

# Charty zostały... czyli o generowaniu wierszy sylabicznych

Łukasz Dębowski

ldebowsk@ipipan.waw.pl



Instytut Podstaw Informatyki PAN

**Rygorystyczny Opis Języka Naturalnego, 17.12.2003, 15.12.2004**

# Zarys referatu

1. Formalny opis wiersza.
  - definicja strofy, metrum i rymu
2. Transkrypcja fonetyczna polszczyzny pisanej.
  - pakiet `speak::Polish.pm`
3. Ekscerpcja randomizowana.
  - program `turpis.pl`
4. Próby wyjścia poza ekscerpcję.

# Formalny opis wiersza

## ■ Typy strofy:

limeryk: A3a\_A3a\_A2b\_A2b\_A3a

saficka: 11a\_11b\_11a\_05b

tercyna: 11a\_11b\_11a\_;\_11b\_11c\_11b\_;\_...

triolet: A3a\_A3b\_A3a\_!1\_A3a\_A3b\_!1\_!2

A3, A2, 11, 05 – typy metrów, a, b, c... – typy rymów,

!k – powtórzenie k-tego wersu, ; – odstęp.

## ■ Typy metrów:

A3: AUSUUSUUSU?      11: AAASAAAAASU

A2: AUSUUSU?      05: AAASU

A – 1 sylaba dowolna, U – 1 sylaba nieakcentowana,

U? – (hiper)kataleksa, S – 1 sylaba akcentowana.

# Definicja strofy, typu metrum i typu rymu

**Strofa** — ciąg wersów o danym typie metrum i typie rymu

**Wers** — ciąg słów mogący się pojawić w tekście

**Typ metrum** — wyrażenie pasujące do metrum wersu

**Typ rymu** — zmienna globalna równa rymowi wersu

**Metrum wersu** — funkcja transkrypcji fonetycznej wersu

**Rym wersu** — funkcja transkrypcji fonetycznej wersu

**Transkrypcja fonetyczna wersu** — pewna funkcja wersu

- Każdy typ rymu ukonkretnia się do innego rymu.
- Ostatnie słowa niepowtórzonych wersów o tym samym typie rymu muszą być różne.

## Pakiet `Speak::Polish.pm`

Biblioteka funkcji (napisana w Perlu):

`Już konie w stajnię wzięto, już im hojnie dano,`

- przybliżona transkrypcja fonetyczna

`[juš końe fstajńe vżento juž jym xojńe dano]`

- akcentowana transkrypcja fonetyczna

`[j?(u)š k!(o)ń.(e) fst!(a)jń.(e)  
vż!(e)nt.(o) j?(u)ž j?(y)m x!(o)jń.(e)  
d!(a)n.(o)]`

- metrum i rym

`[?!..!..??!.!.] [ano]`

`www.ipipan.waw.pl/~ldebowsk/software.html`

Saša Rosen napisał także `Speak::Czech.pm`.

# Przykładowa aplikacja `sortrhymes.pl`

## Wykrywanie błędnej średniówki w *Panu Tadeuszu*:

[?!?!?!?!..!.] Raz w bitwie, gdzie podział się? szukać Bonaparta! -  
[!..!.??...!.] Sąsiadka zaśmiała się, lecz nie powiedziała,  
[!..!.?!..!.] Inaczej bawiono się w drugim końcu stoła,  
[!..!.?!..!.] Tadeusz przyglądał się nieznannej osobie;  
[!..!.?!..!.] Od dni kilku zbiera się na sądy graniczne,  
[!..!.?!..!.] A chłopci żegnali się mówiąc: że po świecie  
[!..!.?!..!.] Często zastanawiał się w powiatowym mieście;  
[:..!.?!..!.] Wydawała, że mnich ten nie zawsze w kapturze  
[!..?!?!..!.] Rumienił się, serce mu biło nadzwyczajnie;  
[!..!.?!..!.] Ujrzała, zaśmiała się i klasnęła w dłonie,  
[?!..!.?!..!.] Sam Podczaszyc, jaki był, opisywać długo,  
[?!..!.?!..!.] Lub wymowy uczyć się w paryskich kawiarniach.  
[!..!.?!..!.] Z ziemi włoskiej stara się przyciągnąć do Polski,  
[?!..!.?!..!.] Rzekł z uśmiechem, a był to uśmiech jadowity:  
[!..!.?!..!.] I dwie twarze w jeden się rumieniec oblekły.

**W 98% przypadków akcent jest na sylabie 6. a nie 5.**

# Ekscerpcja randomizowana

1. Weź dowolny korpus tekstów i sporządź tabelę wszystkich ciągów słów, które się pasują do zdefiniowanych typów metrów:

A3	akjem	brakjem	dwa tygodnie tłumacząc to brakiem
A3	akjem	dżobakjem	środowisko z józefem dziobakiem
A2	ytu	kredytu	udzielenie kredytu
A2	ytu	tu	chudsze chyba że i tu
A2	yva	leńyva	jest istotą leniwą
A2	yva	utśażllyva	ale dość uciążliwą

2. Dla zadanego typu strofy losuj z tabeli kolejne wersy tak, by metra pasowały, rymy powtarzały się odpowiednio, zaś słowa końcowe występowały jednokrotnie.

Procedura losowania może się nie powieść z braku wersów — wówczas losowanie powtarzamy do skutku.

# Program turpis.pl

- TURPIS (The Ultimate Retriever of Poetry Ignoring Sense)

`www.ipipan.waw.pl/~ldebowsk/software.html`

- Obecnie tworzy po polsku i po czesku:

Alexandr Rosen, *Vínek z veršů TURPISových*, 2004.

(w tomie okolicznościowym na cześć Vladimíra Petkeviča)

- Lista obsługiwanych form poetyckich (około 20):

dystych, haiku, limeryk, oktawa, „pantadeusz”, saficka, sekstyna, sonet, tanka, tercyna, triolet, vilanella, kwartet, ballada, decyma, nona, tercet, tyrada, rondel, pantum



## Dzieła wybrane (z *Gospodarstwa*)

widać że okolica obfita  
o piękności metrykę nie pyta  
co by rzekł wojewoda  
teraz nie wiem czy moda  
raz wraz skrzypi i napój w koryta



opatrzonych w broszurki i w różne  
oczy wkoło obracał ostróżne  
starych więcej książkowej  
drugim końcu do nowej  
lub bez nogi przyjąwszy jałmużnę

## Dzieła wybrane (z *Gospodarstwa*)

jak refektarz z wypukłym sklepieniem  
sędzia milczał on jeszcze skinieniem  
śród nich jedna króluje  
widzę i opisuję  
że za pierwszym na miejsce wejrzeniem



zacność domu fortuny szczodroty  
tuż myśliwców herbowne klejnoty  
panny córki choć stary  
nagle pierwsze zamiary  
krucze długie zwijały się sploty

## Dzieła wybrane (z *Gospodarstwa*)

może by była największa różnica  
bo już sam wewnątrz siebie czuł choroby  
cały las wiąże w jedno i jakoby  
znalazł podobne oczy usta lica

nim leży fedon i żywot katona  
kropelkę wina w szklankę panny róży  
mu jak ołtarzyk złoty zawsze służy  
brzęczało ciągle jako naprzykrzona

pamiętam chociaż byłem wtenczas małe  
wicher rozerwie i ręce pod stołem  
lasu wracało towarzystwo całe

tadeusz na którego niespodzianie  
weszli w porządku i stanęli kołem  
jak mógł najprędzej niedzielne ubranie

## Dzieła wybrane (z *Gospodarstwa*)

sam na nim padnie dalej w polskiej szacie  
choć sędzia z dokumentów przekonywał  
rejtan żałośny po wolności stracie  
ciszę przerywał



wrogów a praga już się wkoło pali  
pan nigdy przyjąć nie chciałby do ręki  
na szańcach pragi na stosach moskali  
wstąg jasnych pęki

## Dzieła wybrane (z *Gospodarstwa*)

i płynnie mówił i z rozsądkiem wiedział  
chart z kształtu jeśli równie chwytny chwytny  
wieczerzy będzie z mnóstwem gości siedział  
pies faworytny



boku rzuciwszy wzrok na tadeusza  
same portrety na ścianach wisiały  
było z łez które wylotem kontusza  
charty zostały



wszystkim należy lecz każdemu inna  
ma dotąd pierwsze na świecie ogary  
jest bez grzeczności i miłość dziecinna  
pierwsze zamiary

## Dzieła wybrane (z *Gospodarstwa*)

pan świata wie jak długo pracować potrzeba  
moja ty jesteś jak zdrowie ile cię trzeba

...

i z dziecinną radością pociągnął za sznurek  
kolego lecz nim odszedł wyskoczył na wzgórek

...

## Dzieła wybrane (z *Gospodarstwa*)

ruskie przysłowie  
lubił bardzo myślistwo  
narody giną



raz zawitała  
przystojna i młoda jej  
z niezwykłej ich



becząc w ulice  
iżby je w kalendarzu  
zjechać do dworu

## Dzieła wybrane (z dzieł modelu pentagramowego)

często spwojen lubią czytelniki  
przeczalnych elektora waszym dziestu  
bieżności każdy niewnych satelniki  
prawachlarzone koń mi po pójdziestu  
piosek ilości w drobną prawdzięczościć  
jeszczyków ach umienia wątpliwościć



## Dzieła wybrane (z dzieł modelu pentagramowego)

lecz jego tak gdybym posiemdzień jada  
po pisał pogotowanaściej rady  
nośniego dzieckman pozwoju zagady  
je się zorgany wprowa usa nada

no jakby kto piątego w obrzymają  
sobiegl się pewnie wyrażę no także  
rozwinia histo lutego praszakże  
krakować w ogół niego kupowstają

hajdeckich żołądkownictwa przedniony  
obracił się na spokajał na pani  
robotę bo problemiosłe skorony

w swojennych łopa i gnia ruch pod może  
robistopiero wresztą koma kani  
transmisja niezorientacji doktorze

## Dzieła wybrane (z dzieł modelu pentagramowego)

ten fetorycznej sto całe rzemyśle  
krakowym pan swe tym dla uczesnaście  
rzucili co ja wielkie te przemyśle  
można jest właściwie



dla przedzaju do za chciał wreszczenie  
słuszno i niezwyńczą jeszczzenie  
o fidenci przystaje  
uzbrodnicznie świadaje  
jednoczony w grupierwsza i przenie

## Dzieła wybrane (z dzieł programu M. Rudolfa)

to nie muzeum jak płaskie rozdroże  
wróci rozpacz uwierzy że łożę  
życie rzeki pęknięta  
droga ziemi pamięta  
zapomniany nóż mieszka w jeziorze



jutrzenki drzemie przepaść poszukuje  
nie czeka na niepokój wciąż od nowa  
upadły liść jest twardy zginie czuję  
łzy jak zmysłowa



wiotkie serce blasku śpi  
harfa opuściła łzy

## Dzieła wybrane (z dzieł programu M. Rudolfa)

naprawdę słodka lilia błądzi szklana  
wędnąca pleśń umiera gołąb płonie  
nieskończoności znów ciężki grzmot tonie  
pęka gdzieś morze płonie pogrzebana

wieniec zabite oko leczy czy wiesz  
marzy o ludziach niczym mętna czerń las  
o ziemi czarna myśl snu chwali ból czas  
tęcza odchodzi echo leczy ja też

podmuch to nie cierń widziałem niepokój  
o wodzie błyskawica opuściła  
wiecznie sęp to twarz wilk zapomniał spokój

ognisko zadźwięczało wciąż od nowa  
czemu bezradna zapałka rozmyła  
noc szepcze nigdy nie wiesz czy zmysłowa

## Dzieła wybrane (z *Rzeczpospolitej*)

niego co dla borysa jalcyna  
dla nas niezastąpionym krystyna  
święty jerzy ze smokiem  
restaurację z widokiem  
jaką są spadkobiercy lenina



miękka otoczka instytucjonalna  
których ofiary umierały sądy  
zostanie utrzymana liberalna  
że ich zarządy

## Dzieła wybrane (z *Rzeczpospolitej*)

finansowy financial service  
naszą pożał się boże prawicę  
społeczeństwa jest święcie  
nie jest wrogie przejęcie  
wiednia przez pozostałe stolice



do finału dotarły drużyny  
skazanego lub jego rodziny  
jeśli teraz poddamy  
cztery lata poznamy  
sytuuje się w środku drabiny

## Dzieła wybrane (z *Rzeczpospolitej*)

uniwersalny to nie jest krzak róży  
do samochodu wywieźli do lasu  
czyli zmieszczenie się w limicie czasu  
pojazdu przez pracowników w podróży

kilka dni później fundamentaliści  
łukaszewicza jasiek arkadiusza  
zaprezentują pana tadeusza  
wytyczonymi ścieżkami artyści

z pagórków dwóch krzyków tego dnia kiedy  
ciągle mieszkają w ziemi świętej tu spór  
wychodził na boisko tylko wtedy

grabowska prawo co dnia kodeks karny  
na nilu tamy asuańskiej to wzór  
arka przymierza i ołtarz ofiarny

## Alexandr Rosen, *Vínek z veršů TURPISových*

TURPIS je tak trochu plagiátor, ale plagiátor velmi plodný a vlastně tak nějak originální. Kromě toho je umanutým pedantem ve věcech formy a zcela ignoruje vše ostatní, pro což by badatel v říši negativních pravidel mohl mít pochopení.

[...]

Když jsme se seznámili s jeho tvorbou v jazyce Mickiewiczově, nebylo možné odolat. Stačilo upravit fonetickou transkripci a repertoár básnických forem pro jazyk český, předhodit TURPISovi několik MB textů, a od té doby se nedaří mohutný básnivý proud zarazit.

[...]

Múzy působící prostřednictvím generátoru náhodných čísel vedou TURPISovu ruku lovcí v databázi.



# Turpisova tvorba česky

dusě soudruha stalina cítí  
ale všichni si zlomili vaz  
krátce ze světa sovětský svaz  
dusě soudruha stalina cítí  
v severočeském revíru svítí  
skladů mají již za krátký čas  
dusě soudruha stalina cítí  
ale všichni si zlomili vaz

# Turpisova tvorba česky

jedním uchem tam a druhým ven  
k manifestantům promluvil člen  
policejního sboru  
aby po jejich vzoru  
nezávislosti o právech žen



vysoká bílá  
obrovská tvář z plakátů  
jediné slovo



je to její první vrh  
věřím ve svobodný trh

# Turpisova tvorba česky

pouze mechanický akt  
zapomenout každý fakt



chaosu života víra  
jednoho portýra  
do něho nalili rumu  
přichází k rozumu

## Turpisova tvorba česky

na mazlení myšičko myš  
a já brouček a já brouček  
mír nehne se myška ni již  
a já brouček sekal souček  
to je vnouček malý klouček  
se v hrsti spočítám je hned  
to je bába to je vnouček  
hned jedna dva tři čtyři pět

pan doktor mě povídal cos  
vsěchno snědla udělala  
nos větší nežli nosí kos  
padala tralala mala  
běhala myška běhala  
bramboru položil ji před  
na něj padala tralala  
hned jedna dva třičtyři pět

# Co dalej?

## 1. Zastosowania TURPISa:

- Pomoc naukowa do ćwiczeń z poetyki.
- Wygaszacz ekranu (typu `starwars`).

## 2. Wyzwania:

- Jak ulepszyć transkrypcję fonetyczną?
- Jak reprezentować rymy niedokładne?
- Jak losować wersy poprawne syntaktycznie?

# Jak losować drzewa rozbioru o danym ry(t)mie?

- Program POETA M. Rudolfa:

gramatyka losująca drzewa poczynając od korzenia

- Aby uzyskiwać produkcje o danym ry(t)mie, wszystkie jednostki frazowe należy wzbogacić o parametry ry(t)mu

subst : sg : nom : f : [ ! . ] : [ aba ] → laba | Raba | żaba

- Wzbogacenie — dość złożona transformacja gramatyki frazowej, przy wykorzystaniu słownika wymagań czasownikowych i słownika form fleksyjnych.

- Rozwiązanie zastępcze:

Losowanie drzew aż do uzyskania produkcji o pożądanym własnościach. (*Czekaj tatka latka...*)

## Co z POETY można odsiać w rozsądnym czasie?

już wiem  
odchodzi  
może nie  
to ogród

odchodzi  
wiecznie  
za późno  
odeszłaś

wiecznie  
odchodzi  
odeszłaś  
może nie

łóże śpi  
poza mną  
poza mną  
poza mną

## Co z POETY można odsiać w rozsądnym czasie?

jaszczurka zwiędła  
lekki rytuał płonie  
tak to się skończy



cienkie ognisko dnia tonie w otchłani  
stłuczone serce poszukuje źródła  
łóże księżycy pije nieskończoność  
nic nie zostało



taniec to nie bolesny grobowiec  
srebro blasku nie pragnie miłości  
płaskie oko przebacza  
jak pragnienie księżycy  
tylko dzban jest naprawdę zwycięski



## Co z POETY można odsiać w rozsądnym czasie?

szary głód dnia zapomni o nicości  
kot świata śpiewa niczym jabłko nocy  
mściwa melodia rzeki chce młodości  
kryje się w nocy



latarnia to nie fioletowe morze  
srebrzysta chmura pogrąża się w srebrze  
spragniony kot świtu tonie w jeziorze  
oddech śpi w srebrze