Some Properties of Rhyme

by

Mihaiela Lupea
Maria Rukk
Ioan-Iovitz Popescu
Gabriel Altmann

Studies in Quantitative Linguistics

Editors

Fengxiang Fan   (fanfengxiang@yahoo.com)
Emmerich Kelih   (emmerich.kelih@univie.ac.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 Vol. 4*. 2014, VIII+148 pp.
15. Best, K.-H., Kelih, E. (eds.), *Entlehnungen und Fremdwörter: Quantitative Aspekte.* 2014. VI + 163 pp.
16. I.-I. Popescu, K.-H. Best, G. Altmann, G. *Unified Modeling of Length in Language.* 2014, VIII + 123 pp.
17. G. Altmann, R. Čech, J. Mačutek, L. Uhlířová (eds.), *Empirical Approaches to Text and Language Analysis dedicated to Luděk Hřebíček on the occasion of his 80th birthday.* 2014. VI + 231 pp.
18. M. Kubát, V. Matlach, R. Čech, *QUITA Quantitative Index Text Analyzer*. 2014, VII + 106 pp.
19. K.-H. Best, *Studien zur Geschichte der Quantitativen Linguistik*. 2015. III + 158 pp.

20. P. Zörnig, K. Stachowski, I.-I. Popescu, T. Mosavi Miangah, P. Mohanty, E. Kelih, R. Chen, G. Altmann, *Descriptiveness, Activity and Nominality in Formalized Text Sequences.* 2015. IV+120 pp.

21. G. Altmann, *Problems in Quantitative Linguistics Vol. 5*. 2015. III+146 pp.

22. P. Zörnig, K. Stachowski, I.-I. Popescu, T. Mosavi Miangah, R. Chen, G. Altmann, *Positional Occurrences in Texts: Weighted Consensus Strings.* 2015. II+178 pp.

23. E. Kelih, R. Knight, J. Mačutek, A.Wilson (eds.), *Issues in Quantitative Linguistics Vol. 4*. 2016. III + 231 pp.

24. J. Léon, S. Loiseau (eds.), *History of quantitative linguistics in France*. 2016. II + 232 pp.

25. K.-H. Best, O. Rottmann, *Quantitative Linguistics, an Invitation*. 2017, III+171 pp.

Contents

1. Introduction

It is impossible to write a full analysis even of a short text. A text, just like the world, has as many aspects as we are able to discover or to define in it. Since science develops, the number of aspects increases, too. The aspects are determined by us according to our interest. Something in the text is important for literary science, something else for politics, culture, psychology, cooking, or the news reported by it. There are many books written about text analysis; in them, one has discovered a lot of regularities some of which attained the status of laws, i.e. they have been derived from a theory constructed in the background, its individual hypotheses have been positively tested and one has found links between the given and other properties of the text. Quantitative text analysis becomes slowly unfathomable and the production of books increases quickly using mechanical computing machines. Here we shall restrict ourselves to a smaller circle of problems, namely those of rhyme and shall not try to construct a theory but rather derive some concrete hypotheses and describe some properties of the given texts. The result should help to make a step into the depth.

Rhyme is represented by the phonetic similarity of words occurring – in our case – at the end of the lines of a poem. There are, of course, also internal or initial rhymes but we can omit them here. Since they are constructed of words or parts of words, e.g. syllables, morphemes, clitics, or simply vowels, they have a number of properties many of which have been at least stated and classified. There are descriptions in many languages. Inventories of properties can be found also on the Internet in many languages (e.g.https://de.wikipedia.org/wiki/ Reim# Reime_nach_ phonologischer_Struktur or https://en.wikipedia.org/wiki/ Rhyme). Needless to say, in every language there are also poems without rhyme.

Rhyme, just as any other linguistic entity, has as many properties as we are able to define. If it consists of words, it has all the properties of a word. These properties can be quantified, measured, and ordered according to size and one can fit a mathematical model to this order. One fits it either tentatively, inductively, i.e. one seeks a function or distribution which expresses the order of numbers adequately; or one works deductively and starts with an hypothesis containing some factors which are active during the writing of the text.

Rhyme is a linguistic construct and as such it is a part of a dynamic system – called language – underlying self-organization and self-regulation. The simplest process concerning rhyme is the development of its properties in time. One can conjecture that without any development, rhymes can become trivial, expected and aesthetically inefficient. In order to avoid this state, there are already rhyme dictionaries in some languages. Since rhyme has an aesthetic function, it must change its form just as anything else in art. Further, since it is a linguistic structure, its properties must be in some association with other linguistic properties; that means, there are some links between the rhyme-type and the other words of the verse, or between the rhyme and some properties of language, or between the various properties of rhyme. However, it is not always easy to find them. Are the rhyme properties associated with the properties of the given poem?

But how should one define the properties of poems? And what is more: how do we quantify and measure them?

The classification of rhymes is merely a problem of definitions, the first activity in any type of research, but the quantification of their properties, interrelations and development is a continuation of this first step landing finally somewhere in mathematics. Needless to say, one cannot scrutinize everything at once but one can show a way which could be followed, corroborated or rejected. One can try to look behind the curtain of surface and make a step into the depth.

Here we shall be concerned with the numeric evaluation of some properties both with individual writers and in language as a whole. There are few historical investigations up to now. We can mention e.g. the study of the evolution of the open rhyme in Slovak which displays a clear tendency (cf. Štukovský, Altmann 1965, 1966). Does it exist in other languages, too?

In what follows, we shall investigate the length of the rhyme word, its appartenance to a part of speech, the open and closed rhyme words, and the pairs of rhyme words; we shall compare poems, writers, languages, types of poetry. We shall try to quantify each aspect, set up models and perform significance tests. It is not possible to study all forms of rhyme because the number of types is enormous and we must restrict ourselves to a small number of languages. For each problem we use some poems in different languages but not necessarily all of our data. A detailed analysis of all poems by the Romanian poet M.Eminescu has already been presented (cf. Popescu et al. 2015), here we shall try to generalize it.

It must be remarked that the last word of the line is taken into account only if it forms a rhyme with the last word of some other verse. The not rhymed verses were simply omitted; if three verses formed a rhyme than all three were counted and there were two subsequent pairs.

As will be shown, the mathematical modeling of rhyme properties is no simple matter. Since we have to do with different languages, different authors who wrote the poems in their private development, different age, different contents, different styles, etc., we must expect that every point of view brings at least one model, and there will always be exceptions that cannot be satisfactorily captured. But the striving for high artistic performance forces the poet to go his own way which is not the same as with other authors. From our point of view, we strive merely to capture formally the regularities – if there are any – leaving the comment concerning exceptions to a collaboration of literary scientists and "qualitative" linguists.

A very complex question which has not been even touched up to now is the relation between rhyme words. The question is: are the rhymed words in some relation to one another? Consider for example the rhymed words in the first strophe of *Der Erlkönig* by Goethe: *Wind – Kind; Arm – warm* (wind – child; arm – warm). There is surely no semantic relation. But there may be poems in which one finds something. Hence poems can be distinguished also according to the degree of semantic relations of rhymed words. The question is how to measure the degree of semantic or grammatical, or phonetic relation. Could one take into account historical issues, earlier occurrence together, the extent of phonetic

similarity, etc. Could one consider the rhyme words as belonging to various hrebs (= entities defined by L. Hřebíček) and measure their weight? An analysis of hrebs in Romanian can be found in Tătar; Lupea; Altmann 2014.

One may obtain an indicator which has many properties, but the differences must be testable and one must place this phenomenon in some control cycle. There are a number of questions that may be asked and we hope that these aspects will be inserted in a future rhyme theory.

The types of rhyme may depend both on the writer, on the language and its evolution, on the poetry type, onomatopoetic coloring, etc. It is not possible to study all aspects but step by step one can obtain individual results. It can even be expected (and hoped) that for each aspect of rhyme there will be created a different theory.