/n4475.pdf> [dostęp: 10 maja 2019].
- Apel Willi (1969), *Harvard Dictionary of Music*, Harvard University Press.

36 <https://bitbucket.org/jsbien/parkosz-traktat/wiki/> [dostęp: 10 maja 2019].

37 Niniejszy artykuł został przygotowany – jak wszystkie moje artykuły – za pomocą wspomnianego wcześniej systemu TeX. Na życzenie Redakcji został on skonwertowany do formatu Worda (konkretnie do akceptowanego przez Worda formatu ODT) za pomocą programu make4ht. Program ten wymagał jednak pewnych poprawek i zmian, które wykonał jego autor Michal Hoflich. Udzielił on mi również wielu pożytecznych rad – jestem mu za to bardzo wdzięczny.

- Bandtkie Jerzy Samuel, Raczyński Edward, red. (1830), *Jacobi Parkossii de Żorawice antiquissimus de orthographia polonica libellus*, Wilh. Deckeri et Societatis, Posnaniae, <http://www.wbc.poznan.pl/publication/115430> [dostęp: 10 maja 2019].
- Bień Janusz S. (2004), *Standard Unicode 4.0. Wybrane pojęcia i terminy*, „Biuletyn GUST”, t. 20, s. 9–14, <https://sunsite.icm.edu.pl/pub/GUST/bulletin/20/jsb04.pdf> [dostęp: 10 maja 2019].
- Bień Janusz S. (2011), *Efficient search in hidden text of large DjVu documents*, w: *Advanced Language Technologies for Digital Libraries*, red. Raffaella Bernardi i in., Lecture Notes in Computer Science (Theoretical Computer Science and General Issues) 6699. Springer, s. 1–14, [https://link.springer.com/chapter/10.1007/978-3-642-23160-5\\_1](https://link.springer.com/chapter/10.1007/978-3-642-23160-5_1) [dostęp: 10 maja 2019].
- Bień Janusz S. (2014), *The IMPACT project Polish Ground-Truth texts as a DjVu corpus*, „Cognitive Studies | Études Cognitives”, t. 14, s. 75–84, <https://ispan.waw.pl/journals/index.php/cs-ec/article/view/cs.2014.008> [dostęp: 10 maja 2019].
- Bień Janusz S. (2016a), *Elektroniczne indeksy fiszek słownikowych*, „Kwartalnik Językoznawczy”, z. 2, s. 12, [http://kwartjez.amu.edu.pl/teksty/teksty2016\\_2\\_26/Bien.pdf](http://kwartjez.amu.edu.pl/teksty/teksty2016_2_26/Bien.pdf) [dostęp: 10 maja 2019].
- Bień Janusz S. (2016b), *Problemy kodowania znaków w korpusach historycznych*, w: *Semantyka a konfrontacja językowa*, red. Danuta Roszko, Joanna Satola-Staśko-wiak, t. 5, Instytut Slawistyki PAN, Warszawa, s. 67–76.
- Bień Janusz S. (2016c), *Traktat Parkosza w epoce prawników i Internetu*, w: *Dawne z nowym łącząc... In memoriam Mariani Kucala*, red. Joanna Klimek-Grądzka, Małgorzata Nowak, Towarzystwo Naukowe KUL i Katolicki Uniwersytet Lubelski Jana Pawła II, Lublin, s. 389–396.
- Burnard Lou (2015), *Czym jest Text Encoding Initiative?*, przeł. Joanna Bilińska, DELab UW, Warszawa, <http://bit.ly/2Vf2hYK-TEI> [dostęp: 10 maja 2019].
- Cappelli Adriano (1912), *Lexicon Abbreviaturarum: dizionario di abbreviature Latine ed Italiane usate nelle carte e codici specialmente del Medio-Evo*. 2th. Wersja internetowa. Ulrico Hoepli, Milan, <http://www.hist.msu.ru/Departments/Medieval/Cappelli> [dostęp: 10 maja 2019].
- Cappelli Adriano (1961), *Lexicon Abbreviaturarum: dizionario di abbreviature Latine ed Italiane usate nelle carte e codici specialmente del Medio-Evo*. 6th. Interakcyjna wersja internetowa. Ulrico Hoepli, Milan, <https://www.adfontes.uzh.ch/en/ressourcen/abkuerzungen/cappelli-online> [dostęp: 10 maja 2019].
- Constable Peter (2003), *Proposal to Encode Additional Phonetic Symbols in the UCS*. Spraw. tech. L2/03-190. ISO/IEC JTC 1/SC 2/WG 2, <http://www.unicode.org/L2/L2003/03190-add-phon-syms.pdf> [dostęp: 10 maja 2019].
- Digital Scholarly Editions as Interfaces. Abstracts and Programme* (2016), University of Graz, <http://bit.ly/2WxMCft-Graz16> [dostęp: 10 maja 2019].

- Elliot John, Brand Peter, Harvey Chris (2004), *Proposal to Add Four SENĆOŦEN Latin Characters*. Spraw. tech. L2/04-170. ISO/IEC JTC1/SC2/WG2, <https://www.unicode.org/L2/L2004/04170-sencoten.pdf> [dostęp: 10 maja 2019].
- Everson Michael (2004), *Proposal to add Latin letters and a Greek symbol to the UCS*. Spraw. tech. L2/06-266. ISO/IEC JTC1/SC2/WG2, <http://std.dkuug.dk/jtc1/sc2/wg2/docs/n3122.pdf> [dostęp: 10 maja 2019].
- Everson Michael i in. (2005), *Preliminary proposal to add medievalist characters to the UCS*. Spraw. tech. N2957. ISO/IEC JTC1/SC2/WG2, <http://www.unicode.org/L2/L2005/05183-n2957-mufi.pdf> [dostęp: 10 maja 2019].
- Everson Michael i in. (2007), *Proposal to add additional characters for Coptic and Latin to the UCS*. Spraw. tech. L2/07-085R. ISO/IEC JTC 1/SC 2/WG 2, <https://www.unicode.org/L2/L2007/07085r-n3222r-coptic-adds.pdf> [dostęp: 10 maja 2019].
- Freytag Asmus, Whistler Ken (2005), *Proposal to add nine lowercase characters*. Spraw. tech. N2942. ISO/IEC JTC1/SC2/WG2, <http://std.dkuug.dk/jtc1/sc2/wg2/docs/n2942.doc> [dostęp: 10 maja 2019].
- Gieysztor Aleksander (2009), *Zarys dziejów pisma łacińskiego*, Wydawnictwo Naukowe PWN, Warszawa.
- Górski Konrad, red. (1955), *Zasady wydawania tekstów staropolskich: projekt*, Zakład im. Ossolińskich – Wydawnictwo Polskiej Akademii Nauk, Wrocław.
- Haugen Odd Einar, red. (2003), *MUFI character recommendation version 1.0*. Medieval Unicode Font Initiative, <http://bora.uib.no/bitstream/handle/1956/2004/MUFI-standard-1.0.pdf> [dostęp: 10 maja 2019].
- Haugen Odd Einar, red. (2015), *MUFI character recommendation version 4.0*. Medieval Unicode Font Initiative, <http://hdl.handle.net/1956/10699> [dostęp: 10 maja 2019].
- Houston Keith (2015), *Ciemne typki. Sekretne życie znaków typograficznych*, d2d.pl, Kraków.
- Jensen Joshua M., Pentzlin Karl (2012), *Proposal to encode a Latin Capital Letter L with Belt*. Spraw. tech. N4228. ISO/IEC JTC1/SC2/WG2, <http://unicode.org/L2/L2012/12080-l-with-belt.pdf> [dostęp: 10 maja 2019].
- Jodłowski Stanisław (1979), *Losy polskiej ortografii*, Państwowe Wydawnictwo Naukowe, Warszawa.
- Klemensiewicz Zenon (1961), *Historia języka polskiego*, t. 1: *Doba staropolska*, Państwowe Wydawnictwo Naukowe, Warszawa.
- Korpela Jukka K. (2006), *Unicode Explained*, O'Reilly Media, Inc.
- Kucała Marian (1985), *Jakuba Parkosza Traktat o ortografii polskiej*, Państwowe Wydawnictwo Naukowe, Warszawa, <http://ebuw.uw.edu.pl/publication/220504> [dostęp: 10 maja 2019].

- Lisowski Tomasz (2010), *Pisownia polska. Główne fazy rozwoju*, „Kwartalnik Językoznawczy”, z. 3–4, s. 117–130, [http://www.kwartjez.amu.edu.pl/teksty/teksty2010\\_3-4\\_3-4/Lisowski.pdf](http://www.kwartjez.amu.edu.pl/teksty/teksty2010_3-4_3-4/Lisowski.pdf) [dostęp: 10 maja 2019].
- Łoś Jan (1907), *Jakóba syna Parkoszewego traktat o ortografii polskiej*. Osobne odbicie z t. II Materiałów i prac Komisji językowej Akademii Umiejętności w Krakowie, Kraków, <http://ebuw.uw.edu.pl/publication/238219> [dostęp: 10 maja 2019].
- Maciejowski Waław Aleksander (1852), *Piśmiennictwo polskie od czasów najdawniejszych aż do roku 1830*, t. 3 [Dodatki], Nakładem i drukiem S. Orgelbranda, Warszawa, <http://bc.wbp.lublin.pl/publication/15543> [dostęp: 10 maja 2019].
- Michałowska Teresa (1995), *Średniowiecze*, Wydawnictwo Naukowe PWN, Warszawa.
- Mika Tomasz (2015), *O temacie konferencji „Jak wydawać teksty dawne?”*, <http://classica-mediaevalia.pl/wp-content/uploads/2015/07/SSJ2-informacje.pdf> [dostęp: 10 maja 2019].
- Mühlberger Günter, Colutto Sebastian, Kahle Philip (2017), *Preprint: Handwritten Text Recognition (HTR) of Historical Documents as a Shared Task for Archivists, Computer Scientists and Humanities Scholars. The Model of a Transcription & Recognition Platform (TRP)*, <http://bit.ly/2VZyiEU-TranskribusHTR> [dostęp: 10 maja 2019].
- NSAI, National Standards Authority of Ireland (2012), *Request to change the names of three Teuthonista characters under ballot*. Spraw. tech. N4296. ISO/IEC JTC1/SC2/WG2, <https://www.unicode.org/L2/L2012/12269-n4296-latin-chi.pdf> [dostęp: 10 maja 2019].
- Pieniąż Aneta (2013), *Narzędzia internetowe w nauczaniu paleografii*, „Studia Źródłoznawcze”, t. 51, s. 131–143, <http://rcin.org.pl/ihpan/publication/53130> [dostęp: 10 maja 2019].
- Priest Lorna A. (2004a), *Revised Proposal to Encode Additional Latin Orthographic Characters*. Spraw. tech. L2/04-372R. ISO/IEC JTC 1/SC 2/WG 2, <http://www.unicode.org/L2/L2004/04372r-latin-orth.pdf> [dostęp: 10 maja 2019].
- Priest Lorna A. (2004b), *Revised Proposal to Encode Additional Latin Phonetic and Orthographic Characters*. Spraw. tech. L2/04-348. ISO/IEC JTC 1/SC 2/WG 2, <http://www.unicode.org/L2/L2004/04348-phonetic.pdf> [dostęp: 10 maja 2019].
- Priest Lorna A., Constable Peter G. (2005), *Proposal to Encode Additional Latin Phonetic and Orthographic Characters*. Spraw. tech. L2/05-097R. ISO/IEC JTC 1/SC 2/WG 2, <http://unicode.org/L2/L2005/05097r2-latin-orth.pdf> [dostęp: 10 maja 2019].
- Reeve Jonathan (2016), *Git-Lit: an Application of Distributed Version Control Technology toward the Creation of 50,000 Digital Scholarly Editions*, w: *Digital Humanities 2016: Conference Abstracts*, Jagiellonian University, Pedagogical University, s. 657–658, <http://dh2016.adho.org/static/data/455.html> [dostęp: 10 maja 2019].

- Robinson Peter (2016), *Why Interfaces Do Not and Should Not Matter for Scholarly Digital Editions*, w: *Digital Scholarly Editions as Interfaces. Abstracts and Programme*, University of Graz, <http://bit.ly/2HtFghD-Robinson-interfaces> [dostęp: 10 maja 2019].
- Seaward Louise, Kallio Maria (2017), *Transkribus: Handwritten Text Recognition technology for historical documents*, w: *Digital Humanities 2017*, Alliance of Digital Humanities Organizations, Montreal, Canada, <https://dh2017.adho.org/abstracts/649/649.pdf> [dostęp: 10 maja 2019].
- Semkowicz Władysław Aleksander (2007), *Paleografia łacińska*, Universitas, Kraków.
- Seroczyński Grzegorz (2014–2017), „*Traktat o ortografii polskiej*”, w: *Dawne ortografie, gramatyki i podręczniki języka polskiego. Internetowe kompendium edukacyjne*, red. Wanda Decyk-Zięba i in., Uniwersytet Warszawski, <https://gramatyki.uw.edu.pl/book/3> [dostęp: 10 maja 2019].
- Seroczyński Grzegorz (2014), *Jakuba Parkosza „Traktat o ortografii polskiej”*, „Poradnik Językowy”, z. 4, s. 98–103, <https://www.ceeol.com/search/article-detail?id=88873> [dostęp: 10 maja 2019].
- Spiro Lisa (2016), *Evaluating GitHub as a Platform of Knowledge for the Humanities*, w: *Digital Humanities 2016: Conference Abstracts*, Jagiellonian University, Pedagogical University, s. 688–690, <http://dh2016.adho.org/abstracts/225> [dostęp: 10 maja 2019].
- Wikipedia (2018a), *Jakub Parkoszowicz – Wikipedia: wolna encyklopedia*, [https://pl.wikipedia.org/wiki/Jakub\\_Parkoszowicz?oldid=53462555](https://pl.wikipedia.org/wiki/Jakub_Parkoszowicz?oldid=53462555) [dostęp: 14 marca 2019].
- Wikipedia (2018b), *L with bar – Wikipedia, The Free Encyclopedia*, [https://en.wikipedia.org/w/index.php?title=L\\_with\\_bar&oldid;=847452659](https://en.wikipedia.org/w/index.php?title=L_with_bar&oldid;=847452659) [dostęp: 10 maja 2019].
- Wikipedia (2019a), *Jakub Parkoszowicz – Wikipedia: wolna encyklopedia*, [https://pl.wikipedia.org/wiki/Jakub\\_Parkoszowicz?oldid=56469598](https://pl.wikipedia.org/wiki/Jakub_Parkoszowicz?oldid=56469598) [dostęp: 10 maja 2019].
- Wikipedia (2019b), *Ÿ – Wikipedia, Die freie Enzyklopädie*, <https://de.wikipedia.org/w/index.php?title=%C5%B8&oldid;=185744142> [dostęp: 10 maja 2019].
- Wikipedia contributors (2019a), *Cedilla – Wikipedia, The Free Encyclopedia*, <https://en.wikipedia.org/w/index.php?title=Cedilla&oldid;=893809898> [dostęp: 10 maja 2019].
- Wikipedia contributors (2019b), *ISO 5426 – Wikipedia, The Free Encyclopedia*, [https://en.wikipedia.org/w/index.php?title=ISO\\_5426&oldid;=893056043](https://en.wikipedia.org/w/index.php?title=ISO_5426&oldid;=893056043) [dostęp: 10 maja 2019].
- Wikipedia contributors (2019c), *Pan-Nigerian alphabet – Wikipedia, The Free Encyclopedia*, [https://en.wikipedia.org/w/index.php?title=Pan-Nigerian\\_alphabet&oldid;=894928524](https://en.wikipedia.org/w/index.php?title=Pan-Nigerian_alphabet&oldid;=894928524) [dostęp: 10 maja 2019].



Wikipedia contributors (2019d), *Viz.* – *Wikipedia, The Free Encyclopedia*, <https://en.wikipedia.org/w/index.php?title=Viz.&oldid;=894492673> [dostęp: 10 maja 2019].

Wójcik Rafał, Wiesław Wydra (2008), *Jakub Parkoszwic's Polish Mnemonic Verse about Polish Orthography from the 15th Century*, w: *Culture of Memory in East Central Europe in the Late Middle Ages and the Early Modern Period*, red. Rafał Wójcik, Biblioteka Uniwersytecka, Poznań, s. 119–127, <http://hdl.handle.net/10593/6037> [dostęp: 10 maja 2019].

Janusz S. Bień

### **Parkosz's treatise. An experimental electronic edition**

The fifteenth-century Latin manuscript presenting the proposal of Polish spelling formulated by Jakub Parkosz is an interesting challenge to the editors primarily because it introduces special characters that were not used later. The text of the treatise was presented and discussed in a book by Marian Kucala (available under open licence) which was used as the basis for an experimental electronic edition of the treatise in the form of an interactive PDF file, also available as a LuaLaTeX source. For transliteration, the electronic edition uses the existing Unicode letters with some additions from the private use area of the Medieval Unicode Font Initiative. This paper presents the rationale for the transliteration and discusses some possible alternative forms of electronic editions.

**KEYWORDS:** Polish spelling, history of Polish, Parkosz, Unicode.

**prof. dr hab. Janusz S. Bień** – emeryt, ostatnio zatrudniony w Katedrze Lingwistyki Formalnej Wydziału Neofilologii Uniwersytetu Warszawskiego; informatyk i lingwista (z wykształcenia matematyk); nadal interesuje się dygitalizacją dawnych tekstów polskich.