

Empirical Approaches to Text and Language Analysis

*dedicated to Luděk Hřebíček
on the occasion of his 80th birthday*

edited by

Gabriel Altmann, Radek Čech,
Ján Mačutek, Ludmila Uhlířová

**2014
RAM-Verlag**

Studies in quantitative linguistics

Editors

Fengxiang Fan (fanfengxiang@yahoo.com)
Emmerich Kelih (emmerich.kelih@uni-graz.at)
Reinhard Köhler (koehler@uni-trier.de)
Ján Mačutek (jmacutek@yahoo.com)
Eric S. Wheeler (wheeler@ericwheeler.ca)

1. U. Strauss, F. Fan, G. Altmann, *Problems in quantitative linguistics 1*. 2008, VIII + 134 pp.
2. V. Altmann, G. Altmann, *Anleitung zu quantitativen Textanalysen. Methoden und Anwendungen*. 2008, IV+193 pp.
3. I.-I. Popescu, J. Mačutek, G. Altmann, *Aspects of word frequencies*. 2009, IV +198 pp.
4. R. Köhler, G. Altmann, *Problems in quantitative linguistics 2*. 2009, VII + 142 pp.
5. R. Köhler (ed.), *Issues in Quantitative Linguistics*. 2009, VI + 205 pp.
6. A. Tuzzi, I.-I. Popescu, G. Altmann, *Quantitative aspects of Italian texts*. 2010, IV+161 pp.
7. F. Fan, Y. Deng, *Quantitative linguistic computing with Perl*. 2010, VIII + 205 pp.
8. I.-I. Popescu et al., *Vectors and codes of text*. 2010, III + 162 pp.
9. F. Fan, *Data processing and management for quantitative linguistics with Foxpro*. 2010, V + 233 pp.
10. I.-I. Popescu, R. Čech, G. Altmann, *The lambda-structure of texts*. 2011, II + 181 pp.
11. E. Kelih et al. (eds.), *Issues in Quantitative Linguistics Vol. 2*. 2011, IV + 188 pp.
12. R. Čech, G. Altmann, *Problems in quantitative linguistics 3*. 2011, VI + 168 pp.
13. R. Köhler, G. Altmann (eds.), *Issues in Quantitative Linguistics Vol 3*. 2013, IV + 403 pp.
14. R. Köhler, G. Altmann, *Problems in Quantitative Linguistics 4*. 2014. V + 148 pp.
15. K.-H. Best, E. Kelih (eds.), *Entlehnungen und Fremdwörter: Quantitative Aspekte*. 2014. VI + 163 pp.
16. I.-I. Popescu, K.-H. Best, G. Altmann, *Unified Modeling of Length in Language*. 2014, IV+123 pp.
17. G. Altmann, R. Čech, J. Mačutek, L. Uhlířová (eds.). *Empirical Approaches to Language and Text Analysis*. 2014. VI + 217 pp.

ISBN: 978-3-942303-24-8

© Copyright 2014 by RAM-Verlag, D-58515 Lüdenscheid

RAM-Verlag

Stüttinghauser Ringstr. 44, D-58515 Lüdenscheid

RAM-Verlag@t-online.de

<http://ram-verlag.de>

A letter to Luděk Hřebíček from the editors

Dear Luděk Hřebíček,

The years have quickly passed since the time when we were happy to dedicate to you a volume „*Text as a Linguistic Paradigm: Levels, Constituents, Constructs*“ (Quantitative Linguistics, Vol. 60, Wissenschaftlicher Verlag Trier)“ to celebrate your 65th birthday. Now, fifteen years later, we have another great opportunity to congratulate you most sincerely: Please accept the content of the present book as a kind of scientific metaphor and paraphrase of “*Happy birthday to you*“... composed by 25 scientists from 10 countries on the occasion of your unbelievable 80th birthday. This festschrift is meant to be an expression of a deep respect for you as a prominent Czech orientalist, a theoretician of language and text structure, an author of numerous books, studies, and articles published in English, German and Czech in publishing houses across Europe, and last but not least, for you as an excellent essayist who knows perfectly how to persuade your readers that to build a theory of text is an everlasting and fascinating adventure.

The editors

A Postscript

All of us, who are Hřebíček’s readers, colleagues, students and successors, know well that perhaps the most characteristic and, probably, the most impressive feature of Hřebíček’s scientific work – from the very beginning of his career – is a unique and original methodological line of thinking and reasoning. For several decades, he has been construing his quantitative text theory as an interdisciplinary field, and as a testable theory, very near to those done/performed/ implemented in mathematics and the natural sciences. He formulates models, hypotheses, and conjectures, using various mathematical procedures to test them. He hates chaos (with the only exception of *deterministic* chaos). He loves order, rules, laws, and synergy. He views and interprets linguistics as a *science of language*. He is one of the well-known representatives of such exact thinking in the European scientific context, as well as a singular voice in the context of present-day Czech linguistics. (Of course, not a singular voice in the sense of an isolated promoter of certain linguistic ideas, but in the sense of a perfectly “sharpened” thinker – a linguist and philosopher of language.)

The festschrift mentioned above contains Hřebíček’s selected bibliography of monographs, textbooks, articles, book reviews and popular articles (not to speak about his translated novels and works for the stage, mainly from Turkish). Still, comparing the rich bibliography published there with what he has written since 2000, it is clear that the period of the last fifteen years has been most productive and most fruitful. In addition to articles and papers published both in the Czech Republic (e. g., in the journals *Slovo a slovesnost* or in *Vesmír*), and

abroad (in various volumes, mostly in English), three books have appeared: *Variation in Sequences* (2000, Prague: Oriental Institute); *Vyprávění o lingvistických experimentech s textem* (2001, Prague: Academia); *Text in Semantics* (2007, Prague: Oriental Institute). In addition, Hřebíček is a co-author of a voluminous collective book written together with G. Wimmer, G. Altmann, S. Ondrejovič, and S. Wimmerová *Úvod do analýzy textov* (2003, Bratislava: Veda). The books represent a further decisive step in the development of modern text linguistics. Hřebíček's ideas are acknowledged, accepted, approved and further developed, especially by the younger generation. The present festschrift is good evidence of it. It is not "only" a birthday present, but a manifestation of the undeniable fact that Hřebíček's ideas represent a big challenge to follow.

Contents

Gabriel Altmann The study of hrebs	1
Sergey Andreev Representations of Tutchchev's style: one poet or two?	14
Jan Andres The Moran-Hutchinson formula in terms of Menzerath-Altmann's law and Zipf-Mandelbrot's law	29
Radek Čech, Ludmila Uhlířová Adverbials in Czech: Models for their frequency distribution	45
Fan Fengxiang, Zhou Pianpian, Su Hong The use of the <i>POR</i> in macro-lexical analyses	60
Wei Huang, Haitao Liu The phoneme-grapheme relation in the scheme of the Chinese Phonetic Alphabet	69
Emmerich Kelih, Andrij Rovenchak, Solomija Buk Analysing h-point in lemmatised and non-lemmatised texts	81
Reinhard Köhler The fractal structure of linguistic motifs	94
Miroslav Kubát Moving window type-token ratio and text length	105
Lu Wang Part-of-speech Concentration in Chinese	114
Ján Mačutek, Gejza Wimmer A measure of lexical text compactness	132
Jiří Mácha, Olga Richterová The Quantum of Plurality. The relationship of singular and plural (and singularia and pluralia tantum) in Czech nouns	140

Tomi S. Melka Palindrome-like structures in the <i>rongorongo</i> script	153
Georgios Mikros, Jiří Milička Distribution of the Menzerath's law on the syllable level in Greek texts	181
Haruko Sanada The choice of postpositions of the subject and the ellipsis of the subject in Japanese	190
Kamil Stachowski The volume of Ottoman lexical influence on Romanian	207
Addresses of authors	229

The volume of Ottoman lexical influence on Romanian

Kamil Stachowski

0. Rationale

This paper continues the pursuit of two of the goals set by quantitative linguists. On the one hand, it wants to help implement Luděk Hřebíček's wish (1990: 371), and demonstrate one way of how quantitative and qualitative methods and results can be used jointly to gain a kind of insights that either of them would find difficult to obtain on its own. On the other hand, it will advance the work of K.-H. Best (2005) by further exploring Turkic influence on the languages of Europe, and the possibilities and limitations of the so-called Piotrovskij-Altman law. In this sense, this paper is a continuation of Stachowski (2013), where a similar approach was applied to Turkic glosses in Hungarian and Polish.

In particular, we investigate the vocabulary of Oriental origin that Romanian acquired as a result of the Ottoman presence in the Balkans. The problem is an extensive and complex one, and the current paper does not intend to offer a full analysis. It will be limited to establishing just the overall, general quantitative characteristic of the said vocabulary.

I will: 1. introduce the historical background, its linguistic results, and the assumptions made in this paper, 2. explore selected quantitative aspects of Ottoman influence on Romanian as a whole (2.1), and on the individual regions where Romanian is spoken (2.2), and lastly 3. summarize the conclusions.

1. Introduction

This section provides background information without which the analytic part of the paper may be unclear (1.1–1.2), and introduces Suciú (2009, 2010) as the source of the material discussed in it, together with the adopted assumptions and abbreviations (1.2–1.4). In particular, the note on the terms “Ottoman” and “Romanian” in 1.3 should not be skipped.

1.1 Political history

Romanian is spoken in a territory that encompasses several regions whose histories are often separate, complicated, and marked by long periods of conflicting foreign interferences which had resulted in an unclear political situation

(see e.g. Ágoston, Masters 2009; Pop, Bolovan 2006; Sallanz 2005; Sugar 1977; see also Fig. 1 below).

Wallachia, i.e. Muntenia and Oltenia taken together, is home to one of the first Romanian principalities, founded in 1330 by Basarab I, probably of Turkic (Cuman) origin himself. The country fell again under Hungarian suzerainty soon afterwards, and in 1394, it sent tribute to the Ottomans, an act that the Porte considered an acknowledgment of vassalage. The Ottoman grip on Wallachia tightened further in 1476, after the death of the infamous Vlad the Impaler, but the region was never incorporated into the Empire as a province. The next two hundred years were not free of turmoil; finally, in 1715, the so-called Phanariot epoch began, when voivodes were appointed by the Ottomans from among the Greeks of Istanbul. In the years 1716–39, Oltenia, the western part of the region, was under Austrian rule, and in the second half of the 18th century, the entire country was occupied more than once by Russians and by the Habsburgs. In 1829, Wallachia was together with Moldavia under Russian military rule, in 1859 both elected the same ruling prince, thus forming a personal union, and in 1878 they were united with Dobruja giving rise to the sovereign Kingdom of Romania.

Moldavia, before becoming an independent state in 1359, remained mostly under Hungarian control and was repeatedly invaded, among others, by Turkic-speaking Cumans and Tatars. In 1377 it sent taxes to the Ottomans for the first time. Nevertheless, the country had managed to retain a degree of independence until the late 15th and early 16th century, and never became a province of the Ottoman Empire. The two hundred years before that time, and the following two hundred, were a period of almost constant attempts to regain full independence, with Moldavia's own forces, or by various alliances, and of the consequent Ottoman retaliations. In 1601–18, the country recognized Polish suzerainty. Beginning from 1711, its voivodes were chosen from among the Phanariots. In 1775, the Habsburgs gained control over Bukovina, i.e. the northwestern part of the region, and in 1812, its eastern part, Bessarabia, was seized by the Russian Empire. The remainder came under Russian control after 1829, and became part of independent Romania in 1878.

Dobruja came under Bulgarian control in late 12th century, and was repeatedly invaded, by, among others the Romanian Principality of Wallachia, and by the Turkic Cumans and Pechengs, until it was incorporated into the Ottoman Empire in 1419 to remain one of its provinces until 1878 when it was united with Moldavia and Wallachia as a part of the newly created sovereign Kingdom of Romania. For almost half a millennium, the region was inhabited primarily by Turks and Tatars while Bulgarians and Romanians were in a minority. In 1878, the population of the Constanța County was composed of Tatars (38%), Romanians (23%), Turks (18%), Bulgars (13%), and other nations (8%). These proportions changed dramatically and rapidly, and by 1930, the same county was mainly inhabited by Romanians (66%) while the Turkic element constituted less than 10% of the total population, still about twice as many as it does presently.

Conquered in late 9th and early 10th century, Transylvania remained under Hungarian control, albeit with an increasingly high degree of autonomy, until the defeat at Mohács in 1526 when it effectively became independent, and an object of struggle between Austrians, Hungarians, and Ottomans. The country began to pay tribute to the Porte in 1543, and remained her vassal until the late 17th century. It was occupied by the Habsburgs since 1687, but the Ottoman Empire did not concede the loss until 1699. After World War I, Transylvania proclaimed unification with Romania; the act was internationally recognized in 1920.

Three regions have for some time been considered parts of Transylvania. In the southwest, Banat passed from Hungary to the Ottoman Empire in 1552, and from the Empire to the Habsburgs in 1718. After World War I, it was divided between Hungary, Yugoslavia, and Romania which obtained the largest, eastern part of the land. In the west, Crișana was divided between Transylvania and Austria until 1660 when it was seized by the Ottomans for twenty-six years before being incorporated into the Habsburg Empire until 1918. In the northwest, Maramureș shared the fate of the core part of Transylvania in the 16th and 17th century. It fell gradually under Habsburg control, was fully annexed in 1732, and remained Austrian till the end of World War I.

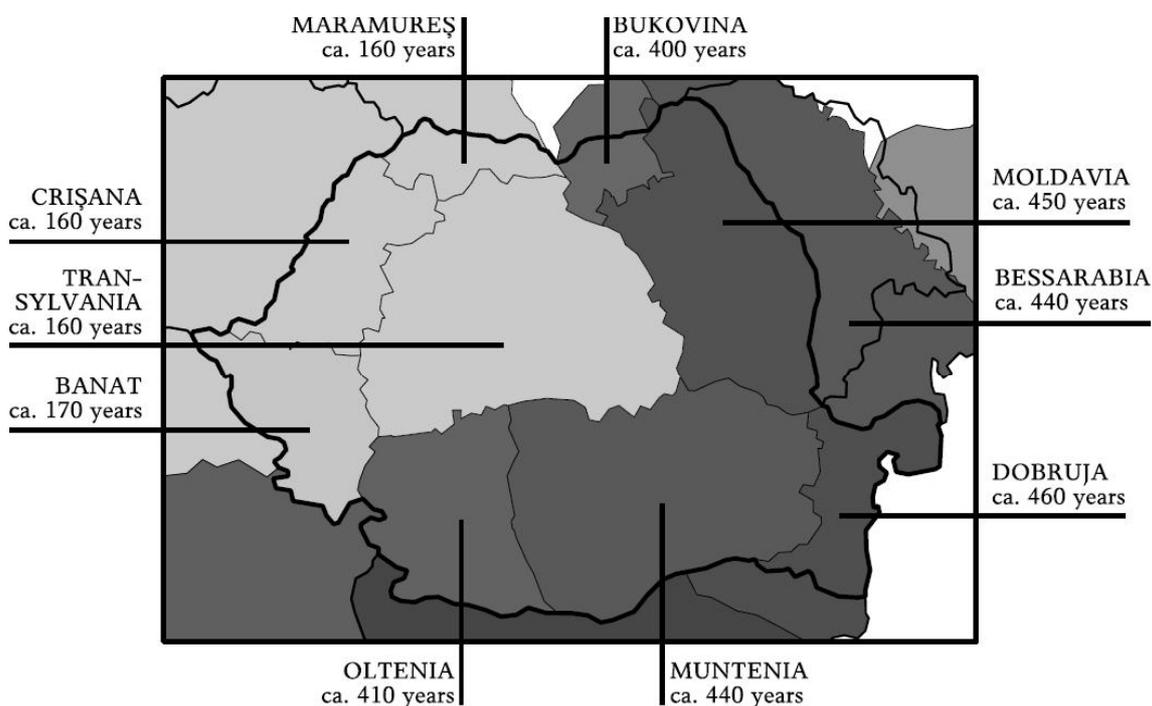


Figure 1. The approximate duration of Ottoman authority over historical regions where Romanian has been spoken, and over the neighbouring lands. Modern borders are added for reference.

1.2 Linguistic results

Remarkably, Romanians have carried their national and linguistic identity through the turbulent times, but traces of old partitions can still be seen in the language today. Dialectal division of Romanian has been the object of a long debate, see e.g. Ursan (2008: 77f) which primarily focused on the status of the Transylvanian variety, while the dialects of Banat, Moldavia, and Muntenia appear to be relatively clearly marked.

There is also a subtler result of this situation, which will be perhaps best illustrated by the word *taban*, which Suciu (2010: 709) defines so:¹

TABÁN² s[ubstantiv] n[eutru] (reg[ional]; 1879–); pl[ural] *-uri* □ *-e*.

1. (Tehn[icǎ], Agric[ulturǎ]; Dobr[ogea]) ‘Patin de la charrue’: **tabán** 1884–1885. **2.** (Constr[uc]□*ii*; Ban[at]) ‘Socle d’un four’: **tǎván** sec[olul] XX/1. **3.** (Constr.; Mold[ova], S[ud] Transilv[ania]) ‘Poutre qui forme la base d’une construction’: **tabán**, **taván** sec. XX/1. **4.** (Constr.; Mold., Munt[enia], Olt[enia]; □*i* la pl., cu sens de s[in]g[ular]; în Mold. determinat prin *de jos*) ‘Plancher’: **taván** sec. XX/1. **5.** (*P[rin] ext[ensie]*; Constr.; rar) ‘Terrasse (d’une maison paysanne)’: **tabán** 1939. **6.** (Constr.; Mold.) ‘Planche longue et mince; latte’: **tabán** 1879, **taván** 1895, **tabá** sec. XX/1, **tabón** mijl[ocul] sec[olului] XX. **7.** (Vestim[enta]□*ie*; Mold.) ‘Semelle intérieure, semelle première’: **tabán** 1900. **8.** (*P. anal[ogie]*; Vestim.; Mold., rar) ‘Bordure de pelisse d’un vêtement’: **tabán** 1924. **9.** (Med[icină]; V[est] Munt.) ‘Enflure du sabot des bêtes’: **taván** 1967. **10.** (Agric.; Dobr.) ‘Cadre de bois auquel on attelle les chevaux dans l’aire de battage’: **tabán** sec. XX/2.

– Din t[ur]c[ă](-osmanlie) **taban** „idem (1, 2, 3, 4, 6)” □*i*, în general, ‘semelle; plante du pied; socle, support, base, partie inférieure’ [...] probabil □*i* *,idem (10)”, **taban[astarı]** „idem (7)” [...], **[iç] taban** „idem (7)” [...], **taban [kabartısı sau şışı]** „idem (9)” [...] – Cf. s[ârbo]cr[oa]tă *taban* „talpa piciorului; talpa plugului; fund de brazdă” [‘foot, ploughshare, bottom of the furrow’], n[eo]gr[eacă] *ταμπάνι* „grindă, stâlp” [‘beam, pole, pillar’], arom[ână] *tăbáne* „grindă” [‘beam’], alb[aneză] *tabán* „talpa piciorului; talpa încăl[ă]minteii, pingea; cre[ș]tetul capului” [‘foot, sole, crown’] [...]

It seems easy to count loanwords based on just the head of the entry. On closer inspection, however, it quickly becomes apparent that the results so obtained may be debatable and, in fact, severely underestimated. Of course, Suciu (2009, 2010) is also aware of these difficulties, and discusses them in more detail in the first volume of his work (2009: 58f).

¹ For readability, certain parts of the originally rather long entry have been omitted here without notice. In particular, sources have not been copied past the date of the first attestation, and Romanian definitions of meanings have been left out and only the French translations preserved. Where these were missing, English ones have been added in square brackets. Abbreviations are only explained where they appear for the first time, and their meaning is not always given in full (e.g. “Tehn.” stands for ‘tehnică, termen tehnic (*technique*)’, but here it is only expanded to “Tehn[icǎ]”).

On the Romanian side, Dobr. *taban* ‘ploughshare’ (meaning 1) is, for example, quite unlikely to have resulted from the same act of borrowing as e.g. Ban. *tăvan* ‘oven base’ (meaning 2), or even Dobr. *taban* ‘wooden block to which horses are attached during threshing’ (meaning 10), i.e. the same phonetic shape used in the same region but with a rather different semantics. On the Turkic side, *taban* ‘sole, heel, base, floor, &c.’ does not necessarily have to be considered the same etymon as the phrases *taban kabartısı* or *taban şişi* ‘swelling of *taban*’ which yielded western Munt. *tavan* ‘swelling of the inner hoof in cattle’ (meaning 9), and where *taban* is merely an attribute.

Meanings 5 and 8 are marked by Suciú as Romanian innovations (indicated by the qualifiers “P[rin] anal[ogie]” ‘by analogy’ and “P[rin] ext[ensie]” ‘by extension’, see Suciú (2010: 8)), highlighting the possibility of internal borrowing between Romanian dialects. Cases where semantic or phonetic change did not occur in the process may be virtually indistinguishable from words borrowed directly from Turkish. Also, note the presence of Alb. *taban* ‘foot, sole, crown’ and SCr. *taban* ‘foot, ploughshare, bottom of the furrow’, both of which might well have mediated the word from Turkish to Romanian. The eventual Turkic source may also have been Gagauz or Tatar in some cases, especially in 20th century borrowings in Dobruja and Moldavia (e.g. *amlă* ‘sitting board at the end of a fishing boat’, attested in Dobruja since 1971, or *saman* ‘straw’, attested in eastern Moldavia since 1978).

Thus, the actual number of Romanian words is open to discussion, and so is the number of etymons that they stem from, and the routes that they have taken. Technically, words borrowed via Serbian or any other language – including Romanian dialects – are not *Turkish* loanwords. Leschber (2011) discusses in more detail almost seventy Romanian words with such complicated etymologies, and several interesting cases can also be found investigated in Mitu (2002).

1.3 Treatment of the material

Since etymology is often unable to resolve the three issues listed above, and the answers are necessary before a quantitative analysis can begin, simplifications had to be made.

The material will be considered first as a whole, and then as a collection of independent, regional wholes. In the first case, Suciú’s judgement will be followed and every entry counted as one loanword. In the second, entries will be split, and each individual meaning with a regional qualifier treated as a separate borrowing – unless it is believed by Suciú to have evolved on Romanian ground. For example, *taban* above will count as one loanword in Romanian in the first part of the analysis, while in the second, as one borrowing in Banat (meaning 2), one in Oltenia (4), one in southern Transylvania (3), two in Dobruja (1, 10), two in Muntenia (4, 9), and four in Moldavia (3, 4, 6, and 7), and meanings 5 and 8

will be ignored as Romanian innovations. In the second part, also entries (meanings) without regional qualifiers will be omitted even when they are marked as literary. The beginnings of the Romanian literary language reach back to the 16th century (Ghe \square ie, Mare \square 1985: 451) but it cannot be expected to have already been universally and fully adopted in the remote past.

This unequal treatment has been dictated by that the first part will be considering the heterogeneous sum of regional varieties and the literary language, while the second part will be dealing with each in isolation as if with a separate language.

The limitations mentioned in 1.2 and the methodological choice made above will also affect use of terms:

- “Romanian” will denote ‘Daco-Romanian, as a whole or any of its (sub)dialects’,
- “Ottoman” will be used to refer to ‘any Turkic language from which word(s) have penetrated into the Balkans as a result of the presence of the Ottoman Empire in the region’, i.e. not so much to the nation and the high language of its elites as to the political entity it ran. (See e.g. the ethnic composition of Dobruja mentioned in 1.1.)

The material analyzed in this paper has been primarily extracted from Suciu (2010). The datation of loanwords is also given in a table in Suciu (2009: 619f), and it was consulted where the notation was unclear (e.g. s.v. *bairac-agă*). More importantly, however, these series of dates are not always identical. The differences are discussed in 2.1.1, but essentially of no consequence for the main points of the paper.

1.4 Piotrovskij-Altman law

First suggested by R.G. Piotrovskij in 1960, the so-called Piotrovskij-Altman law has since evolved in Altmann (1983) and other works into a convenient formula for quantitative modelling of change in language (eq. 1).

$$(1) \quad p(t) = \frac{1}{1 + ae^{-bt+ct^2}}$$

It has three variants, that describe: (1.) a reversible (nonmonotonic) change ($C \neq 0$), or (2.) a non-reversible (monotonic) change, which can be either represented (2a.) in absolute terms ($C = 0$), or (2b.) as a fraction of the complete metamorphosis ($C = 0, c = 1$). For a more detailed introduction, see Strauss, Altmann ([2013]), and Stachowski (2013: 109f).

Variant 1. will be used as is and commented on in 2.1.1. Variant 2a. will be used in a variation with linguistically more meaningful coefficients (eq. 2, see Stachowski (2013: 110)). Variant 2b. has no application in our case as there is no fixed maximal number of loanwords that a language can possibly absorb.

$$(2) \quad p(t) = \frac{1}{1 + ae^{-b(t-A)}}$$

2. Analysis

This section is composed of two parts, one where the entire Romanian territory is considered en bloc (2.1), and the other where the varieties of specific regions are analysed in isolation (2.2). The material is not treated equally in both parts, see 1.3.

2.1 Romanian as a whole

In this subsection, the Romanian literary language is treated as one with all of its regional and stylistic varieties. Each entry in Suciu (2010) counts as one loanword, regardless of whether it is literary or dialectal, or whether it appears in all regions with the same or different meanings and phonetics. More assumptions apply, see 1.3.

Suciu (2010) collects as many as 2776 words. Unlike most dictionaries, he does not limit himself to the first attestations but also provides the approximate time when the given word ceased to be used in Romanian. This allows the Piotrovskij-Altmann law (see 1.4) to be tested on a new kind of data, and significantly enriches the overall quantitative picture of Ottoman influence on Romanian. In 2.1.1, its total volume is discussed, and in 2.1.2, the longevity of loans from different periods.

2.1.1 Chronology

The Piotrovskij-Altmann law was fitted for three series: the number of loanwords newly appearing during the given century (cumulatively, symbolized ● in Fig. 2, fitted to eq. 2), the number of disappearing loanwords (also cumulatively, symbolized ■, also fitted to eq. 2), and the number of loanwords that were in use at the given time (symbolized ▲, fitted to equation 1). The results have proven to be very precise; see Figure 2 and Table 1.

The last of these series requires additional attention. It was mentioned in 1.3 that Suciu (2009; the monographic volume) and Suciu (2010; the dictionary volume) provide dates that do not match up. Suciu (2010) ignores uncertain attestations and does not inform us about gaps, e.g. *părpăi* ‘Oriental-style white woolen trousers, wrinkled, loose at the top and tight at the bottom, worn by Romanian peasants’ is marked as “probably used” in the first half of the 19th century in Suciu (2009), and as “1884–” in Suciu (2010), while e.g. *hamailău*

‘amulet, talisman’ is marked as first attested in the second half of the 16th century and then throughout the 19th century in Suciu (2009), and as “1594–sf[âr□itul] sec[olului] XIX” in Suciu (2010). In some cases, the nature of the inconsistency is less clear, e.g. *pazar* ‘market, fair; buying and selling, transactions, haggling’ is marked in Suciu (2009) as attested in the second half of the 15th century, and

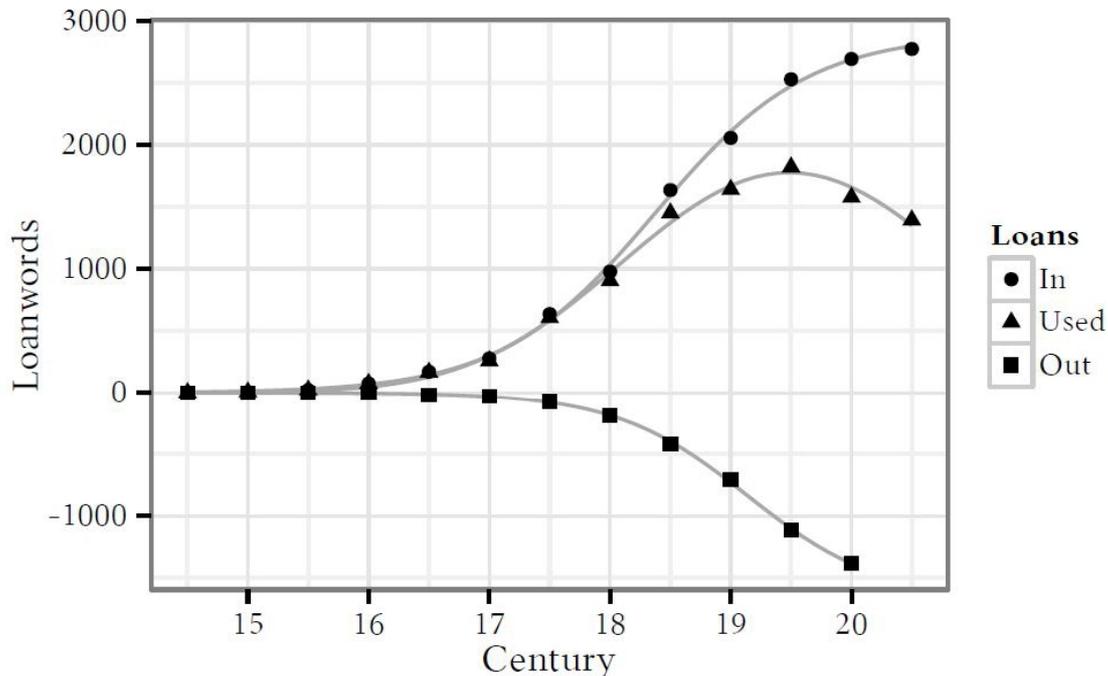


Figure 2. Fitting of the Piotrovskij-Altman equation to the Ottoman-Romanian data. The three series are: loanwords coming into use (●, cumulatively), in use at the time (▲), and going out of use (■, cumulatively, and multiplied by -1 for graphical clarity). The middle series is based on Suciu (2010). Century numbers were decreased by 19 to enable fitting; see the main text. Fitted with R .

then from the first half of the 19th century on, and as “1500–” in Suciu (2010). Also, Suciu (2010: 826f) adds sixteen new words and withdraws one. Both volumes distinguish between words that were normally used in the given period, and those that were seen as dated or obsolete. In the material extracted here from Suciu (2010), both are treated as still in use. Suciu (2009), on the other hand, provides two series of summary dates, one that includes uncertain attestations and such archaic words, and one that does not.

All three series can be accurately modelled with eq. 1, as is shown in Fig. 3. However, in order to achieve the match, century numbers had to be decreased. This is not unacceptable in itself as dates are a pure convention, but it was my arbitrary decision to decrease them by precisely 19. In the case of eq. 2 (non-reversible change, the “In” and “Out” series in Fig. 2), the meanings of coefficients are clear: A represents time (the location of the curve on the horizontal

axis),² b the intensity (the slope of the curve), and c the strength (the height of the curve). With eq. 1 and $C \neq 0$ (reversible change), however, their function cannot be so clearly defined. They hinge on by how much the dates have been decreased, and their variation is also much greater (compare Tables 1 and 3).

Table 1

The empirical (actual) and theoretical (fitted) number of Ottoman loanwords in Romanian. Coefficients for loanwords in use at the time (\blacktriangle) refer to eq. 1; the other two series to eq. 2. The R^2 coefficient of determination represents essentially the proportion of variability in the dependent variable that is accounted for by the model.

Century	● In (cumul.)		▲ Used		■ Out (cumul.)	
	empirical	theoretical	empirical	theoretical	empirical	theoretical
14.5	1	6.6	1	0.9	0	0.4
15.0	6	14.4	6	3.8	0	0.9
15.5	22	31.5	22	14.5	0	2.2
16.0	75	68.2	75	46.8	4	5.5
16.5	170	145.4	166	128.7	20	13.5
17.0	278	300.8	258	299.4	29	33.3
17.5	636	587.1	607	587.5	73	80.3
18.0	979	1036.9	906	973.5	185	186.4
18.5	1637	1593.9	1452	1373.3	416	397.6
19.0	2058	2111.0	1642	1671.9	709	730.7
19.5	2531	2477.4	1822	1778.2	1116	1102.1
20.0	2695	2690.4	1579	1660.2	1381	1385.2
20.5	2776	2800.1	1395	1353.5	NA	NA
	$A = 18.373$		$a = 5.011$		$A = 19.141$	
	$b = 1.571$		$b = 0.306$		$b = 1.821$	
	$c = 2899.267$		$c = 10050$		$c = 1674.76$	
			$C = 0.3149$			
	$R^2 = 0.999$		$R^2 = 0.996$		$R^2 = 0.999$	

² Note that A is in the same units as the original data. In this paper, time is given in centuries, so that e.g. $A = 18.37$ corresponds to the year 1737 rather than *1837.

In their present form, the coefficients do not in themselves add to the linguistic value of the Piotrovskij-Altman law in its “reversible change” variant. It is not clear to me how the formula can be improved. Possibly, it might be beneficial to view nonmonotonic change as a sequence of two monotonic changes, and do away with eq. 1 altogether, concentrating only on eq. 2. One might also want to note, as a separate issue, that all the three series presented in Fig. 3 have in fact a rather sharper peak than the fitted curves suggest. Perhaps just a coincidence, the same phenomenon can be observed in the data collected and analyzed by Imsiepen (1983), unaffected by smoothing in Best, Beöthy, Altman (1990).

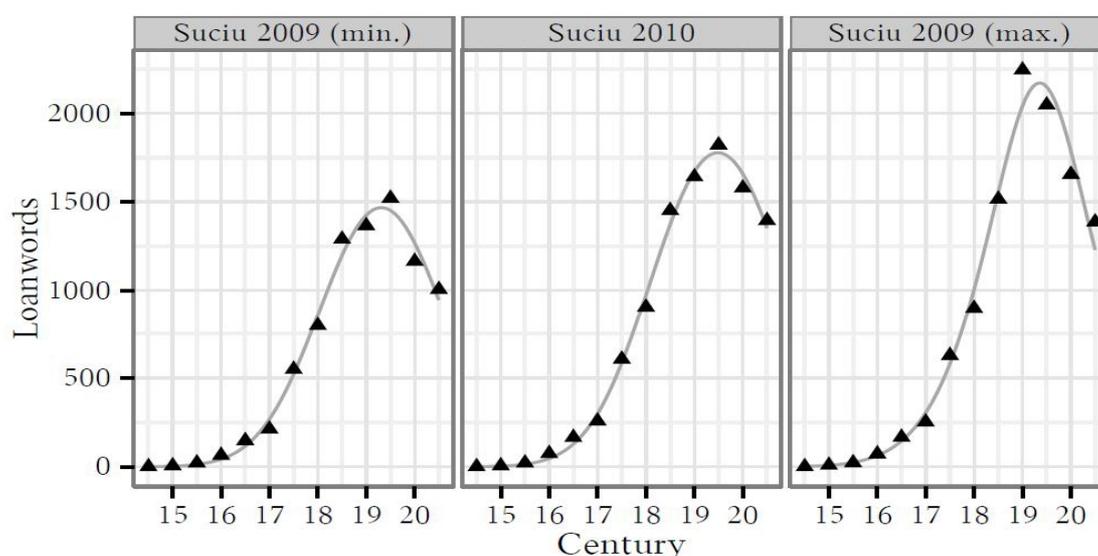


Figure 3. The number of Ottoman loanwords in use in Romanian at different times, and the fitted Piotrovskij-Altman curves (eq. 1). The three series are a result of different selection criteria. Century numbers have been decreased by 19 in all cases, see the main text. Fitted with R .

Table 2

The coefficients used in the fitting of eq. 1 to the number of Ottoman loanwords in use in Romanian at different times. The three series are a result of different selection criteria. Century numbers have been decreased by 19 in all cases, see the main text. Fitted with R .

	Suciu (2009; min.)	Suciu (2010)	Suciu (2009; max.)
a	14.74	5.011	-2.1880
b	0.2043	0.306	0.1849
c	22430	10050	-2426.9160
C	0.3283	0.3149	0.2597
R^2	0.990	0.996	0.987

The very precise fitting obtained here is, no doubt, due in some degree to the fact that many words are only dated with the accuracy of half a century. This forced binning of the data has ironed out many possible outliers and improved the goodness of fit but, at the same time, obscured the correlation with historical events; see also Stachowski (2013: 111f). Naturally, the material being essentially the sum of regional varieties and the literary language, precise correspondence could not have been expected in any case.

What can be observed is that the influx does not gain impetus until the first half of the 17th century and once acquired, maintains it till the beginning of the second half of the 19th century – which is also when the outflux, after having visibly accelerated about a hundred and fifty years earlier, begins to surpass the influx, and the number of loanwords in use starts finally to drop.

A historical explanation can easily be found for the closing date: severely weakened by the end of the 18th century, the Ottoman Empire loses its authority over most of Romanian lands in the first half of the 19th century, and in 1878, they unite and form the independent Kingdom of Romania.

The starting date is less obvious. One may perhaps speculate about the intensification of Ottoman control in that period, but another factor might also need to be taken into account. The oldest surviving document in Romanian is dated 1521, and the 16th century is generally when the literary language begins to form, see e.g. Ghe[□]ie, Mare[□] (1985: 450f). The increase of the number of recorded loanwords does not seem to correlate with one particular event in the political history but it does coincide with the rise of writing in Romanian. It is difficult to estimate to what extent this has affected our results, but the issue certainly deserves attention. A similar phenomenon can be observed in the regional data, see 2.2.1 below.

2.1.2 Longevity

Thanks to the fact that Suci^u (2009, 2010) provides not only the dates of first attestation, but also the approximate time of disappearance of Ottoman loanwords from Romanian, we are allowed a rare chance to inspect their longevity. Suci^u himself devotes much thought to this question in respective chapters in (2009: 109f), but he does not seem to proceed to make general statements.

Fig. 4 might create the impression that, with the exception of the 20th century, the later the borrowing took place, the more short-lived it has proven to be. A detailed examination of these words, however, fails to reveal clearly why it should be so.

Let us divide Fig. 4 into three parts: the first stretching from the very beginning to the first half of the 18th century, the second from the second half of the 18th to the second half of the 19th century, and the last encompassing only the first half of the 20th century.

Politically, the first part covers the period of almost undisputable domination of the Porte over most or all of the Romanian regions. Similarly to what was said in 2.1.1 above, one might wish to explain the visible rise in the number of short-lived loanwords during that period by speculating about tightening of the Ottoman grip. But again, if the history of Romanian writing is factored in, together with the fact that the oldest surviving purely Romanian documents are not older than the 16th century, another explanation becomes available, namely that the distribution of longevity of borrowings might have been in fact relatively stable throughout the period, only the oldest examples were simply either never attested or the documents that contained them did not survive till our times.

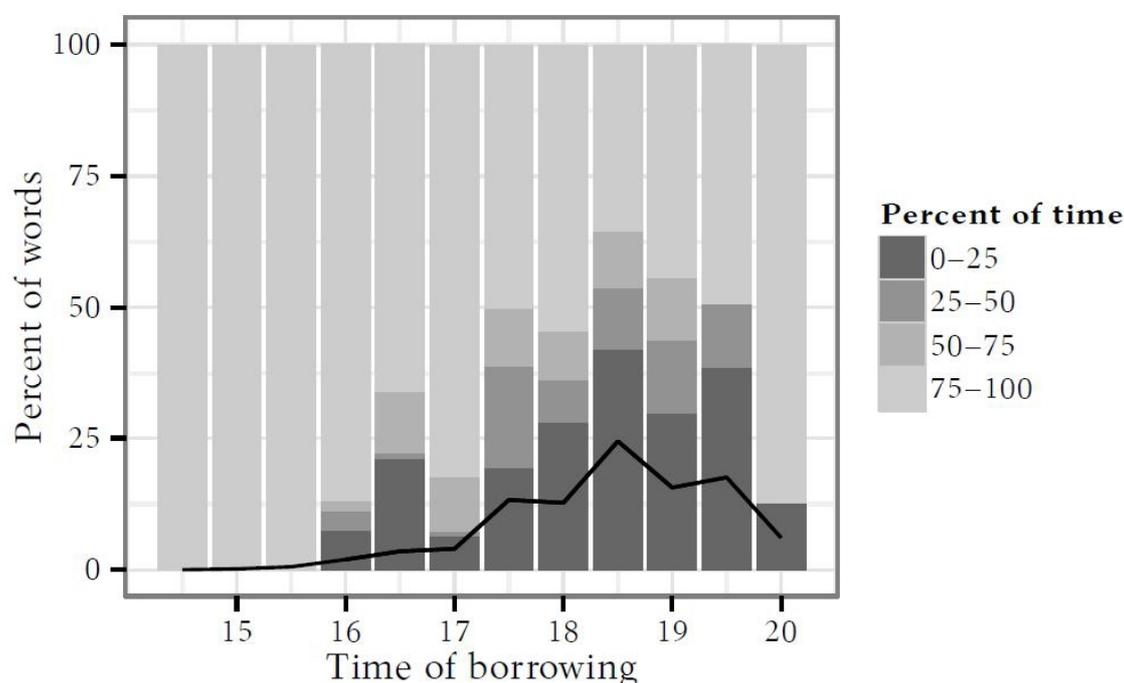


Figure 4. Longevity of Ottoman loanwords in Romanian. Bars represent the percent of words borrowed at the given time that have survived in the language for a specific percent of the time till nowadays. For example, all of the words borrowed in the second half of the 14th century (the bar at 14.5) have remained in Romanian for 75–100% of the six centuries that have passed since their appearance; at the same time, more than 27% of words borrowed in the second half of the 18th century have disappeared after 0–25% of the two centuries that have since passed. Words marked as *historic* by E. Suciú are not treated in a special way. The line shows what percentage of the total number of loanwords comes from the given time. For example, almost a quarter of all Ottoman loanwords that were ever in Romanian were borrowed in the second half of the 18th century.

The second period is the time of the fight for independence, and of much increased volume of writing. Possibly, many loanwords were simply the need of the hour, that happened to have appeared at the right time to become recorded.

Finally, in the last period the representation of longevity as a percentage somewhat breaks down. With the best available accuracy of half a century, words borrowed in the first half of the 20th century could have either survived for 0% of the maximum time, i.e. disappeared in the same period, or 100% of the maximum time, i.e. still have existed in the second half of the century. The visible drop in the former category is surely a result of regaining independence in 1878.

It remains to be observed that the absolute number of loans in the given period correlates quite well with the number of the most short-lived and that of the most long-lived borrowings (Spearman's $\rho = 0.979$ and 0.986 , respectively), but not so well with the two middle categories ($\rho = 0.856$ and 0.614). The relation appears to be logarithmic rather than linear in nature, but requires a more detailed examination before conclusions can be drawn or explanations suggested.

2.2 Regional varieties

In this subsection the material collected in Suciu (2010) is treated as belonging to separate varieties. (The term *dialect* is deliberately not used here as Suciu's qualifiers refer to historical regions rather than the still disputed dialectal division of Romanian, see e.g. Ursan (2008).) Individual meanings are considered independent words (see 1.3 for justification), and only the words with regional qualifiers are taken into account. Also the words marked as "literary" are discarded because they are often additionally labelled as being used "today mostly in ..." which suggests that even when known in other regions, they do not strictly belong to the regional speech.

By such criteria, Suciu (2010) contains as many as 1513 words from six regions: Banat, Dobruja, Moldavia, Muntenia, Oltenia, and Transylvania. Occasionally, locations beyond the core of the modern range of Romanian are marked. They will not be discussed here, and the six words that are not attested outside of these areas will be ignored (northwestern Bulgaria: *chimur*, *peringi*, northern Serbia: *pai*, *tere*, southern Ukraine: *otac*, *tarciniu*). Additional geographical specifications are also sometimes given, such as "eastern", "north-western", &c. They will not be taken into account here except in "northern Moldavia" which denotes 'southern Ukraine' and "eastern Moldavia" which stands for the 'Republic of Moldova', see Suciu (2010: 6). The former will be ignored, as was stated before, and the latter will be treated as a separate region for historical reasons (independent from Romania since 1947).

Regional collections are not nearly as rich as the general Romanian material. Apparently, the historic data are particularly incomplete, as will be seen in 2.2.1. For this reason, longevity of borrowings will not be discussed, and the

subsequent parts will be limited to only those loanwords that are still in use today.

2.2.1 Chronology

The general Romanian material was analyzed with respect to both the time of appearance and that of disappearance of loanwords (2.1.1). The regional data, however, do not appear to be sufficiently representative.

It can be seen from Fig. 5 that the Piotrovskij-Altman law can very accurately model the influx of Ottoman loanwords into regional varieties of Romanian. The correlation with historic events, however, is below expectations.

The rise of the rate of influx begins before suzerainty in Banat and Transylvania, and inversely in the remaining five regions. But in all the regions, it can be said to gain impetus in the first half of the 16th century, i.e. together with the growth of writing in Romanian. Similarly then to the general material discussed in 2.1 above, the historic data appear to reflect more the volume of Romanian writing than that of Ottoman influence.

Data on the decrease of the rate of influx come from newer and better attested times, and appear to be more representative. In almost all the regions, the period of rapid growth ends with the second half of the 19th century, i.e. just when Romania gained independence. Only Dobruja is an exception, most probably because Turkic-speaking peoples remained a sizeable part of its population well into the 20th century. If predictions can be made based on the Piotrovskij-Altman law, a significant drop in the rate of influx is not to be expected before mid-22nd century. Intuitively, a more imminent decrease would seem more likely.

Banat and Transylvania were under Ottoman authority for a relatively short time. The former was fully incorporated into the Empire, the latter managed to maintain a degree of autonomy, but the influx of Ottoman words proceeded in both at a very similar, staid pace. In all the other regions, a sharper spike can be observed in the second half of the 19th century. Supposedly, it was caused primarily by B.P. Hasdeu's extensive dialectological questionnaire started in 1884 (the manuscript of the responses comprised eighteen volumes, Suciu (2010: 17) s.v. *H*), L. ăineanu's *Influența orientală asupra limbii și culturii române* (1900), and by other publications prompted by the general growth of interest in folklore at the time. The political situation and fight for independence might have contributed to the zeitgeist but they would be an unlikely cause of the increased influx of Ottoman loanwords.

Overall, the regional data are strongly influenced by the volume of Romanian writing at the given time. Pre-16th century parts do not at all seem complete, but their accuracy appears to improve with time, and they might be expected to have reached a fair degree of representativeness with the end of the 19th century.

For material ranging from the second half of the 14th to the second half of the 20th century, this is quite late. The number of loanwords going out of use, therefore, and the number of loanwords currently in use at different times, will not be discussed here.

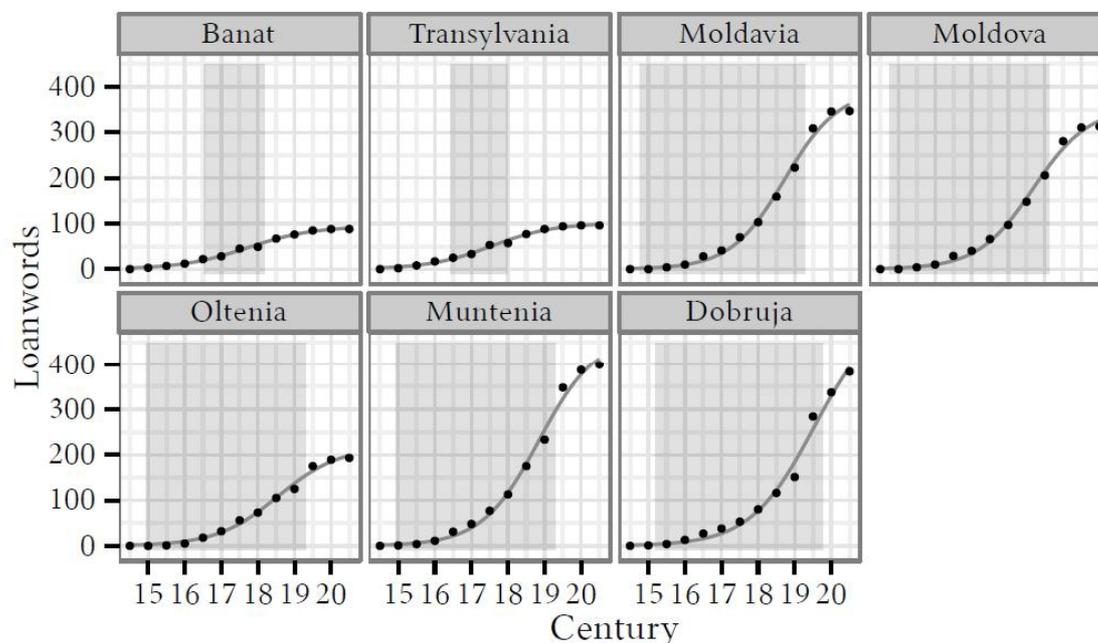


Figure 5. The influx of Ottoman loanwords into regional varieties of Romanian, and the fitted Piotrovskij-Altman curves (eq. 2). Regions are ordered to resemble the geographical distribution rather than alphabetically. Grey rectangles show the duration of Ottoman authority over respective regions; short episodes of dependency to other countries are not marked (e.g. Austrian rule in Oltenia in the years 1716–39, see 1.1). Fitted with R .

Table 3

The coefficients used in the fitting of eq. 2 to the influx of Ottoman loanwords into regional varieties of Romanian. Fitted with R .

	A	b	c	R^2
Banat	17.717	1.114	94.045	0.995
Dobruja	19.532	1.145	522.093	0.989
Moldavia	18.726	1.353	394.402	0.995
Moldova	18.687	1.324	355.273	0.995
Muntenia	18.849	1.281	463.288	0.994
Oltenia	18.554	1.187	217.97	0.994
Transylvania	17.553	1.166	101.446	0.994

2.2.2 Duration and volume of influence

One might be tempted to link the volume of influence to the duration of Ottoman authority over individual regions, and the time that has elapsed since its end. Historical data not being fully representative (see 2.2.1), let us limit the test to those words only that are still in use today.

Settling on the exact number of loanwords was shown in 1.2 to be a neither obvious nor an entirely objective procedure, and so is determining the precise number of years a region was, and then was not, under Ottoman authority. Some were incorporated into the Empire; others retained varying degrees of autonomy; in some the Turkic presence was considerably more visible at certain times than in others. The brief summary in 1.1 cannot do justice to the complex history of the region.

For the purpose of the experiment, the numbers shown in Fig. 6c were used. Two formulas were tested, yielding the following relations (l stands for ‘loanwords in use today’, t_O for ‘duration of Ottoman authority’, and t_n for ‘time since the end of Ottoman authority’; fitted with R):

$$(3a) \quad l \approx 0.4384 \cdot t_O - 0.6382 \cdot t_n + 205.1582$$

$$(3b) \quad l \approx 0.5020 \cdot (t_O - t_n) + 154.881$$

The accuracy is effectively identical in the two cases, $R^2 = 0.784$. This is a disappointingly low result. As it happens, the correlation between the mutually independent t_O and t_n dwarfs them both (Pearson’s $\rho = -0.9628736$).

It is possible that the problem lies in precision. Muntenia and Oltenia are often regarded as merely the two parts that together constitute Wallachia. This is a perfectly justified practice from the historical point of view, and when applied to our case, it yields the following relations:

$$(4a) \quad l \approx 0.6824 \cdot t_O - 0.3436 \cdot t_n + 82.5277$$

$$(4b) \quad l \approx 0.57398 \cdot (t_O - t_n) + 167.72428$$

which have a considerably improved accuracy of $R^2 = 0.908$ for eq. 4a, and 0.907 for 4b. This effect must be akin to what is observed when fitting the Piotrovskij-Altman law to grouped, and to ungrouped data (see 2.1.1).

2.2.3 Similarity of regions

The duration and extent of Ottoman authority differed considerably from region to region. The unification of the entire Romanian territory happened in several steps over almost a century, and was followed by the loss of Moldova and parts of Bukovina and Dobruja in 1947. It is to be expected, then, that some regions

will be more similar to one another than to others. As in 2.2.2 above, the discussion will be limited to words that are currently still in use.

Perhaps the most conspicuous feature in Fig. 7 is that the similarities appear to more reflect historical divisions than geographical proximity, which suggests a relatively low level of communication between the regional varieties even during the past century when they were parts of one state. Three partitions are especially visible: Transylvania with Banat, Moldavia with Moldova, and the remaining four regions which formed the Kingdom of Romania in the years 1878–1913.

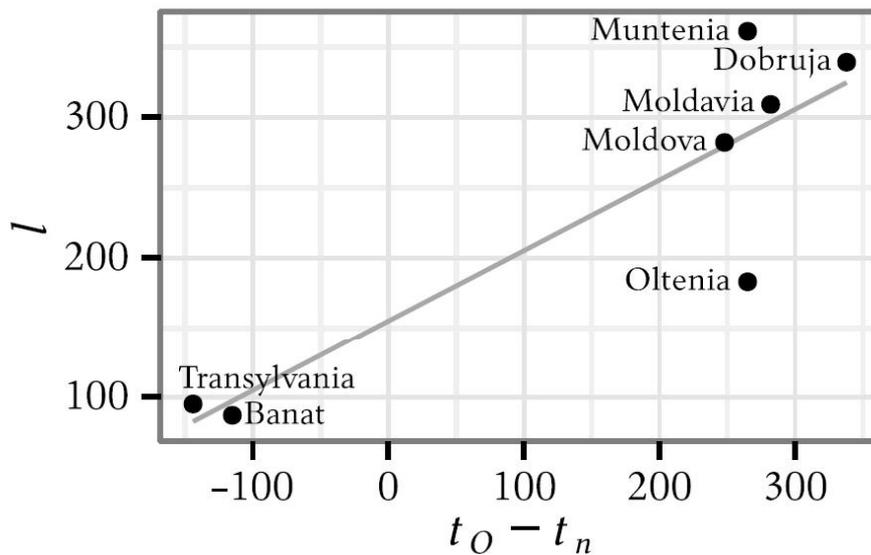


Figure 6a. Muntenia and Oltenia considered separately (eq. 3b). See Table 4.

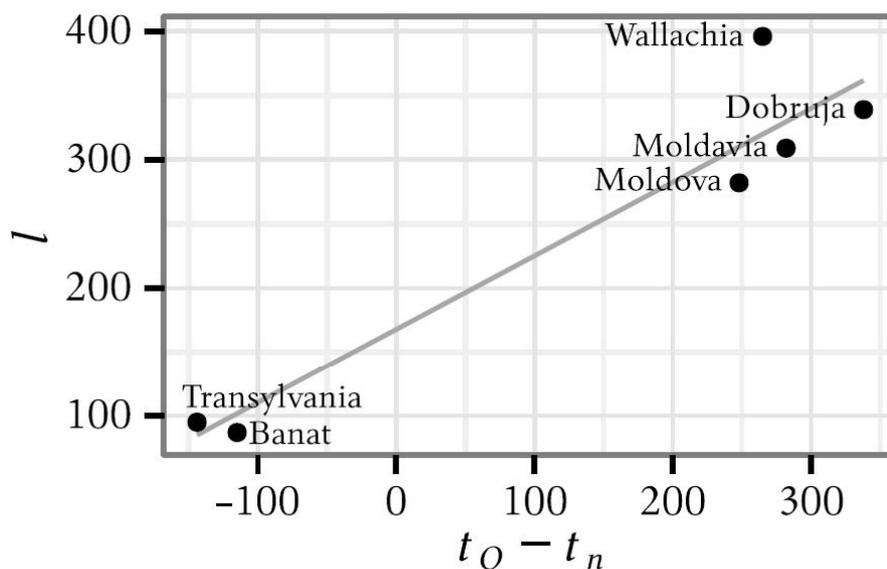


Figure 6b. Muntenia and Oltenia considered together (eq. 4b). See Table 4.

Banat is more alienated than any other region. Its stock of Ottoman loanwords only resembles that of Transylvania, but not at all even that of the neighbouring Oltenia. This could be explained by its peripheral location, but Transylvania which lies in the very centre of the country also does not really seem to have assimilated its stock to Moldavia's or Wallachia's, or acted as a middleman between the two. Banat and Transylvania share a large part of their history, and were only officially tied to the other regions after World War I, although they have been populated in a considerable proportion by Romanians for at least several centuries. Note also that the separateness of the Banatian dialect does not raise doubts, while the Transylvanian variety has eluded clear categorization for a very long time (see 1.2).

Table 4

The relation between the number of Ottoman loanwords in use today in different regions (l), the time these regions remained under Ottoman authority (t_o), and the time that has since passed (t_n), with reference to the year 1999, as Suciu's data covers the period till the end of the 20th century). See also Figures 6a and 6b.

	l	t_o	t_n
Banat	87	166	281
Dobruja	339	459	121
Moldavia	309	452	170
Moldova	282	435	187
Muntenia	361	435	170
Oltenia	183	435	170
Transylvania	95	156	300
Wallachia	396	435	170

Moldavia and Moldova were considered one region for some time. In fact, Suciu (2010) refers to the latter as 'eastern Moldavia'. In the years 1947–91, however, it was part of the USSR and this proved sufficient to kindle a national consciousness. In the sphere of Ottoman loanwords, nonetheless, the effect appears to have been minimal. The stock is very similar in the two regions, and has relatively many similarities to the other parts of the first Kingdom of Romania. Only Moldova and Dobruja are further part, due perhaps to the geographical wedge of Budjak.

Muntenia and Oltenia are commonly seen as parts of one region, therefore their relative proximity does not surprise, and neither does the fact the the stock of the eastern one of the two, Muntenia's, is more similar to the stock of Dobruja. Oltenia's similarity to Moldavia and Moldova is perhaps slightly more unusual as

they do not have a shared border. The mediation must have occurred via Muntenia. Geographical proximity appears to have only played a more important role in the case of Dobruja.

Table 5

The similarity of the stocks of Ottoman loanwords between Romanian regions.

The index used is the Jaccard distance, $\frac{|A \cap B|}{|A \cup B|}$, where 1 denotes identity, and 0 no similarity at all.

	Banat	Dobruja	Moldavia	Moldova	Muntenia	Oltenia	Transylvania
Banat	1						
Dobruja	0.052	1					
Moldavia	0.073	0.232	1				
Moldova	0.076	0.190	0.870	1			
Muntenia	0.067	0.296	0.376	0.331	1		
Oltenia	0.098	0.165	0.224	0.214	0.374	1	
Transylvania	0.422	0.061	0.107	0.102	0.096	0.121	1
all	0.104	0.262	0.418	0.388	0.380	0.219	0.115

3 Conclusions

Although unusually rich, the material collected in Suciu (2009, 2010) has not proven to be sufficiently representative for a quantitative diachronic analysis. Especially older parts of the collection depend heavily on the volume of writing in Romanian that has survived to our times. The problem affects the general Romanian and the regional data equally (2.1.1–2.1.2, 2.2.1). The issue is not of major importance for the present paper, but it highlights a possible weakness in some of the results obtained so far for the Piotrovskij-Altmann law.

Remarks were also made in 2.1.1 about the “reversible” variant of the law. It was pointed out that its coefficients lack clear linguistic meaning, and that the resulting curve fails to reflect what might be a characteristic spike in nonmonotonic phenomena in language.

Regardless of these shortcomings, the discussion in 2.1.2 appears to suggest that the total number of loanwords from a certain period depends primarily on those borrowings which are about to disappear very quickly, and those which are

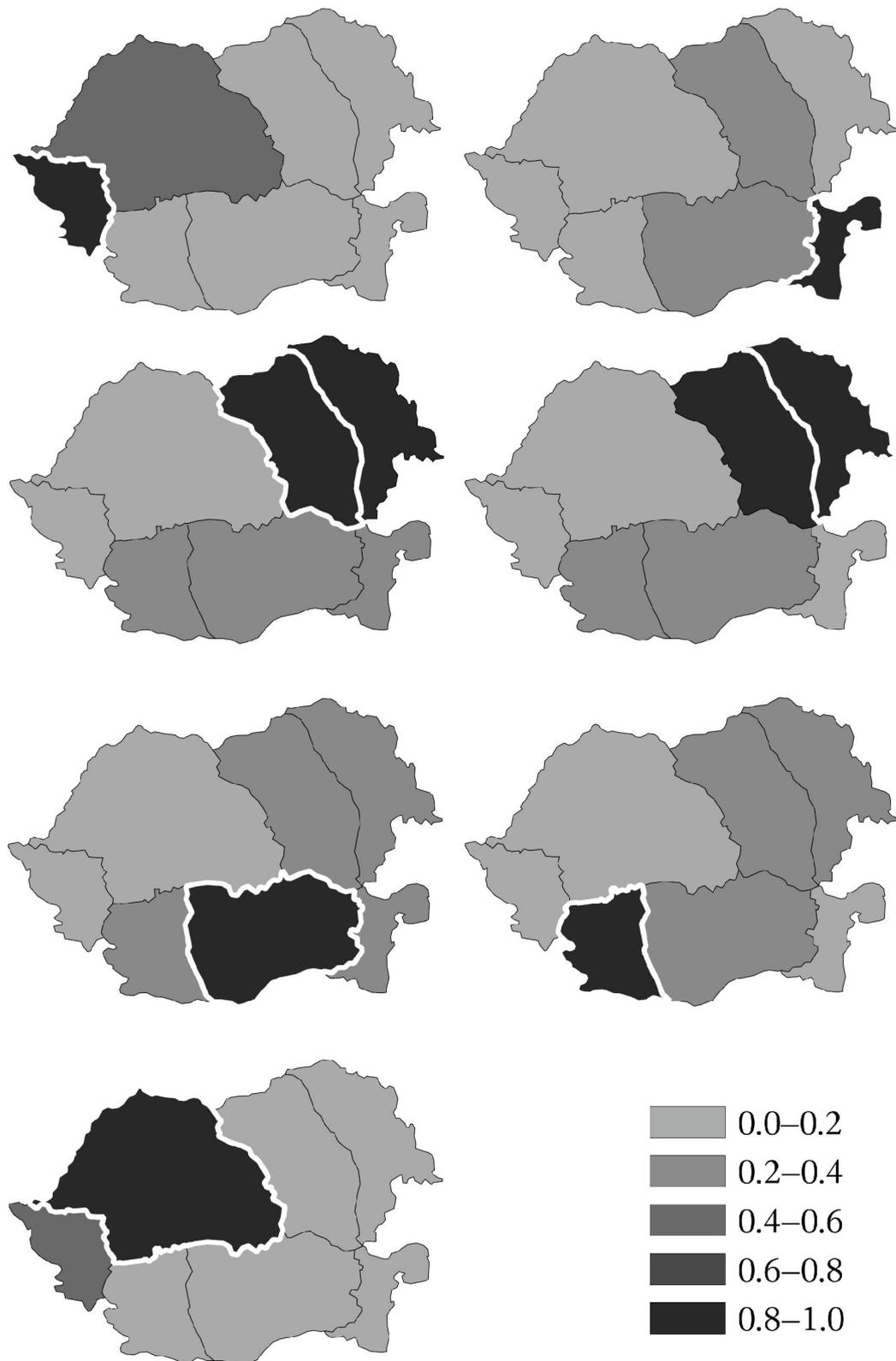


Figure 7. The similarity of the stock of Ottoman loanwords in the highlighted, and the remaining regions of Romania. The index used is the Jaccard distance, J , where 1 denotes identity, and 0 no similarity at all.

to survive for a rather long time. The medium-lived ones seem to be the least numerous group, and more randomly distributed across the centuries.

The number of surviving loanwords was shown in 2.2.2 to bear some correlation with the duration of Ottoman authority, and the time that has elapsed since its end. Two regions, Muntenia and Oltenia, did not quite conform to the tendency set by the other five. When they were considered together, however, as is often the historically justified practice, the accuracy of the fitting significantly improved.

Lastly, the stock of Ottoman loanwords in different regions was compared in 2.2.3 to reveal that the old political partitions have apparently had a greater impact on transmission of borrowed vocabulary than did geographical proximity. The present paper does by no means exhaust the topic or the material contained in Suciú (2009, 2010). This would require a much larger enterprise. Here, just an overall quantitative characteristic has been presented, together with some theoretical observations, and hopefully a way of combining qualitative and quantitative research and findings to obtain a new kind of insights.

Abbreviations

Alb. = Albanian | **Ban.** = Banatian | **Dobr.** = Dobrujan | **Munt.** = Muntenian | **SCr.** = Serbo-Croat

References

- Ágoston, G., Masters, B. (eds.) (2009). *Encyclopedia of the Ottoman Empire*. New York: Facts On File.
- Altmann, G. (1983). Das Piotrowski-Gesetz und seine Verallgemeinerungen. In: In: Best, K.-H., Kohlhase, J. (eds.), *Exakte Sprachwandelforschung. Theoretische Beiträge, statistische Analysen und Arbeitsberichte*. Göttingen: Edition Herodot, 59–90.
- Best, K.-H. (2005). Turzismen im Deutschen. *Glottometrics 11*, 56–63.
- Best, K.-H., Beöthy, E., Altmann, G. (1990). Ein methodischer Beitrag zum Piotrowski-Gesetz. *Glottometrika 12*, 115–124.
- Best, K.-H., Kohlhase, J. (eds.), (1983). *Exakte Sprachwandelforschung. Theoretische Beiträge, statistische Analysen und Arbeitsberichte*. Göttingen: Edition Herodot.
- Gheşie, I., Mareş, A. (1985). *Originile scrisului în limba română*, Bucureşti, Editura ştiinţifică şi Enciclopedică.
- Hřebíček, L. (1990). Quantitative studies. In: Hazai, G. (ed.), *Handbuch der türkischen Sprachwissenschaft 1*, Budapest: Akadémiai Kiadó, 371–387.
- Imsiepen, U. (1983). Die *e*-Epithese bei starken Verben im Deutschen. In: Best, K.-H., Kohlhase, J. (eds.), *Exakte Sprachwandelforschung. Theoretische*

- Beiträge, statistische Analysen und Arbeitsberichte*. Göttingen: Edition Herodot, 119–141.
- Leschber, C.** (2011). Lehnwege einiger Orientalismen und Wörter eurasischer Herkunft im Rumänischen und den sonstigen Balkansprachen. *Studia Etymologica Cracoviensia* 16, 33–61.
- Mitu, M.** (2002). Orientalizmy leksykalne w języku polskim i rumuńskim. In: Rusek, J., Boryś, W., Bednarczuk, L. (eds.), *Dzieje Słowian w świetle leksyki*, Kraków: Wydawnictwo Uniwersytetu Jagiellońskiego, 301–306.
- Pop, I.-A., Bolovan, I.** (eds.) (2006). *History of Romania. Compendium*. Cluj-Napoca: Romanian Cultural Institute. Center for Transylvanian Studies.
- Sallanz, J.** (ed.) (2005). *Die Dobrudscha: Ethnische Minderheiten – Kulturlandschaft – Transformation. Ergebnisse eines Geländekurses des Instituts für Geographie der Universität Potsdam im Südosten Rumäniens (= Praxis Kultur- und Sozialgeographie 35)*. Potsdam: Universitätsverlag Potsdam.
- Stachowski, K.** (2013). The influx rate of Turkic glosses in Hungarian and Polish post-mediaeval texts. In: Köhler, R., Altmann, G. (eds.), *Issues in Quantitative Linguistics 3, dedicated to Karl-Heinz Best on the occasion of his 70th birthday*. Lüdenscheid: RAM-Verlag, 100–116.
- Strauss, U., Altmann G.**, *Change in language*. In: *Laws in Quantitative Linguistics*. http://lql.uni-trier.de/index.php/Change_in_language [July 2013]
- Suciu, E.** (2009). *Influența turcă asupra limbii române, 1: Studiu monographic*. București: Editura Academiei Române.
- Suciu, E.** (2010). *Influența turcă asupra limbii române, 2: Dicționarul cuvintelor românești de origine turcă*. București: Editura Academiei Române.
- Sugar, P. F.** (1977). *Southeastern Europe under Ottoman Rule, 1354–1804*. Washington: University of Washington Press.
- Ursan, V.** (2008). Despre configurația dialectală a dacoromânei actuale. *Transilvania* 1, 77–85.