The Köhler knick: Prolegomena on the synergetics of academic and bio-demographic politics

Peter Grzybek

1 The problem

This contribution is of both analytic and historiographic orientation. It intends to shed some light on a specific chapter of the history of quantitative and synergetic linguistics, which has hitherto been largely neglected; at least, it has remained generally ignored and slipped the due academic attention. Yet, it deserves to be treated in detail, the more since it offers an innovative approach to the synergetic integration of historical, geo-political, methodological and biographical information, culminating in Reinhard Köhler's works. Taking into account the multi-disciplinary and multi-factorial aspects of this issue, it is no wonder that it took even him 60 years to come to terms with the complexity of this matter.

In the course of this contribution, an attempt shall be made to analyze and explain Reinhard Köhler's personal and academic success at the University of Trier, a more or less small and (as some people would think) provincial town with ca. 100 000 inhabitants in the German Federal State of Rheinland-Pfalz. More specifically, the demographic development of Trier's population size since Reinhard's personal presence there will be the focus of this contribution. It is an open secret that there was a remarkable rise of population in Trier, after Reinhard came there, and rumors never ceased that this was due to (some people say: thanks to) a very special kind of personal engagement of his among particular parts of Trier inhabitants. It is not the right place here to go into further details, as far as these rumors and their background are concerned. Instead, it is a major objective of this contribution to rebut these rumors and to demonstrate that they are absolutely unfounded. Rather, the true reasons of Trier's demographic development will turn out to be located on a very different (much higher) level of abstraction, and it will be seen that they are closely related to theoretical interests of Reinhard's in linguistic and synergetic modeling, rather than trivial and earthbound interests, as some ill- or short-minded people have previously tended to assume.

Before going into details with regard to the city of Trier – which, in a particular sense, will be our 'text' to be analyzed – it may seem worthwhile, however, to call into the readers' minds some basic cornerstones of the broader circumstances, i.e. the Federal State of Rheinland-Pfalz (our 'context').

2 The context: Rheinland-Pfalz

2.1 Historical Foundations

Rheinland-Pfalz – sometimes in English texts, the translation *Rhineland-Palatinate* is preferred – today is one of the sixteen states of the Federal Republic of Germany, with ca. four million inhabitants. Situated in Western Germany, Rhineland-Pfalz borders a number of other Federal States and countries. From the north and clockwise these are: North Rhine-Westphalia, Hessen, Baden-Württemberg, France, Saarland, Luxembourg and Belgium.

Rheinland-Pfalz was established as a Federal State on August 30, 1946; the constitution was legally confirmed by referendum on May 18, 1947. It was formed after World War II from the Northern part of the French Occupation Zone which, among others, included not only parts of Bayern (so-called *Bavaria*), namely, the Pfalz (also known as *Rhenish Palatinate*), but also the southern parts of the Prussian Rhine Province (Rheinprovinz, or Rheinpreußen), and parts of Rheinhessen (Rhinehessen), on the western banks of the Rhine.

2.2 Rheinland-Pfalz: Population development

At the time of its formal establishment as a Federal State, approximately three million inhabitants were living on the territory of the area of Rheinland-Pfalz, which covers ca. 20 000 square kilometers. In order to better historically understand the situation, Table 1 presents data covering the time from 1871 to our days, thus offering a general outline of the demographic development of Rhineland-Pfalz.

Table 1: Demographic development of Rhineland-Pfalz (1871–2010)

Year	Inhabitants	Year	Inhabitants
1871	1 832 338	1975	3 665 777
1900	2 303 725	1980	3 642 482
1910	2 541 697	1985	3 615 049
1925	2 690 165	1990	3 763 510
1933	2 841 395	1995	3 977 919
1939	2 959 994	2000	4 034 557
1950	3 004 784	2005	4 058 843
1961	3 417 116	2010	4 003 745
1970	3 645 437		

The data are graphically illustrated in in Figure 1, which contains the year data in the original form, as given in Table 1.

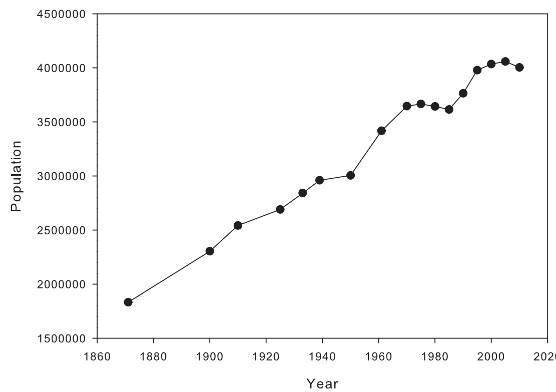

Figure 1: Demographic development of Rhineland-Pfalz (1871-2010)

Generally speaking, there seems to be a constant rise of population over the whole time span; this is, without a doubt, the major trend. Yet, at closer inspection, two periods can be detected which slightly deviate from the overall trend, so that, on the whole, one can speak of three trends, or phases, within the overall development:

1. it can clearly be seen that there was a more or less constant rise of population until the 1970s;
2. there is an obvious decline of population starting in the 1970s, and ending in the 1980s;
3. third, there is a clear increase of population, beginning in the later 1980s, which seems to be slightly weakened only at its very end, i.e., in our days.

The distinction of these phases will provide important background information for us, when we take a closer look at the development of the city of Trier. Trier is not, by the way, the capital of Rheinland-Pfalz. Instead, Mainz with its 197 778 inhabitants (in 2010), is held to be the official capital. In terms of population size, Mainz is followed by Ludwigshafen, Koblenz, and then Trier. But this is, as insiders report, mainly for carnival reasons. In quantitative linguistic circles, however, it is a well-established fact that the true capital (not only of Rheinland-Pfalz, of course) is in fact Trier. Or at least, it should be – which can easily be proven.

2.3 Rheinland-Pfalz: city size distribution

Ordering all Rheinland-Pfalz cities (with more than 10 000 inhabitants) according to their size in decreasing order results in what is used to be called the city size distribution. For almost one hundred years, beginning with Felix Auerbach's (1913) attempts to find a 'law of population concentration', many scholars have tried to find an appropriate mathematical formula to model corresponding distributions. Most of these attempts have started from Zipfian ideas, and even more did not go beyond his initial assumptions, limiting themselves by some dogmatic interpretation and subsequently starting competitions in wondering why the model is not flexible enough to cover the empirically observed variations. In this respect, it is important to note that a recently discussed model originating from quantitative linguistics seems to turn out a useful alternative (cf. Grzybek 2011a,b). In its continuous form, this is the well-known Zipf-Alekseev function, or the Zipf-Alekseev distribution, after norming (cf. Wimmer and Altmann 1999), which is given as

$$y = \frac{C}{x^{-(a+b\cdot\ln x)}} \; . \qquad (1)$$

There is no need to go into details here, as far as mathematical derivation of (1) is concerned.[1] In any case, fitting this function[2] to the city size distribution of Rheinland-Pfalz, limited to those 33 cities with more than 10 000 inhabitants, results in a determination coefficient of $R^2 = 0.99$, with parameter values $C = f_1 = 197778$ (see above), $a = -0.27$ and $b = -0.15$. Figure 2 presents the result of fitting the continuous model according to (1) in graphical form.

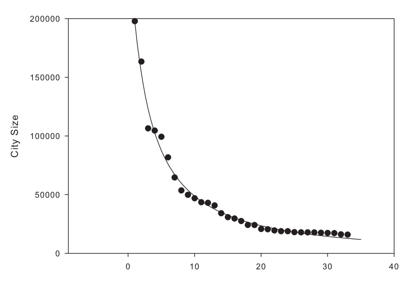

Figure 2: Fitting function (1) to the city size distribution of Rheinland-Pfalz (2010)

1. Suffice it to say that there have been different attempts to derive this function, e.g. by Hammerl (1990), Hřebíček (1997: 42f., 2000: 14f.), and Hřebíček and Altmann (1996: 56f.) who interpret the Zipf-Alekseev function with regard to the Menzerath law, ultimately considering the Menzerathian structure to be some kind of 'auto-iterator'.
2. Fitting the corresponding discrete distribution, in its right-truncated version, yields an excellent result, as well, with discrepancy coefficient $C = X^2/N = 0.0069$, for parameter values $a = -0.32$ and $b = -0.13$.

It seems obvious that we are concerned with a convincing model for the city size distribution of Rheinland-Pfalz. It goes without saying that such a model can originate only from theoretical modeling in quantitative linguistics as it is being promoted in Trier, so that, seen from a higher (theoretical level), it is, without a doubt, the city of Trier which should be seen the true capital of this state – *quod erat demonstrandum.*

With this in mind, we can now turn to a more detailed ('context-embedded') analysis of Trier as our textual object, and this analysis will provide us with further arguments in the direction outlined above.

3 From context to text: Trier

3.1 Trier: Historical Foundations

As has been already mentioned, the Federal State of Rheinland-Pfalz has been established in 1946. Reinhard Köhler was born five years later, in 1951. He was born in Dortmund, not in Trier. So from the very beginning, he had the advantage of knowing what it means to live in a large city and, in fact, what a large city is. As we know today, this circumstance would later predestine him to enter into a fruitful symbiotic relation with the city of Trier which, at that time, counted about 75 000 inhabitants, and which has been struggling more and more to transcend the 100 000 person limit.

Ultimately, it is just this struggle which is very much related to Reinhard's personal involvement and engagement in Trier. Of course neither Reinhard himself nor the Trier local authorities could at that time presage the fruits the symbiotic combination of his person, on the one hand, and the place of Trier, on the other, would carry decades later. But to fully understand this success story, it seems reasonable to cast a glance back to history for a moment and to recall some major historical cornerstones of the city of Trier, which today indeed is proud of accommodating more than 100 000 inhabitants, thus having the official right to call itself a "large city".

It is assured that the city of Trier is of Roman origin, founded in or before 16 C.E. as *Augusta Treverorum* (i.e., Augustus' city in the Treveri country), a name which in the second half of the 3rd century became common as *Treveris*. From the beginning of Trier's existence, there has been a pronounced need to accommodate talented linguists; the reason for this has to be seen in the fact that the name of the town, or rather its etymology, has remained basically unclear. Although it is a well established fact that the Treveri were a tribe of Gauls, speaking Gaulish as an early variety of Celtic, the etymology of Treveri (sometimes the variant Trev*i*ri can be found) has remained far from being clear: whereas Ancient Greek texts give Τρηούΐροι, transliterated as *Trēouïroi* (cf. Long 1854), variants such as Tre*b*eri or Τρίβηροι (Tr*i*bēroi) can be found in the works of Pliny and Ptolemy. The name has been etymologically interpreted as referring to a “flowing river”, or to “crossing the river”. Following Swiss linguist Rudolf Thurneysen (1857–1940) the name of Treveri might be interpreted as a Celtic *trē-uer-o*, with the element *trē-* < **trei* ‘through’, ‘across’ (cf. Latin *trans*) and *uer-o* ‘to cross a river’; following this interpretation, the name Treveri could thus mean “the ferrymen”, because these people’s job was to help other people to cross the Mosel river. The root *treuer-* can also be compared to the Old Irish *treóir* ‘guiding, passage through a ford’, ‘place to cross a river’. The word *uer- / uar-* can be related to an Indo-European word meaning ‘stream’, or ‘river’ (Sanskrit: *vār*, Old Norse *vari* ‘water’), that can be found in various river and place names. Given this background, some people claim that Trier is the oldest German town; to side with such an assumption is, of course, a terminological and political question of what one considers to be a ‘city’, a ‘town’, or a ’settlement’, etc.

It is just for this reason, with its important and far-reaching implications for cultural identity, that there has always been the need to have competent linguists in Trier. It goes without saying, that linguistics has very much changed and extended its shape over the centuries. Thus, whereas etymologists were needed in the early historical periods of Trier, later experts in dynamical and synergetic linguistics would be needed, due to the changed situation. In any case, at the beginning of the fourth century, Trier was the Roman Emperor’s residence, and with ca. 80 000 inhabitants it was the largest town north of the Alps.

3.2 Trier: Population development

In the course of time, the population continually decreased; only in the 18th century, the number of inhabitants started to rise again, reaching up to 8 820 in 1801. Up to this year, data were usually based on estimation; later data from results of a population census or from statistical recordings are more exact, but still not fully reliable because they were based on different principles of data collection. Due to industrialization, the population growth accelerated in the 19th century: in the year 1900, more than 43 000 people were living in Trier, the number of inhabitants rising to twice as much in 1939, with more than 88 000 inhabitants. During World War II, Trier lost ca. one third of its inhabitants, and the population again went down to ca. 57 000, in 1945. Only by way of taking administrative measures, i.e. by incorporating smaller communities into the larger municipality of Trier, the population increased by more than 20 000 by the end of the 1960s, being larger than 100 000 in the 1970s for the first time, and rising to 105 260 at the end of 2010.

Table 2 shows the development of the population from 1831 until 2010, thus covering a time span of ca. 180 years.

Table 2: Demographic development of the city of Trier (1831–2010)

Year	(transf.)	Inhabitants	Year	(transf.)	Inhabitants
1831	1.00	14723	1961	43.67	87141
1840	3.33	15717	1970	46.67	103724
1855	8.33	20172	1975	48.33	100338
1858	9.33	20060	1980	50.00	95536
1871	13.67	21442	1985	51.67	93472
1900	23.33	43506	1990	53.33	97835
1910	26.67	49112	1995	55.00	99428
1925	31.67	57341	2000	56.67	99410
1933	34.33	76692	2005	58.33	99843
1939	36.33	88150	2010	60.00	105260
1950	40.00	75526			

The development over the whole period is illustrated in Figure 3; the black dots mark the observed data, as given in Table 2. For technical reasons, the year data (as the independent variable) have been linearly transformed to a smaller scale (cf. the resulting values given in columns 2 and 5 in Table 2).

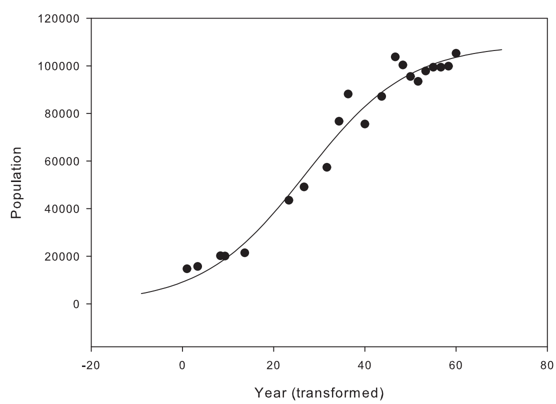

Figure 3: Demographic development of Trier (1831–2010)

In additional to the empirical data, Figure 3 contains an *S*-curve as the graphical representation of a non-linear regression function, modeling the overall tendency. This curve results from a theoretical model which is well-known in quantitative linguistics, namely the Piotrovskij-Altmann law, which is given by (2).

$$y = \frac{C}{\left(1 + a \cdot e^{-bx}\right)} . \tag{2}$$

With iteratively determined parameter values $C = 109\,187$, $a = 10.93$ and $b = 0.09$, the determination coefficient of $R^2 = 0.97$ shows the fit to be very good. Setting $C = x_{max} = 105\,260$ (and thus reducing the number of parameters to be explained by one), the result differs only minimally: in this case, with $a = 12.03$ and $b = 0.10$, the R^2 value remains basically unchanged. As a result, we may thus conclude that the city of Trier has always been a city predestined not only for linguists in general, but for quantitative and synergetic linguists in particular.

However, despite the fact that the overall trend clearly follows a well-known rule from diachronic linguistics, things are not as self-evident as they seem to be: as a careful inspection of Figure 3 shows, the population development over the last decades has not been as smooth as it has been over a period of almost 100 years before (cf. also Table 2. It seems obvious that these "irregularities" are not specific of Trier, but should be seen in a broader context; in fact, one may see here a parallel to the general trend of population development, as has been outlined above (cf. Table 1 and Figure 1).

Yet, with regard to the city of Trier, it is just Reinhard Köhler who has been involved in these processes much more than is generally known. In fact, one can say that without his decisive interventions, the overall development might not have resulted in such a smooth course to end in a model as the one to be seen in Figure 3, and it is the objective of the following sections to prove this in detail.

3.3 From Pillenknick to *RAM*

This is related, of course, to the fact that Reinhard Köhler was invited to come to Trier University, in the late 1980s – an invitation which would finally end with his being appointed ordinary professor there. The reason for his invitation undoubtedly was his expertise in quantitative linguistics. Yet, since that time, rumors never ceased that it was, first and foremost, local demographs, rather than linguists, who paved the way for him to come. Since the background of these historical circumstances is somewhat obscure, at least from today's perspective, it seems worthwhile to once again look at the situation from a retrospective point of view, in order to adequately appreciate Reinhard's role and merits in this affair.

It cannot be denied, of course, that the specific situation at Trier was not independent of the general development (be that the situation in the Federal State of Rheinland-Pfalz or in Germany as a whole). Yet, heuristically concentrating here on the specific situation in Trier, one cannot but state that before Reinhard Köhler was invited to come, there had been a dramatic change in the previous population development: due to the drop of birth rate, primarily related to the contraceptive pill (a phenomenon known as the 'Pillenknick' in German), there had been a clear decrease in population, starting with the 1970s. As a consequence, the population of Trier, which at that time had just surpassed the magic (and for both financial and reputation reasons extremely important) limit of 100 000 inhabitants, rapidly decreased.

Table 3 shows the development, concentrating on the cornerstones of the period from 1871 until the mid 1980s, again linearly transformed to a smaller time scale.

Table 3: Demographic development of the city of Trier (1871–1985)

Year (transf.)	Inhabitants
1	21442
6	43506
8	49112
11	57341
12.60	76692
13.80	88150
16	75526
18.20	87141
20	103724
21	100338
22	95536
23	93472

Demographs at that time, who tried to theoretically model this population development, had convincingly argued in favor of the assumption that the observed demographic tendency of the Trier population did not follow a logistic regression models, but rather a Lorentzian function like (3),

$$y = \frac{C}{1 + \left(\frac{x-a}{b}\right)^2}. \tag{3}$$

According to these analyses, prognoses predicting a negative increase (i.e., decrease) for future development seemed to be inevitable. Indeed, with parameter values $C = 97\,996$, $a = 20.53$, and $b = 12.37$, fitting (3) seemed to be very convincing, yielding a determination coefficient of $R^2 = 0.94$. Figure 4 shows the corresponding results.

Sharp-minded Reinhard Köhler, however, when faced with this situation and starting to analyze it, would correct the demographic approach by first introducing into the discussion three important maxims, all of them well-known from the field of quantitative linguistics:

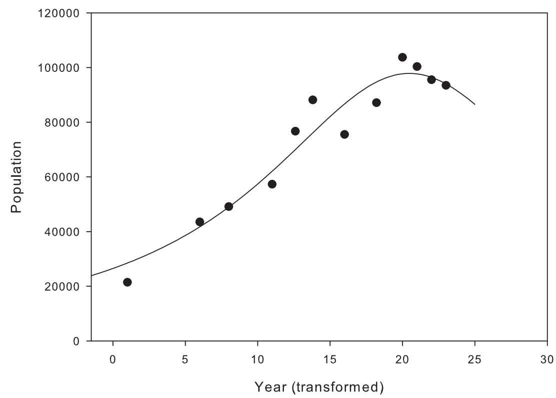

Figure 4: Demographic development of Trier (1871–1985)

Maxim 1
When theoretically modeling data, don't forget to look into the mirror – if you see a beard there[3]*, ask yourself if a razor like Occam's might improve your looks, or at least that of the data and the model thereof!*

Maxim 2
Attempt to interpret the parameters and parameter values of your model, at least as many of them as possible, however absurd your interpretation might seem to others!

Maxim 3
Think positive!

With regard to these maxims, Köhler made a number of suggestions, as early as during his first contacts with the local authorities at Trier University. One cannot but interpret these suggestions as an outcome of his congenial analytical thinking. "First", he said, "the Lorentz function should be used for analytical, not prognostic purposes; therefore, one should set $a = max_x = 23$ in (3), and not calculate any theoretical values beyond this."[4] Second, he drew attention to parameter b: "Look at parameter b", he said, "isn't it obvious to you, that one should use the lucky number '13' here, if one wants to arrive at a positive

3. As is well known, Reinhard Köhler has been wearing a beard since his early years(cf. the photo of his, reproduced on the cover of this book).
4. For some moments, Köhler had thought to interpret the value of 23 in a different way: having spent half of his life in Dortmund (see above), it would have been reasonable to calculate a as the Dortmund zip code (still 46 at that time) divided by 2, thus arriving at 23. But Köhler was clever enough not to present this interpretation, in order not to provoke the Trier authorities' local pride.

interpretation of the situation? Why don't you set $b = 13$?" And finally he was convinced: "If you follow my advice, you are not only likely to obtain a good fit, you will also get a clear hint at how many inhabitants Trier really should have, given its historical long-term development."

In fact, with $a = 23$, $b = 13$, the result of iterative fitting yields $C = 104\,067$, with a determination coefficient of $R^2 = 0.90$ which is, of course, slightly worse than the one obtained above, but which includes the desired positive effect. Figure 5 presents the results in graphical form; this model has later been termed Köhler's 'Realistic Analytical Model' (*RAM*).[5]

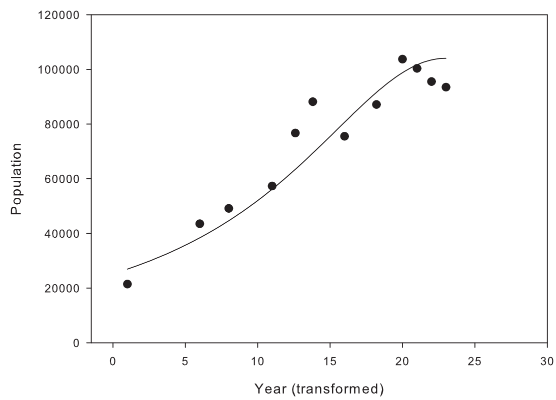

Figure 5: Demographic development of Trier (1871–1985)

The local authorities were of course not only deeply impressed by Köhler's suggestions, but also most enthusiastic about the C value obtained, which clearly indicated an "appropriate" population size of $> 100\,000$.

Köhler, however, despite the sizzling atmosphere around him, kept as cool as a cucumber (i.e., as he is known in quantitative linguistics enviroments). As a well-trained Saussurian follower, he next suggested to concentrate on a synchronous analysis of the situation, before attempting diachronical or even prognostic modeling. Specifically, he recommended to focus on the most recent developments – that is, the situation from the 1970 peak until 1986. This synchronic shift of the focus would result in Köhler's "Positive Prognostic Model" (*PPM*).

5. This name would later even serve as the basis for a publishing house specialized in quantitative linguistics.

3.4 From *RAM* to *PPM*

Table 4 represents the detailed data of Trier's population development from 1970 through 1986, with the transformed time scale in the second column.

Table 4: Demographic development of the city of Trier (1970–1986)

Year	(transf.)	Inhabitants	Year	(transf.)	Inhabitants
1970	1	103724	1979	10	95736
1971	2	103387	1980	11	95536
1972	3	103048	1981	12	95180
1973	4	102228	1982	13	94683
1974	5	101145	1983	14	95070
1975	6	100338	1984	15	94190
1976	7	99107	1985	16	93472
1977	8	97822	1986	17	93076
1978	9	96787			

Reinhard's Köhler sharp and experienced eye found the trend not to follow the Lorentzian curve, as previously assumed by demographers, but instead a well-known extension of the Menzerath-Altmann Law. This extension was, in fact, only very much later derived by Altmann and Wimmer from their "Unified Theory of Some Linguistic Laws" (Wimmer and Altmann 2005, 2006); but Köhler has always been a visionary.

In its simplified continuous form, the differential equation at the basis of the most frequently and extensively used continuous models in linguistics may be written as

$$\frac{dy}{y} = \left(a_0 + \frac{a_1}{x} + \frac{a_2}{x^2} + \frac{a_3}{x^3} + \ldots\right) dx\,, \tag{4}$$

the solution of which is

$$y = C\,e^{a_0 x}\,x^{a_1}\,e^{-a_2/x - a_3/(2x^2) - a_4/(3x^3) - \ldots}\,, \tag{5}$$

with C here symbolizing the integration constant. With $a_i \neq 0$ for $i = 0, 1, 2$, and $a_j = 0$ for $j = 3, 4, 5, \ldots$, equation (6) is obtained from (5):

$$y = C\,e^{-a_0 x}\,x^{a_1}\,e^{-a_2/x}\,. \tag{6}$$

Indeed, applying equation (6) to the Trier demographic data represented in Table 4, results in an excellent fit of $R^2 = 0.98$, with parameter values $a_0 = 0.0004$, $a_1 = -0.0737$, $a_2 = 0.1158$, and $C = 116278$. All three parameters (a_0, a_1, a_2) thus turn out to be negative; Figure 6 shows the results of the analysis in graphical form.

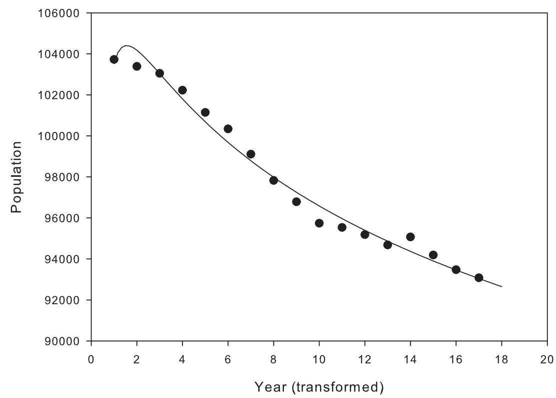

Figure 6: Demographic development of Trier (1970-1986)

It can easily be seen from Figure 6 that the initial increase is in line with what could be expected from the analyses in the previous step. When Reinhard Köhler presented this approach, and the results obtained from it, to the Trier demographs, the latter were as frustrated as stunned. It was only the value of parameter C, with $C = 116\,278$, which gave the local authorities[6] some hope, since they interpreted it to offer a hint at what the "appropriate" population size actually should be.

Yet, Reinhard Köhler at that time was wise enough not to present any future-oriented prognosis as to the increase of population growth. Only when it came to the job interview, in 1989, he was asked, among other things, for his opinion as to future demographic developments, and for possible suggestions how to improve the situation. Questions of this kind are quite common for academic application processes like this, but Reinhard Köhler was clever enough to see the true reason behind this seemingly innocent question. He immediately understood that, now that his brilliance as an analyst has been settled, he was asked not only for his analytical expertise, but also for his prognostic help, in order to improve the situation of a town, the shining star of which was in danger of sinking forever, once having been the prime city of the Northern Roman Empire...

During this job interview, Reinhard Köhler gave a convincing explanation, referring to his replacement of the Lorentzian approach with the general synergetic one. According to him, what had gone wrong in the past, was a typical phenomenon also known as 'overfitting'; or, to put it in simple words: the use

6. Not only linguists and demographs followed this presentation, but also politicians and even a State Secretary who, at those days, just by chance had no other political duties than local Carnival activities.

of too many parameters – this is fully in line with Köhler's Maxim no. 2, see above. Quite logically, the local authorities were almost enthusiastic, when he explained to them that a model with one parameter less would do as well to explain what he now uncovered to be an historical error; in fact, setting $b = 0$, one obtains

$$y = Cx^{a_1} e^{-a_2/x} , \tag{7}$$

and with parameter values $a_1 = -0.075$, $a_2 = 0.11$ and $C = 116\,145$, the fit turned out to be equally good ($R^2 = 0.98$) as for the more complex equation (6).

In the next step, Reinhard Köhler made another impressively clever move, thereby providing the necessary steps for the desired positive prognosis and, by way of this, a solution as to the future development and improvement. Suggesting that it was not parameter a_0 that should be equated with 0, but instead parameter a_2, he arrived at the well-known third special case of the Menzerath-Altmann law as described by Altmann (1981), which even the local experts had heard of through the grapevine before:

$$y = Ce^{-a_0 x} x^{a_1} . \tag{8}$$

Irrespective of the fact that this model – with $R^2 = 0.95$ for parameter values $C = 105\,240$, $a_0 = 0.0050$ and $a_1 = -0.0138$ – results in a slightly worse (though still very good) fit, this turned out to be a necessary pre-condition to arrive at a prognosis model, which would finally result in predicting a positive increase of population.

In the following considerations, Köhler was very much ahead of his time. Thanks to his theoretical way of thinking and arguing, he could convince his audience by pointing out but one crucial factor, namely, that parameter interpretation had simply gone wrong in past: as can be seen, both parameters, a_0 and a_1, turned out to be negative in the last analysis, based on equation (8). In order to improve the demographic situation, Köhler once again referred to his Maxim no. 3 "Think positive!" Integrating this slogan into the framework of synergetic linguistics, Köhler referred to the Zipfian concept of antagonistic forces, which simultaneously exert (differently weighted) influence on a given system and, by way of this, create and guarantee this system's dynamic balance. With regard to the demographic situation of the city of Trier, Köhler sagaciously identified these forces: the first being the general birth rate, the second being related to the contraceptive pill, with its residual risk. It is obvious that these two factors are not independent of each other. In fact, the relation between them turns out to be trivial, at closer sight: the higher the residual risk, the higher the birth rate.

As a consequence, exponents a_0 and a_1 should not only be interrelated, but also, one of them should be positive – in full accordance with Köhler's Maxim no. 3 – in the case of a positive population development. With this in mind, what first seemed to be an open and insoluble question, soon turned out to be what is called a "mown meadow" in Austrian pluricentral version of German, that is: a problem easily solved.

Analyzing the situation even more deeply, Köhler noticed that in the mid 1980s, the birth rate all over Germany was about ca. 11 per 1 000 (i.e., 11‰, or 0.0011, in relative terms). Given this observation, he first suggested to raise this rate up to the lucky number of 13 (resulting in a relative rate of 0.0013). This suggestion seems plausible: first, it had worked before, and second, the magic number '13' would easily meet the interpretational needs of people concerned with the pill's residual risk in a negative way. As a consequence, Köhler suggested to set $a_0 = -0.0013$ in equation (8), to arrive at a positive prognostic model. With regard to the second factor, he then suggested to fix the residual risk of birth control pill failure at 0.02 (a realistic estimation rate in those years), thus setting $a_1 = 0.02$. Finally, with regard to C, he took the starting value from 1989, thus setting $C = 96\,721$.

Figure 7 shows the trend as prognosticated by Köhler in 1989, on the basis of equation (8), and with the parameter values described above. The model – which Köhler termed the Positive Prognostic Model (*PPM*) – covers a period of 15 years (i.e., until 2003), as was desired at that time. Figure 7 also contains the real data for this period, as we know them today (cf. Table 5 below), which were of course not known to Köhler or anyone else, when he suggested this model.

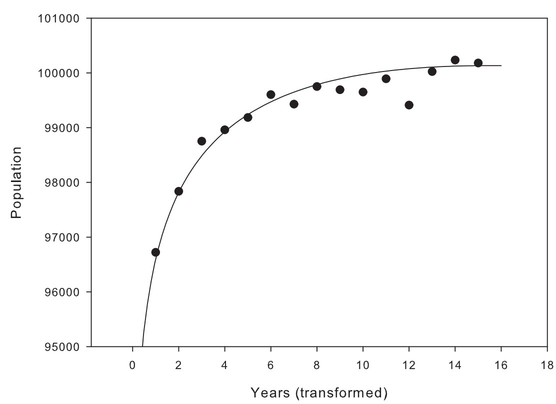

Figure 7: Demographic development of Trier (1989-2003)

Due to this lack of information, Köhler – who had always been a visionary clairvoyant, but no precog – could not, at that time, prove his theory to be correct, by way of statistical testing. Yet we know today – although this became to be known only much later – that Köhler's introduction of the *PPM* into the academic and political sphere was the true reason why he was then offered the chair for linguistic data processing and quantitative linguistics at Trier University. Today, with the available data represented in Table 5, we know of course that with a fitting result of $R^2 = 0.94$, Köhler's model was almost perfect.

Table 5: Demographic development of the city of Trier (1989–2003)

Year	(transf.)	Inhabitants	Year	(transf.)	Inhabitants
1989	1	96721	1997	9	99691
1990	2	97835	1998	10	99650
1991	3	98752	1999	11	99891
1992	4	98958	2000	12	99410
1993	5	99183	2001	13	100024
1994	6	99602	2002	14	100234
1995	7	99428	2003	15	100180
1996	8	99750			

3.5 From *PPM* to *EVA-KKM*

Summarizing thus far, we can say that thanks to Reinhard Köhler's

a. (re-)interpretations and suggestions in terms of his Realistic Analysis Model (*RAM*) for the tremendous decrease of the population in Trier, which had started in the 1970s and continued to the mid 1980s, and
b. development of the Positive Prognosis Model (PPM), which had paved the way for the renewed increase after 1989,

the overall demographic tendency following the Piotrovskij-Altmann law (as was shown at the beginning of this text) could be guaranteed. As can be seen from a detailed historical analysis as the one above, it sometimes takes a good craftsman, eventually with some expertise in synergetic theory, to turn the right screw in the right direction so that the water flows where one expects and wants it to flow. Without such seemingly small corrective measures, historical courses are likely to be thrown off the track; and when such measures are taken, they are likely to be overseen and insufficiently appreciated, since generally valid regularities and laws tend to override them in their long-term perception as mass phenomena.

In the case of the historical events treated with above, this has not only led to the fact that Reinhard Köhler's crucial impact on the historical development has been almost forgotten; first and foremost, his role (a) in interpreting the decreasing tendency, and (b) in re-establishing an increasing tendency have usually be seen (if they have been seen) independent of each other. In fact, however, they represent but two phases of one process. This can clearly be seen from a synopsis of Figures 6 and 7: subsequent to the post 1970 decline, there is a clear increase of population starting in the mid 1980s. Including Reinhard Köhler himself, scientists have always seen or analyzed both tendencies separately; but only taken together, they represent the basis of what may – complementary to the well-known term Pillenknick – adequately be termed the "Köhler knick", to include the increasing anti-tendency as compared to the Pillenknick decrease before.

In the remaining remarks to follow, an attempt shall be made to theoretically describe both tendencies in a common (joint) theoretical model. A standard procedure in establishing such a model for both tendencies would be to determine two specific domains and to regard the overall trend as a composition of the two (in this case opposite) components. Yet, this would not do justice to Reinhard Köhler's merits, and it would not be a happy ending of a story with Reinhard Köhler as major protagonist. As we know, he has never been what in German is called a "man of half affairs", i.e. someone who would be content with 50% solutions, or even with 100% results composed of two 50% partial solutions…

Thus, in order to integrate both tendencies outlined (including the most recent data up to our days[7]), a model shall be presented here by the name of "Köhler Knick Model" (*KKM*). This model results from what is usaually known as Extreme Value Analysis (*EVA*), a branch of statistics which deals with extreme deviations from a given probability distribution. *EVA* is widely used in many disciplines, including engineering, finances, earth sciences, traffic prediction, etc., but it has hardly ever found its way to synergetic modeling. The most common Extreme Value Model is

$$y = a \cdot \exp\left[-\exp\left(-\frac{x-b}{c}\right) - \frac{x-b}{c} + 1\right], \tag{9}$$

but for the specific purposes here, it is reasonable to use its intercept and tailed version

$$y = m - a \cdot \exp\left[\frac{-x+b+c-cd\exp\left(-\frac{x+c\ln d-b}{c}\right)}{cd}\right]. \tag{10}$$

At first sight, this model may seem to be rather complex. But since Köhler has always sided with Mario Bunge in assuming that no one ever said that science must be easy, we will be glad to follow this maxim, too: if reality is complex, a model of it need not be simple… Paying due attention to Köhler's Maxim no. 1, this is less about the complexity of the model (and the numbers

7. The data for the remaining years are: 2004 – 100 163, 2005 – 99 843, 2006 – 103 518, 2007 – 103 888, 2008 – 104 640, 2009 – 104 587, 2010 – 105 260.

of parameters involved), than about the matter of interpretation according to Maxim no. 2 and, eventually, Maxim no. 3. With these maxims in mind, let us attempt a parameter interpretation à la Köhler of the complex five-parameter model (10) above.

1. Setting the additive constant as $\hat{m} = f_1 = 103\,724$, four parameters remain (a,b,c,d).
2. With regard to the first of the four remaining parameters (a), it is reasonable to interpret it as the difference between maximum and minimum of the observed values, i.e. $103\,724 - 93\,076 = 10\,648$; the result must be diminished, however, by the "lucky percentage" of 13.1313% of that value, thus arriving at $\hat{a} = 9\,250$, with three parameters remaining (b,c,d).
3. As to parameter b, there is no problem in finding out that this is the "lucky number" itself, thus resulting in $\hat{b} = 13.13$, with two parameters remaining (c,d).
4. The estimation and interpretation of parameter c is, in fact, the most difficult part of the task: last but not least, this is related to the predilection of static approaches in linguistics (be they oriented towards synchrony or diachrony), and the prevailing neglect of dynamic aspects; the due integration of the temporal dimension, however, makes it clear that a year has 365 days, which results in a parameter estimation of $\hat{c} = 3.65$, and one parameter (d) remaining to be interpreted.
5. As to the remaining parameter d, it is but plain sailing to see that there is only one way to interpret it, namely $\hat{d} = 4$, since this is the fourth parameter to be interpreted.

The result of fitting equation (10) with the parameter values explained above – i.e., with $\hat{m} = f_1 = 103\,724$, $\hat{a} = 9\,250$, $\hat{b} = 13.13$, $\hat{c} = 3.65$, and $\hat{d} = 4$ – can be seen in Figure 8.

Figure 8 for the first time ever shows the famous Köhler knick in detail: the black dots correspond to the empirically observed data, the curve results from data fitting (10) with the parameter values reported above. In fact, as a result one obtains a determination coefficient of $R^2 = 0.91$, indicating an excellent fit, including a satisfying and convincing interpretation.[8]

8. The deviations to be observed over the very last years could not be foreseen; they are due to the fact that the local political authorities in Trier recently changed the way to determine population size in such a way that they could be on the safe side with their counts. Basically, they "artificially" increased population size by taking into account not only main domiciles, but also secondary residence tax payers so that, for example, every person with a camping caravan would additionally count as inhabitant. It goes without saying that such spontaneoius political influences cannot easily be introduced into long-term oriented theoretical models.

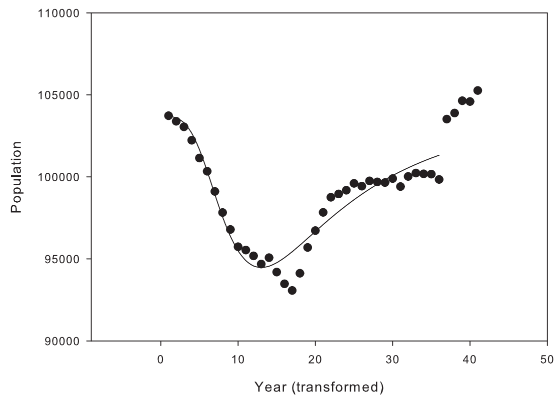

Figure 8: Demographic development of Trier (1970-2010)

4 Conclusion

It should have become clear from the above discussion how fruitful the symbiotic co-existence of Reinhard Köhler and the city of Trier has turned out to be over the years. It is difficult, of course, to say what have would have become of Trier without his being there. What we can say is that, without him, Trier would never have become the capital of quantitative linguistics. Also, it is very likely that the demographic tendencies would have gone a very different way, had they not been accompanied by and re-organized along linguistic laws, rules, and regularities brought into the discussion by Reinhard Köhler. After all, it was his theoretical contributions which, in the end, have guaranteed that Trier has been able to successfully overcome the negative population increase of the Pillenknick, and that, thanks to the Köhler knick, it has become a prospering town, even larger than in the ancient and glorious Roman times…

References

Auerbach, Felix
1913 Das Gesetz der Bevölkerungskonzentration. In: *Petermanns Mitteilungen*, 59; 74–76.

Grzybek, Peter
2011a Semiotics of the City · Semiotics of Cities. The Case of Estonia. [Presentation at the Third Lotman Days «Urban Semiotics. City as a Cultural-Historical Phenomenon». Tallinn, June 3-6, 2011]

2011b Sibiriens Städte gestern und heute: Ein Fall für Zipf? [Presentation at the Bilaterial Symposium «Siberien · Russland · Europa». Graz, September 19-20, 2011]

Hammerl, Rolf
1990 Untersuchungen zur Verteilung der Wortarten im Text. In: Hřebíček, Luděk (ed.), Glottometrika 11. Bochum: Brockmeyer; 142-156.
Hřebíček, Luděk
1997 *Lectures on Text Theory.* Prague: Oriental Institute.
Hřebíček, Luděk
2000 *Variation in Sequences.* Prague: Oriental Institute.
Hřebíček, Luděk; Altmann, Gabriel
1996 The Levels of Order in Language. In: Schmidt, Peter (ed.), *Glottometrika 15. Issues in General Linguistic Theory and The Theory of Word Length.* 38-61.
Long, George
1854 Treveri. In: Smith, William (ed.), *Dictionary of Greek and Roman geography, illustrated by numerous engravings on wood.* London: Walton and Maberly.
Wimmer, Gejza; Altmann, Gabriel
1999 *Thesaurus of univariate discrete probability distributions.* Essen: Stamm.
2005 Unified derivation of some linguistic laws. In: Köhler, Reinhard; Altmann, Gabriel; Piotrowski, Rajmund G. (Eds.), *Quantitative Linguistik · Quantitative Linguistics. Ein internationales Handbuch · An International Handbook.* Berlin, New York: de Gruyter; 791-807.
2006 Towards a Unified Derivation of Some Linguistic Laws. In: Grzybek, Peter (ed.), *Contributions to the Science of Text and Language: Word Length Studies and Related Issues.* Dordrecht, NL: Springer; 329-337.

On the dependency of word length on text length. Empirical results from Russian and Bulgarian parallel texts

Emmerich Kelih

1 Introduction

This paper tackles two basic problems of quantitative linguistics: firstly the "word length" and secondly the text length in terms of type and token numbers. It has to be shown that these two basic properties of a text are directly related. The interrelation between word length and text length can be captured by an appropriate mathematical model; hence a law-like status of the interrelation between word length and text length can be stated. Up to now, no special attention has been paid to this particular kind of self regulation (increase of the word length with increasing text length) of the text structure, which can be explained by a general control of the information flow on the lexical level. After a theoretical explanation of the interrelation of word and text length, some empirical results are presented on the basis of Russian and Bulgarian parallel texts (*The Master and Margarita*).

2 Word length research and text length

Recent word length research is dedicated mainly to problems of theoretical modelling (cf. Wimmer et al. (1994) and Best (2001) with an extensive bibliography) and to the general dependency of word length on the analysed text types (cf. Antić et al. 2006; Kelih et al. 2005). Moreover, word length plays a crucial role in Menzerath's law – cf. Altmann (1980), Altmann and Schwibbe (1989) – and is a central property of synergetic linguistics – cf. Köhler (1986), or Weber (2005) who gives a good overview.

Within word length studies, especially in works on modelling word length frequencies, different aspects have been discussed: the problem of an appropriate word definition, the question of relevant units of measurement (phoneme, syllable, morpheme, etc.), the impact of text types on word length, the sampling problem, etc. – for details see Grotjahn and Altmann (1993). In other words, there are many boundary conditions which have to be taken into consideration in word length studies.

According to our knowledge, the question of a "minimal", "maximal" or "optimal" text length for word length studies has never been discussed in detail,

Sven Naumann
Peter Grzybek
Relja Vulanović
Gabriel Altmann
(eds.)

Synergetic Linguistics

Text and Language as Dynamic Systems

To Reinhard Köhler on the Occasion
of his 60th Birthday

prae
sens

Bibliografische Information der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über http://dnb.d-nb.de abrufbar.

ISBN: 978-3-7069-0700-2

Gedruckt mit Unterstützung
der Karl-Franzens-Universität Graz

sowie der Kulturabteilung der Stadt Wien,
Wissenschafts- und Forschungsförderung