

XXIII. A Very Early Slavic Letter Statistic: Jan Svatopluk Presl (1791-1849)

K R O K. *Weřejný spis všenaučný pro vzdělance národu Česko-Slowanského. Wydáván přispjwánjm mnoha učených vlastenců od Jana Swatopluka Presla, Doktora w lékařstw, c. k. weřejného professora přírodopisu a ředitele c. k. přírodowny na vysokých školách pražských.*

K R O K.

Weřejný spis všenaučný

pro

Wzdělance Národu Česko-Slowanskeho.

Wydáván

přispjwánjm mnoha učených vlastenců

od

Jana Swatopluka Presla,

Doktora w lékařstw, cjs. král. weřejného Professora
přírodopisu a ředitele c. k. přírodowny na vysokých školách
pražských a auda několika společuostj učených.

Spartam, quam nađus es, orna.

Díl druhý.

(Se 7 obrazci.)

W PRAZE

k dostánj u Josefy Fetterlowé, arcibiskupské knihtiskařky, w
Seminarium Nře. 190.

1831.

Letter and sound frequencies obtained an increasing interest in the 19th century, relevant research arising in the first half of the century and being extended later. To a certain extent, the overall interest of this increase was motivated linguistically; in this case, it has to be seen as a specific continuation of comparative approaches in the field of linguistics, ultimately referring back to the Neogrammarians' interest in regularities, or even "laws", underlying language and language development. Later, this predominantly linguistic interest was increasingly accompanied by technical needs, starting from shorthand and stenography, including cyphering and decyphering techniques, different type-writer systems, extending to information transfer, etc.

It goes without saying that in the earliest studies, sound and letter statistics were not always strictly distinguished, and if so, results were highly dependent on the sound definitions chosen. As far as the earliest Slavic statistics are concerned, August Schleicher's data on Old Church Slavonic sound frequencies tend to be regarded as one of the earliest letter or sound statistics (cf. Grzybek/Kelih 2003, 2004 2005): Schleicher had published these data in his *Formenlehre der kirchenslawischen Sprache* (1852: 20f.), understanding them as a complement to the earlier data on German, Greek, Latin and Gothic, which Ernst Förstemann (1846, 1852) had provided in the seminal studies.

In this context, an almost unknown sound statistic must be presented here which was published almost two decades before the studies mentioned above. This statistic seems to be one of the very first letter and sound statistics in general, not only in the Slavic area. It was published in 1831, in the first Czech scientific journal *Krok*, edited by Jan Svatopluk Presl (1791-1849).



Jan Svatopluk Presl was born in Prague. His early interest in medicine culminated in his study at Karl-Ferdinand University in Prague from which he graduated as a doctor of medical sciences, in 1816. After being assistant for zoology and mineralogy, from 1816 to 1818, he became professor for these subjects at Olomouc university, and was then appointed professor for zoology and mineralogy at Prague University, in 1820. He soon became one of the most important scientists of his country, who made important contributions to the (Czech) scientific terminology in botanics, chemics, zoology, mineralogy, geology, paleontology. In fact, with these works Presl inspired the development of Czech scientific terminology, thus making it possible for original Czech research to develop. To be sure, at a time when not Czech, but rather German was the standard

language for official communication (e.g., in Czech elementary schools and the administrative bureaucracy), Presl became very active in promoting Czech language and culture, hereby being in closest contact with the leading thinkers of his time, who were promoters and representatives of the Czech National Revival (*České národní obrození*). This cultural movement – the main purpose of which was to revive Czech [language](#), culture and national identity – was headed by [Josef Dobrovský](#) (1753-1829). and [Josef Jungmann](#) (1743-1847). Whereas Dobrovský has often been called one of the founders of slavistics as an academic discipline, Jungmann is predominantly famous for his *Historie literatury i jazyka českého* (1825) and his five-volume *Slovník jazyka českého* (1835-39), a major lexicographical work, which had a great formative influence on the Czech language.

Together with Dobrovský and Jungmann, but also with Václav Hanka, Jan E. Purkyně, and others, Presl founded a society aiming at the promotion of Czech in academic life, which

initially pursued the idea of founding a Slavic Academy in Prague. Presl became actively engaged in founding the Czech National Museum; together with Jungmann, Hanka, Pavel J. Šafárik (1795-1876) and František Palacký (1798-1876) Presl also founded the Committee for the Scientific Development of the Czech Language and Literature at the museum, from which the literary foundation *Matica česká* emerged in 1831.

In extensive discussions with the group of scholars mentioned above, Presl attempted to prove Czech being an adequate language for scientific discourse. It was on this general background that, in 1821, first Czech scientific journal *Krok* was founded. Most of the above-mentioned scholars contributed to this journal, which appeared with irregular intervals until 1840. Despite his pronounced engagement in Czech cultural affairs, it was only towards the end of his life that Presl became increasingly politically involved: In the revolutionary year of 1848, he was elected into the Czech National Committee; he also became one of the 341 delegates to the All-Slav Congress in Prague, where 42% of the participants were of Czech or Slovak origin. The Congress was headed by Palacký – one of the regular contributors to *Krok*.

Srovnání spoluhlásek a samohlásek v češtině, vlastině a němčině.			
<p>Wyčítá se vůbec češtině, že pro náramné množství spoluhlásek jest tvrdá a těžká k vyslovení. K vyvrácení této výčitky posloužij následující porovnání.</p>			
<p>W 1000 písmenách</p>			
českých	wlaských	německých	
samohlásek 480,0	512,0	326,0	
a . . . 120,0	130,0	80,0 (aa, ah, au, ei aci, ei, eu, ey.)	
e . . . 110,0	135,0	77,5 (ä, ee, oe, eh,*)	
i . . . 90,0	110,0	127,5 (ie, ii, y, i w ai a d.)	
o . . . 100,0	92,0	20,0 (oo, oh)	
u . . . 60,0	45,0	22,0	
spoluhlásek . . .	520,0	488,0	674,0
b . . . 10,0	17,5	38,0	
c . . . 10,0	2,5 (z)	22,5 (z)	
č . . . 4,0	5,0	00,0	
d . . . 50,0	25,0	66,0	
f . . . 00,0	5,0	18,0 (v)	
g . . . 00,0	7,0 (gh)	25,5	
h . . . 2,5	2,5	50,0	
ch . . . 17,5	0,0	12,0	
j (g) . . . 10,0	2,5	2,5	
k . . . 42,5	45,0 (e před a, o, u, kw qu)	22,5	
l . . . 40,0	70,0 (ll)	20,0 (ll)	

	českých	wlaských	německých
m . . .	37,5	22,0 (mm)	29,0 (mm)
n . . .	64,0	90,0 (nn)	127,5 (nn)
p . . .	22,5	27,0	5,0
q . . .	wiz k a w		
r . . .	25,0	43,0	70,0
ř . . .	2,5	00,0	00,0
s . . .	27,5	37,0	55,0
š . . .	30,0	7,5 (sci)	36,5 (sch)
ž . . .	57,0	55,0	52,0
w . . .	27,5	15,0 (v, w w qu)	22,0
x . . .			
z . . .	30,0	wiz s	wiz s
ž . . .	10,0	2,5 (gi)	00,0

Na jednu samohlásku přigde w češtině spoluhlásek 1,083, we vlastině 0,952, w němčině 2,067. Na jednu spoluhlásku přigde w češtině samohlásek 0,923, we vlastině 1,049, w němčině 0,483.

*) e nemá gsau právem wynechána, poněwadž se newyslovugj.

The name of *Krok* is a homonym¹ in Czech, and the name seems to be carefully chosen: On the one hand, it means “step”; and in fact, the foundation of this journal was an important step in Czech cultural development. On the other hand, *Krok* refers far back to Czech (Bohemian) history, being the name of an Ancient Czech (Bohemian) legendary ruler who reigned in the 7th century. Accordings to the legends – cf. the *Chronica Boemorum* by Bohemian writer and

¹ My thanks to Ludmila Uhlířová who directed my attention to this ambivalence.

historian Cosmas of Prague (ca. 1045-1125) – , one of his three daughters, Libuše is the foundress not only of Prague, but also, by her marriage with Přemysl, the foundress of the Přemyslid dynasty who ruled in the Czech lands from 873 or earlier until the 14th century.

The sound statistics published in volume 3 of *Krok* includes data on Czech, Italian, and German. The author's introductory words very clearly put the whole publication into the cultural background described above: after all, the statistics are meant to provide counter-evidence as to the alleged difficulty of pronouncing Czech. Therefore, consonant-vowel combinations are computed, resulting in similar relations for Czech and Italian, as opposed to German.

It goes without saying, that these data need more thorough re-analysis than this can be done in this context, the more since the data are accompanied by some methodological flaws. For example, no information is given as to the text material analysed, or as to the sample size. Furthermore, no raw data are given, but only percentages (or rather promille, ‰), summing up to 1 only in case of Czech and German, thus obviously containing some errors as to Italian. Anyway, the data are supposed to represent sound, not letter frequencies. This can be seen, for example, from the treatment of German where not letter combinations such as <aa>, <ah>, <au> are treated as [a], but also <ei> or <eu>; similarly, not only <e>, but also the <ee>, <ä> or <oe> are counted as [e]. Also, double consonants like <ll>, <mm> or <nn> are counted as [l], [m], or [n], respectively. The letter <z> is counted as [c] if realized as affricate (usually being denoted as [ts], today), voiceless <v> as [f], etc. etc. – for further details see the reproduction above.

As far as we know, the publication in *Krok* is one of the earliest sound statistics we know of, at least with regard to Slavic languages. For the history of quantitative linguistics, it would be interesting and valuable to find earlier references.

References

- Hoffmannová Eva** (1973). *Jan Svatopluk Presl, Karel Borivoj Presl*. Praha.
- Kimball, Stanley B.** (1973). The Austro-Slav Revival: A Study of Nineteenth-Century Literary Foundations. *Transactions of the American Philosophical Society, NS, 63(4)*, 1-83.
- Kolari, Veli** (1973). Notes on Jan Svatopluk Presl as Terminologist. *Scando-Slavica 19*, 187-196.
- Kolari, Veli** (1981). *Jan Svatopluk Presl und die tschechische botanische Nomenklatur: Eine lexikalisch-nomenklatorische Studie*. Helsinki.
- Moritsch, Andreas** (Hrsg.) (2000). *Der Prager Slavenkongress 1848*. Wien etc.
- Moritsch, Andreas** (2000). Revolution 1848 – Österreichs Slaven wohin? In: Moritsch (Hg.) (2000): 5-18.
- Pokorný, Jiří** (2000). Der Prager Slavenkongress und die Tschechen. In: Moritsch (Hg.) (2000): 63-70.
- Weitenweber, Wilhelm Rudolf** (1854). Denkschrift über die Gebrüder Joh. Swat. Und Carl Bor Presl. *Abhandlungen der Königlich-Böhmischen Gesellschaft der Wissenschaften, Folge 5, Bd. 8*, 4-27.

Peter Grzybek

Glottometrics 13

2006

RAM-Verlag

Glottometrics

Glottometrics ist eine unregelmäßig erscheinende Zeitschrift für die quantitative Erforschung von Sprache und Text

Beiträge in Deutsch oder Englisch sollten an einen der Herausgeber in einem gängigen Textverarbeitungssystem (vorrangig WORD) geschickt werden

Glottometrics kann aus dem **Internet** heruntergeladen, auf **CD-ROM** (in PDF Format) oder in **Buchform** bestellt werden

Glottometrics is a scientific journal for the quantitative research on language and text published at irregular intervals

Contributions in English or German written with a common text processing system (preferably WORD) should be sent to one of the editors

Glottometrics can be downloaded from the **Internet**, obtained on **CD-ROM** (in PDF) or in form of **printed copies**

Herausgeber – Editors

G. Altmann	ram-verlag@t-online.de
K.-H. Best	kbest@gwdg.de
P. Grzybek	peter.grzybek@uni-graz.at
A. Hardie	a.hardie@lancaster.ac.uk
L. Hřebíček	hrebicek@orient.cas.cz
R. Köhler	koehler@uni-trier.de
J. Mačutek	jmacutek@yahoo.com
G. Wimmer	wimmer@mat.savba.sk
A. Ziegler	arne.ziegler@uni-graz.at

Bestellungen der CD-ROM oder der gedruckten Form sind zu richten an
Orders for CD-ROM's or printed copies to

RAM-Verlag RAM-Verlag@t-online.de

Herunterladen / Downloading: <http://www.ram-verlag.de>

Die Deutsche Bibliothek – CIP-Einheitsaufnahme

Glottometrics. –13 (2006) –. – Lüdenscheid: RAM-Verl., 2006

Erscheint unregelmäßig. – Auch im Internet als elektronische Ressource unter der Adresse <http://www.ram-verlag.de> verfügbar. -

Bibliographische Deskription nach 13 (2006)

ISSN 1617-8351