

On forgetting the historical past

THIJS POLLMANN

Utrecht University, Utrecht, The Netherlands

In this paper, empirical data are presented related to memory and time perception. The data are the frequencies with which specific calendar years are cited in newspaper texts. When plotted, the curves produced by the time series of these frequencies turn out to be independent of the languages and cultures in which the texts have been written as well as the year and the year density of the text corpus. The frequency of a specific year is inversely proportional to the distance from that year to the year in which the texts were written. It is argued that these curves are forgetting curves. It is suggested that the curves might be explained in terms of the "cognitive distance" between past and present. An argument is presented based on the curve representing the frequencies with which future years are cited in newspaper texts.

The aim of this paper is to present and analyze empirical data related to memory and time perception. A difference from other investigations in this field is that the number of data to be presented is very large. Moreover, the data relate to rather large time spans between present and remembered past, measured in decades and centuries rather than in minutes, hours, days, or weeks. The data are not the outcome of experiments, but they have been collected from large language corpora.

The data provide some empirical evidence that corroborates the intuition that the further the past recedes, the less often we make an appeal to it, and that this decline does not come to an end when we go beyond the beginning of the past we have experienced ourselves. Apparently, the content of our memory for historical matters decays with the passage of time. The data might be arranged as "forgetting curves." This process of decay, however, is a function of real time and apparently not related to the moments in which we experienced these things of the past through education or private reading. Therefore, the data ask for answers to questions about "forgetting."

Unless stated otherwise, here and in the rest of this paper I do not refer to historical knowledge as the knowledge one can find in documents of the past or the publications of historians, but I refer to historical knowledge as a mental property. As far as I know, little is known about the temporal-cognitive structure of internalized historical knowledge, especially the knowledge of a time that we did not experience ourselves. Undoubtedly, his-

torical knowledge in the indicated sense is governed by a temporal organization. We know that Luther's Reformation and the "discovery" of America by Columbus took place in the same period of history. We know that "perestroika" and "glasnost" preceded the fall of the Berlin Wall. Sometimes we remember the time of these events by deriving this information from the way we experienced them or learned about them. But that is not necessary. Sometimes we retrieve aspects of time by considering other historical events of which we do know the year (see Linton, 1975). It is easy to recognize a historical practice in such procedures. If someone asks me when Robespierre died, the answer will be reconstructed out of other historical data: "It was after the French Revolution, which was in 1789, because the revolution brought about the reign of terror, of which Robespierre himself was one of the last victims. It was also after 1792, the year in which the French king was executed. And the Reign of Terror ended before Napoleon began his campaigns; wasn't that in 1795 or 1796? Well then, Robespierre must have died in 1793 or 1794." Such a description of the reconstruction of the time in which an event took place is in accordance with the results of psychological research concerning time memory (Friedman, 1993; Neisser et al., 1996). The contextual information that we use in such a derivation cannot be based even partially on personal experiences—but, as we have seen, neither is this necessary for the derivations of contemporary events.

This means, in the first place, that we must assume that human cognition is able to organize knowledge of the past temporally, regardless of whether the subject has experienced this past or not. But if this were true, what about forgetting? Ought we to expect that knowledge about a past we did not experience ourselves is governed by the normal laws of forgetting, and, therefore, is explained by existing psychological theories of forgetting? When memory is able to determine the time of a historical event in the same way as it proceeds in determining the time of a contemporary event, should we not expect to find forgetting curves for this kind of knowledge?

The author wants to thank Leen Dorsman, Ed Elbers, Lex Heerma van Voss, Carel Jansen, Ed Jonker, Malcolm de Mowbray, Judith Pollmann, Elwin Savelsbergh, and the reviewers of this journal for their comments and advice during the various stages of the research for this paper. Correspondence should be addressed to T. Pollmann, Utrecht University Research Institute for Language and Speech (UiLOTS), Drift 8, NL-3512 BS, Utrecht, The Netherlands (e-mail: thijs.pollmann@let.ruu.nl).

— Accepted by previous editor, Geoffrey R. Loftus

In the next section, I shall present empirical evidence that bears upon these questions.

DATA AND FIRST ANALYSIS

The evidence I shall present is based on the use of calendar years in newspaper texts. Years are used as markers for all kinds of historical events (1066, 1517, 1813, 1917, etc.) but also as points of an imaginary time line. In this paper, I deal with years in this last sense. Years indicated as points on a time line make it possible to measure time spans—in particular, those larger than a few years. As such, years are mental supports to embed our understanding of “earlier” and “later.” They are a means of giving the past some structure.

Of course, years are not historical knowledge. But I take it that years, as used in normal language, are signs that the text makes reference to something in the past (or, the future, for that matter; I restrict myself at first to years of the past, but I will return to future years below).

Years are rather well suited as data in investigations about certain fundamental aspects of processing knowledge of the past. This is due to a series of properties. Years are discrete entities, and they are easy to detect. They are frequent in normal written language, easy to process statistically, and easy to process statistically with “time” as an independent variable. The frequencies with which years are cited in texts form a time series.

In the following, I assume that the concept of historical time is represented or related in one way or another to a line of years. This line of years is a cognitive construct to measure linear time with year as the measuring unit. The line of years is a cognitive scheme in the sense of Michon (1986): It is a “rule-based, generative procedure that guides perception, thinking and action” (p. 55). The linear structure of time is not inconsistent with Friedman’s (1990, p. 55) argument that linear time has only a superficial role in the perception of time and that “cyclic time appears a far deeper current in human experience and would have exerted a stronger force in shaping any natural temporal code” (pp. 43–44). Indubitably, the recurrent sequence of hours of the day, days of the week, and months and seasons of the year is an important aspect of our time perception. But it is also a fact that our knowledge of historical time is pinned to our knowledge of years. If we possess a time indication of a historical event, it is normally the year in which it happened, not the day or the month.

In the framework of research on the use of numbers in natural language, I built up collections of years from corpora of newspaper texts. In the analyses below, I will use the collections based on the following text corpora: the CD-ROM of the 1994 volume of the *International Herald Tribune (IHT)*, an American newspaper for the international market (appr. 18,000,000 words); the 1994 CD-ROM of the *Frankfurter Allgemeine Zeitung (FAZ)*, the famous German newspaper (appr. 26,200,000 words); the “27mlncorpus” of the *Instituut voor Nederlandse Lexicologie* (Institute for Dutch Lexicology) based on the volumes for 1994 and the first part of 1995 of the *NRC/Handelsblad (NRC)*, a leading Dutch newspaper (appr. 27,000,000 words).

The data have been gathered electronically. In the first step, concordances were made of the citations of all numbers between 1200 and 2100 in these text corpora. Subsequently, from the numbers that came up in a relatively low frequency (< 50), those indicating years were selected “by hand”—that is, on the basis of the “sense” of the number/year in the given context. (For an account of the collections in more detail I refer to Pollmann, 1997.) Expressions such as “the 1860s” or abbreviations such as “‘14-‘18” have been omitted.

For practical reasons, I have limited my counts of these materials to years of the period 1200–2100 (without the years 1994 and 1995, the years from which the data came).

Shorter time series (i.e., the frequencies of the years 1930–1993) have been collected from the 1994 CD-ROM editions of the French newspaper *Le Monde*, *The Times* of London, and the Spanish *El Mundo*. In the last case, only the first 6 months of 1994 have been used.

The newspaper corpora I used are rather large. They contain between 18,000,000 (*IHT*) and 41,000,000 words (*Times*). The size of the *El Mundo* corpus is unknown; it might be estimated as large as 10,000,000 words.

Years are cited in large numbers in newspaper texts. In the 1994 volume of the *FAZ*, I counted more than 70,000 year citations related to the past. In the “27mlncorpus” of the Dutch *NRC*, I found more than 63,000 years cited. The 1994 volume of the *IHT* cites 22,000 years. Table 1 gives some further details.

“Year density” values were computed for the year citations for the period 1993–1200. In Table 1, year density values are expressed as a ratio of the number of citations for years of that period to the total number of word

Table 1
Number of Years and Year Density in the Three Language Corpora

Text Corpus	Number of Words in the Corpus	Number of Years			Percentage of Past	Year Density
		Past (1993–1200)	Future (1996–2100)	Total		
<i>FAZ</i> 1994	26,200,000	73,423	7,148	80,571	91.2	1:356
<i>NRC</i> 1994+	27,000,000	63,471	5,969	69,440	91.4	1:425
<i>IHT</i> 1994	18,000,000	22,656	4,177	26,833	84.4	1:794

Note—“Percentage of Past” refers to the percentage of the total number of years that were past years.

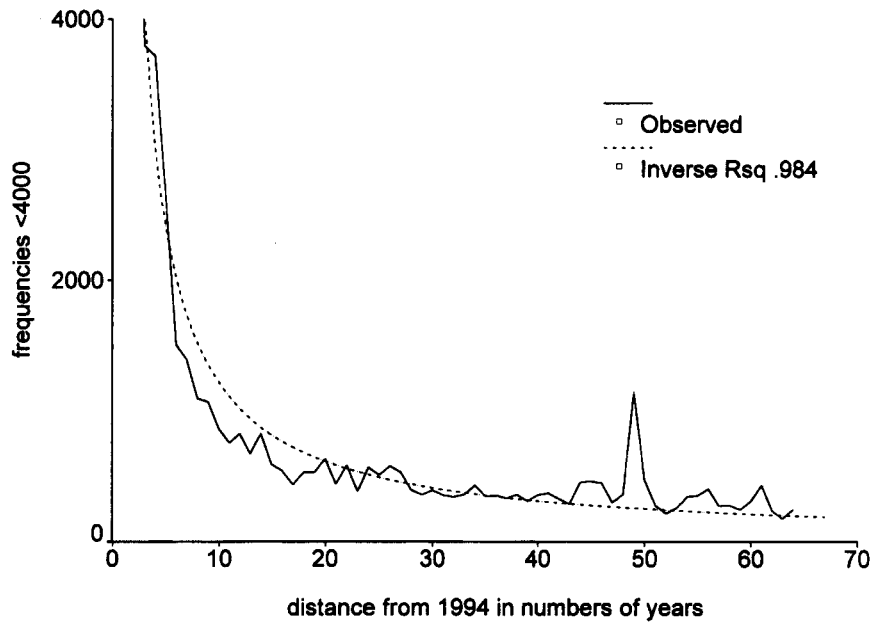


Figure 1. *Frankfurter Allgemeine Zeitung* 1994: Time series 1993–1930.

forms in the corpora. The year densities of the three corpora diverge substantially.

Of all the year citations that appear in the three collections, the share relating to years in the past lies between 91.4% and 84.4%. Given the diverging year densities, these values can be said to be rather close to each other.

Against this background, I will investigate the distribution of the frequencies with which the years are cited over the time line. As might be expected, the frequencies diminish with the passage of time; however, up until 1200, we find enough citations to say that most years are represented at least once. Looking backward, the first year we miss in *IHT* is 1807, 1717 in the Dutch *NRC*, and 1557 in the *FAZ*.

If the frequencies with which each year is cited in an arbitrary part of the time line are computed and then these frequencies against time are plotted, curves such as those in Figure 1 are obtained.

In Figures 1–6, the 1990–1930 part of the time line is depicted, on the basis of the six data collections for this period.

The curves have one large peak and several small peaks. Of the irregularities in the lines, many can be explained either on the basis of the cultural importance of the events of these years or on the basis of their roundness. Some years or periods define part of the identity of a culture; thus, they are cited more often in the newspaper. This, of course, explains the large peaks of 1944 and 1945 but also the smaller ones such as those of 1933 and 1938 in the *FAZ*. “Round years” are, in general, more frequent than other years, just as in normal language “round numbers” are more frequent than other numbers (see Jansen & Pollmann, 1995; Sigurd, 1988, p. 247). A similar “roundness effect” in estimations of elapsed time was found by Huttenlocher,

Hedges, and Bradburn (1990). For a further analysis and a defense of this explanation of the irregular features of the curves, the reader is referred to Pollmann (1997).

The curves also represent an underlying regular form. Apparently, the most recent year is the most frequent. (The years 1991, 1992, 1993, whose values rise still higher, have been omitted from the graph for reasons of presentation.) The frequencies with which the years are cited diminish with the passage of time. We find this regular drop not only in the period 1990–1930 but, in principle, in each part of the time line. Initially, the descent is fast, but the speed of the drop diminishes as the distance from the past increases.

Fit functions have been computed for these data using SPSS 6.1. For the data presented in Figure 1, the inverse function appears to be the best fit, at least among the 11 models that the program is able to compute. A survey of the competing fit functions is presented in Table 2.

Table 2
Survey of Models That SPSS 6.1 Is Able to Compute

Function	$y =$
Linear	$c_0 + c_1 * t$
Inverse	$c_0 + c_1 / t$
Power	$c_0 * t^{c_1}$
Quadratic	$c_0 + c_1 * t + c_2 * t^2$
Cubic	$c_0 + c_1 * t + c_2 * t^2 + c_3 * t^3$
Compound	$c_0 + c_1^t$
Logistic	$1 / (1 / u + c_0 + c_1^t)$
Logarithmic	$c_0 + c_1 * \ln(t)$
Growth	$e^{(c_0 + c_1 * t)}$
Exponential	$c_0 * e^{c_1 * t}$
S	$e^{(c_0 + c_1 / t)}$

Note— $c_0, c_1, c_2,$ and c_3 are constants (C_0 might be left out of the equation); t is an independent variable; u is an upperbound variable. In our case, t is time measured in number of years.

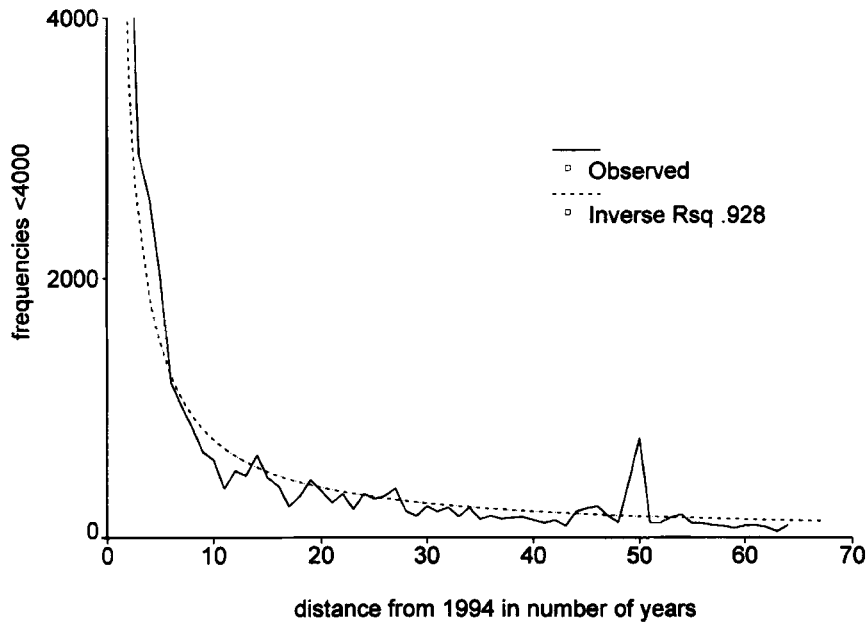


Figure 2. *International Herald Tribune* 1994: Time series 1993–1930.

The inverse function is a power function of a certain type. The function has the following formula: $[y = C_0 + C_1/x]$, where y is the frequency, x is the distance in time, and C_0 and C_1 are constants. As a measure of the degree of fit, Rsq scores can be used. The maximum Rsq score is 1, if every frequency point in the data is one of the points of the fit curve. Actually, I found Rsq scores that lie between .915 (*Le Monde*) and .986 (*El Mundo*) for the curves of Figure 1.

To neutralize the effect of the small and sometimes larger irregularities in the curves (the peaks in the graphs),

the fit functions for the 10-year averages of the rough data in three of the corpora have been computed for longer periods. For the data of the period 1993–1200, the inverse models again appear to give the best fit. The Rsq scores lie between .963 (*NRC*) and .955 (*FAZ*) (Figures 7–9). These scores have to be considered as satisfactorily high.

This means that the frequency of a specific year is inversely proportional to the number of years that has passed between that year and the year in which the texts were written. This is true for each of the collections of years I have built up.

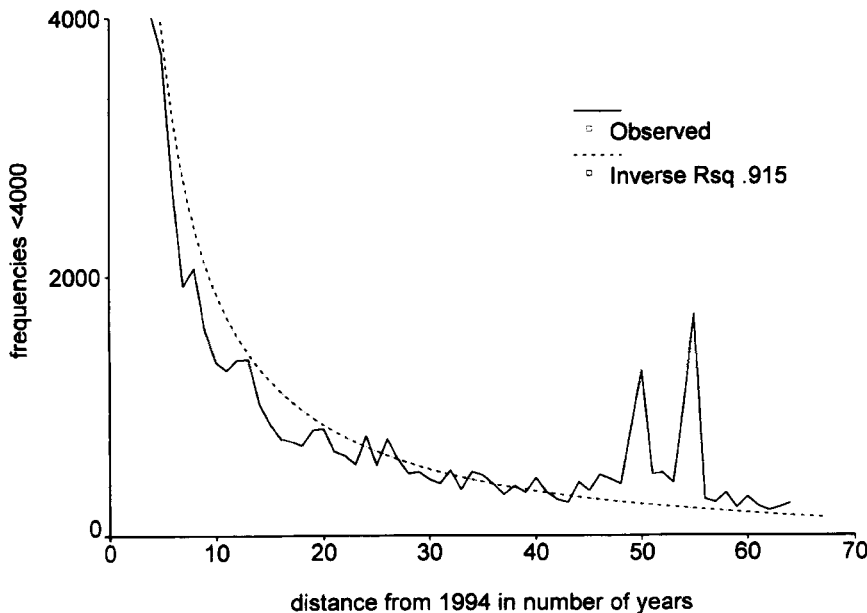


Figure 3. *Le Monde* 1994: Time series 1993–1930.

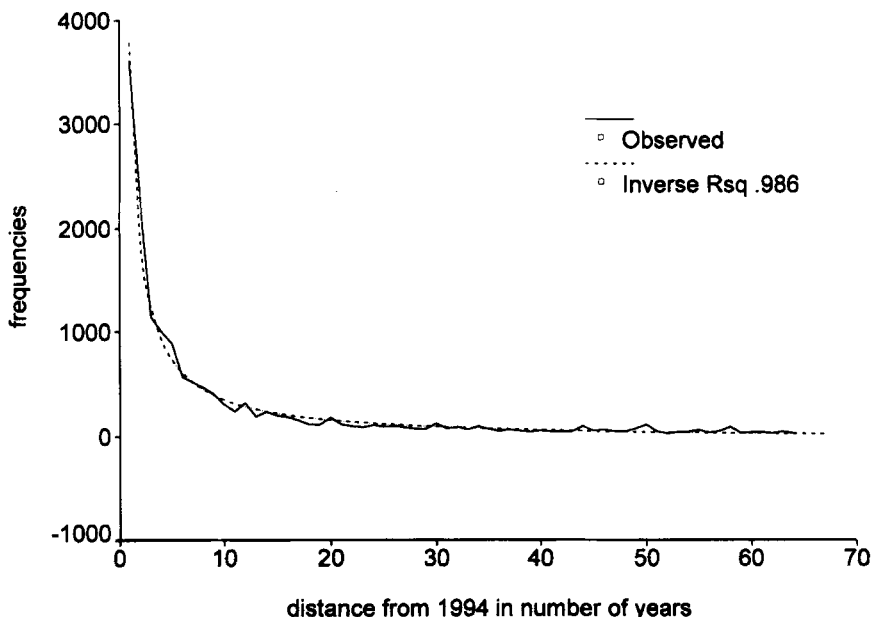


Figure 4. *El Mundo* 1994: Time series 1993–1930.

We might conclude that the overall shapes of the curves are independent of the languages and cultures in which the texts have been written, the year 1994, and the year density of the text corpus. Of course, it is an intriguing question how this remarkable constancy in the decay of attention paid to the past might be explained.

That memory decays is a trivial fact, and it has been investigated in thousands of studies. Normally, the course of this decay in time is mathematically presented as a forgetting curve. It will be clear, at first sight, that the

curves we derived from the data have a remarkable resemblance to these forgetting curves (see Anderson & Schooler, 1991). The question of whether the curves might be conceived as forgetting curves will be the subject of the next section.

FORGETTING THE HISTORICAL PAST

It seems to me that one can speak of two kinds of forgetting, which I shall call *Forgetting 1* and *Forgetting 2*.

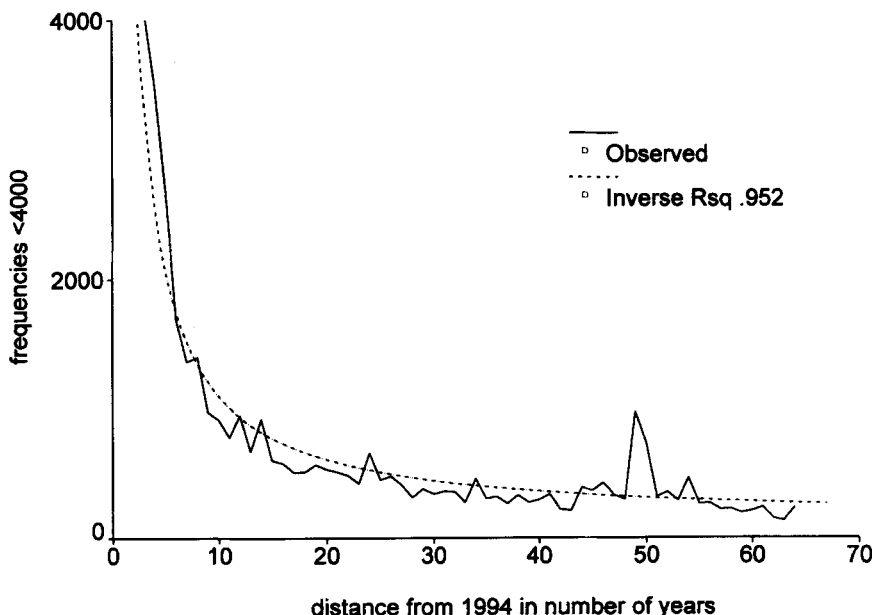


Figure 5. *NRC/Handelsblad* 1994+: Time series 1993–1930.

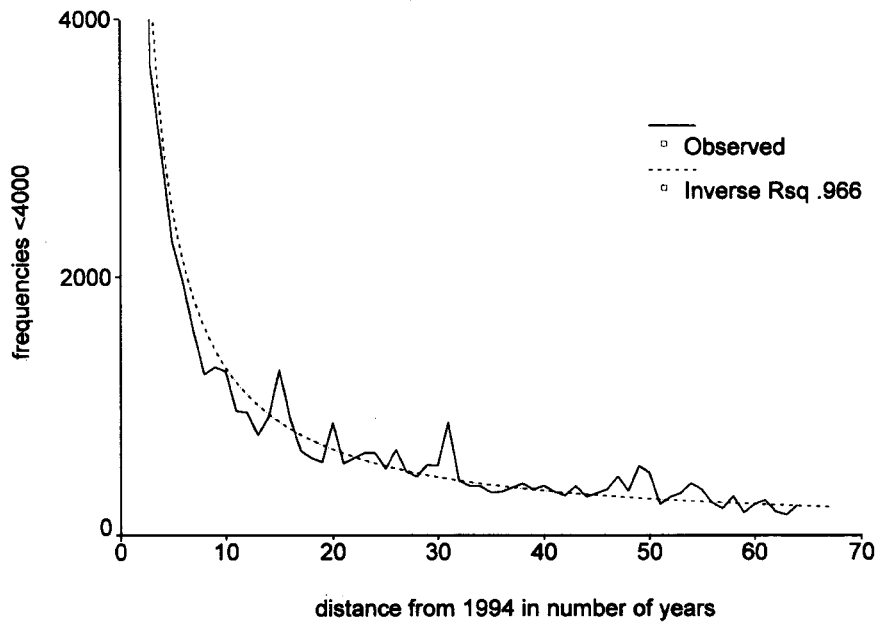


Figure 6. *The Times and Sunday Times 1994: Time series 1993–1930.*

Forgetting 1 is the diminishing of a memory content or the diminishing ease the increase of effort with which a part of the content of memory can be retrieved or correctly reproduced. Forgetting 2, on the other hand, is the decline in the amount of attention paid to a part of the content of memory.

Although both types of forgetting are seldomly explicitly kept apart, I have the impression that psychological science deals more particularly with Forgetting 1. This form of forgetting seems to be relevant in a frame-

work of investigations into the functioning of memory that follows (almost directly) a process of learning. Words, lists of nonsense words, trigrams, combinations of digits, sounds, visual configurations, and so on are to be learned in experimental situations; afterwards (in a time span that seldom exceeds a month), there is an investigation of what the subject remembers or is able to retrieve of these stimuli. Whether this is to be interpreted as a loss of memory content or a decreasing ease of retrieval is a matter for theoretical debate.

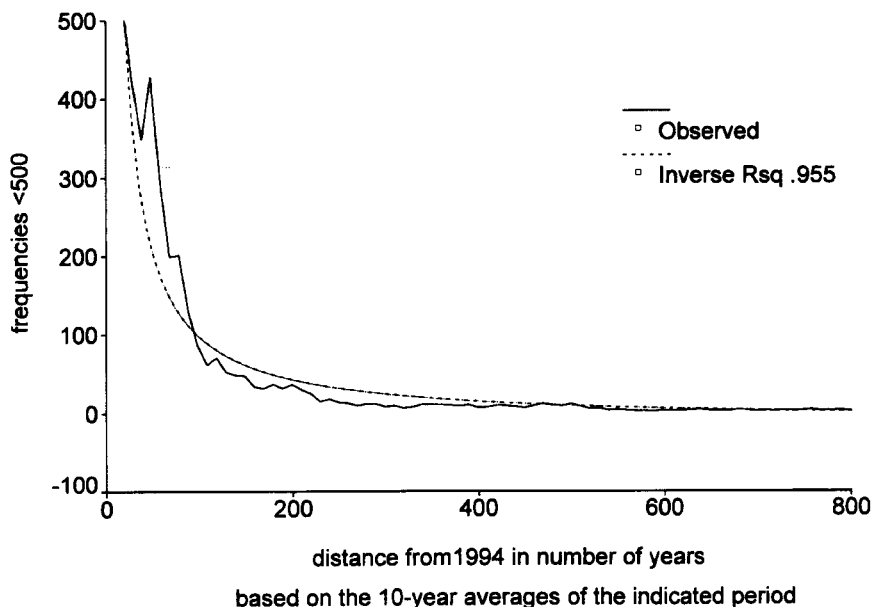


Figure 7. *Frankfurter Allgemeine Zeitung 1994: Time series 1993–1200.*

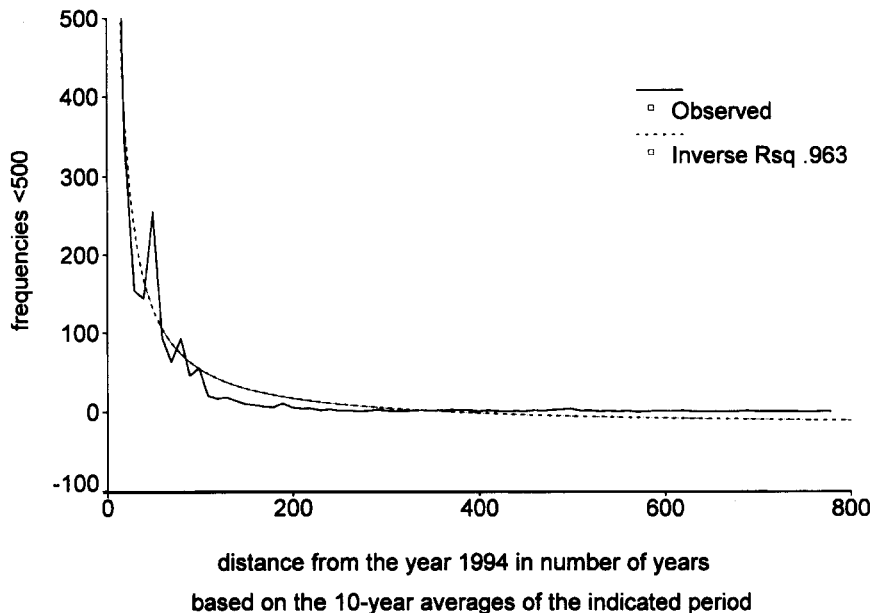


Figure 8. *International Herald Tribune* 1994: Time series 1993–1200.

Forgetting 2, on the other hand, refers to experiences that presuppose a less explicit learning. When we say that we have forgotten a deceased aunt, poverty, a craze, our sorrows, or a natural disaster, we mean that these things no longer claim our attention, whereas, previously, they had claimed our attention. The decline in attention paid to an aspect of the past over the passage of time is Forgetting 2 (see Hölscher, 1989).

As I have said, these two kinds of forgetting are not always carefully distinguished. Certainly, one finds forms

of memory research in which critics explicitly state that repetition of the things the subjects learned experimentally could not have been excluded and that, therefore, the resulting forgetting curve has to be judged unreliable (see, e.g., Kintsch, 1977, p. 55). In these cases, the critics are apparently worried that Forgetting 1 should not be mixed with Forgetting 2. However, there exists a respectable branch of memory research, usually called *autobiographical memory research*, in which forms of Forgetting 1 have to have been mixed with Forgetting 2.

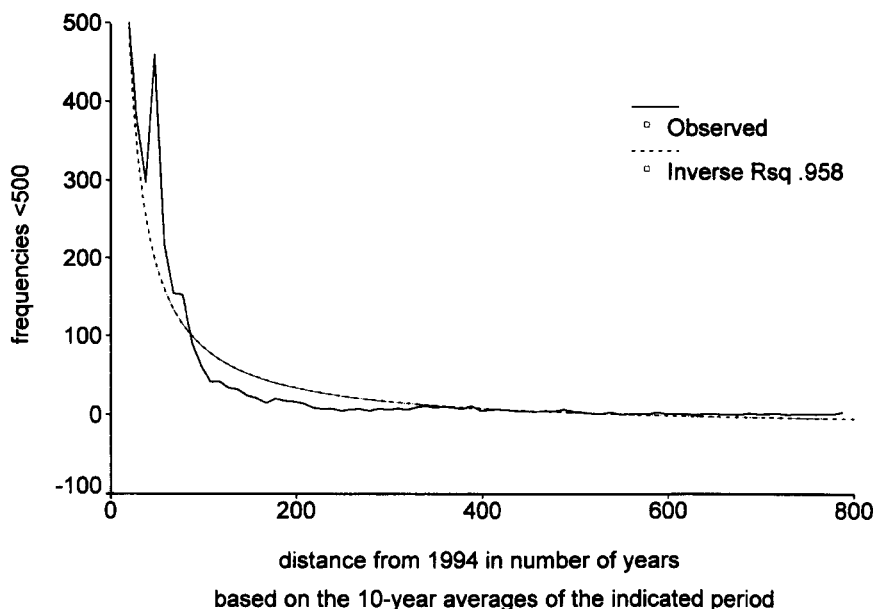


Figure 9. *NRC/Handelsblad* 1994+: Time series 1993–1200.

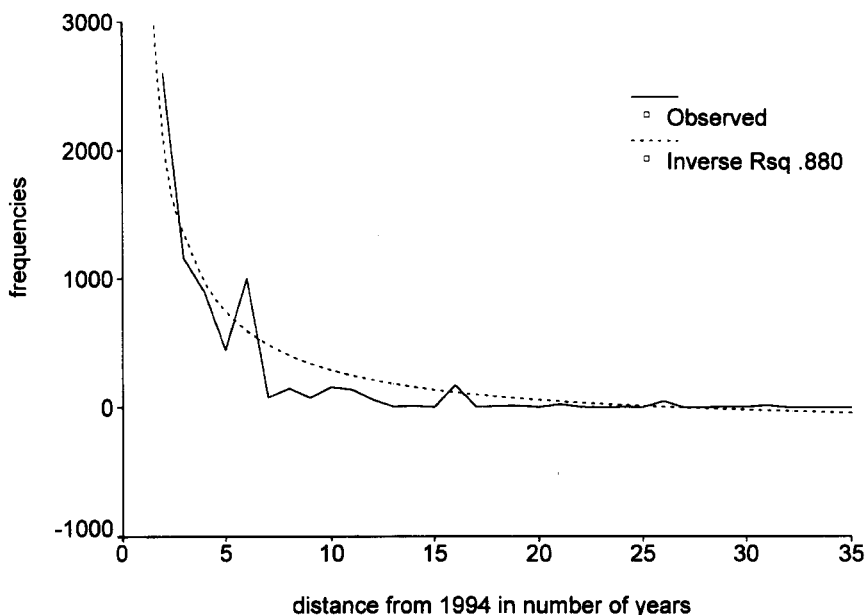


Figure 10. *Frankfurter Allgemeine Zeitung* 1994: Time series 1996–2030.

Wagenaar’s (1986) research is an example. The investigations of Friedman and Wilkins (1985) and Neisser et al. (1996) into the recollections that people had of an earthquake they experienced (directly or indirectly) seem to be another.

The journalist’s task of citing 1 or more years is similar to the task a subject executes in an “indirect memory” experiment (cf. Richardson-Klavehn & Björk, 1988, p. 478). It is quite natural to consider journalists who cite a particular year as subjects who have been urged to pay

attention to an aspect of the past. This urging leads to a mental activity that appeals to their memory or through their memory to an external source. In the latter case, memory will give a certain indication about the event’s location in time, which will be completed by external information.

One can conclude that the presented curves indicate a diminishing attention to phenomena of the past. If they represent forgetting, it is Forgetting 2. This conclusion seems to be in line with the findings of Anderson and

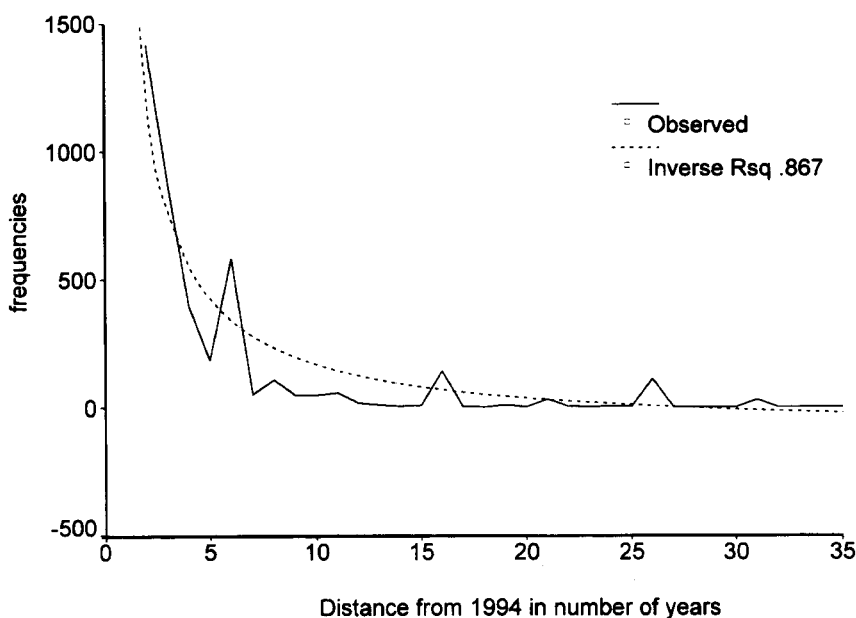


Figure 11. *International Herald Tribune* 1994: Time series 1996–2030.

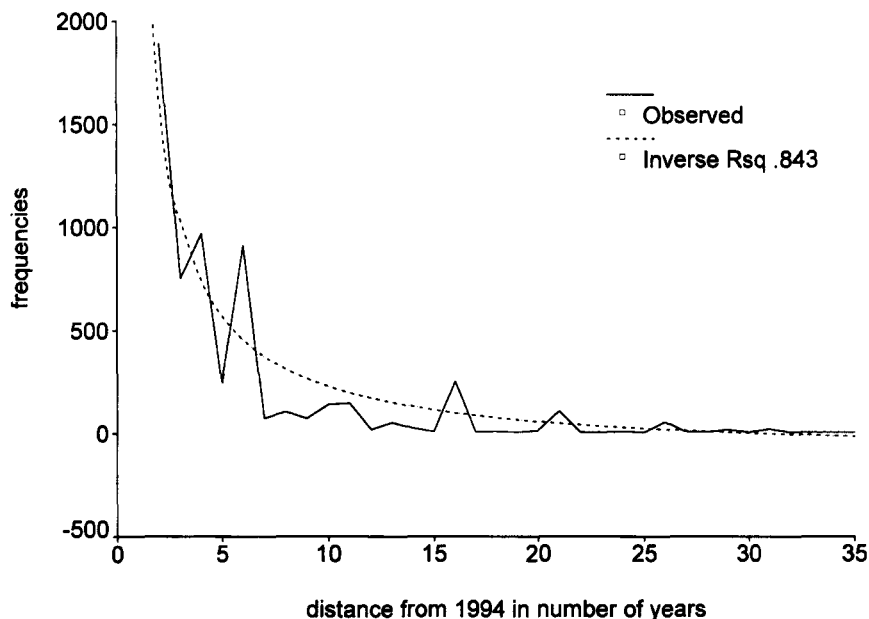


Figure 12. *NRC/Handelsblad* 1994+: Time series 1996–2030.

Schooler (1991), their basic idea being “that at any point in time, memories vary in how likely they are to be needed and the memory system tries to make available those memories that are most likely to be useful” (p. 400).

Explaining the data as Forgetting 2 phenomena is not in conflict with the fact that some years or groups of years depart from the underlying slope: the small or little peaks in the graphs. The regularity in the process of Forgetting 2 is compatible with the culturally defined “extra” attention to some years or periods in history.

How does one explain Forgetting 2 phenomena in a way that is compatible with the curves we found? The data do not seem to say anything about memory contents, about the way the knowledge of the past has been filed in memory, or about our capacity to recall things from the past. In which other direction should we look for an explanation that better covers the facts? I have nothing to offer but a few suggestions.

Apparently, the longer an event has passed, the more difficult it is to relate it to something in the present. I think that one should try to explain the data in terms of the “cognitive distance” between past and present. Not forgetting something in the sense of Forgetting 2 depends on its usefulness in the present—its “need-odds,” to use a term of Anderson and Schooler (1991). This usefulness, in turn, apparently diminishes as the distance between the past and the shifting present grows. In this explanation, Forgetting 2 is not a function of the human capacity to select, store, and retrieve information, but it depends on the amount of attention given to things that are not in the field of our direct attention.

If cognitive distance is decisive for the differences among the frequencies with which particular years are

cited in newspaper texts, then it should be expected that this is also the case for years in the future. One would expect the frequencies for future years to be inversely proportional to the distance between the present and future years. To put it differently, one would expect the distribution of the future years over the time line to have an inverse function as the best fit. This prediction is borne out by the facts.

It is striking, but also very understandable, that round years are widely represented in the future part of the data collections. Even more often than for the past, we rely on estimations about the time of future events. However, this does not alter the fact that we find an overall declining curve for the dates of the future. Figures 10–12 depict the diminishing frequencies of the years for the period 1996–2030 in three text corpora. (Data for the period 2030–2100 are not completely absent, but they appear in the round years only; the conclusions are not affected by omitting them from the graphs.)

The frequencies with which years are cited for this period decline in all collections. I again computed the best fit functions. For all investigated collections, the inverse functions were found to give the best fit. The *Rs_q* scores vary from .880 (*FAZ*) to .843 (*NRC*).

What connects the cognitive domains of the future and the past is their mental distance from the present. Generalizing about the past and the future, the “nonpresence,” we might speculate that the mental difficulty of finding or constructing a connection between a thing in the present and a thing not in the present increases with the distance of the past of future event to the present. The larger this distance, the more difficult it is to bridge the gap between something in the past or future to something

in the present, and, therefore, the less likely it is that the past or future will find a place in the actual universe of discourse.

REFERENCES

- ANDERSON, J. R., & SCHOOLER, L. J. (1991). Reflections of the environment in memory. *Psychological Science*, **2**, 396-408.
- FRIEDMAN, W. J. (1990). *About time: Inventing the fourth dimension*. Cambridge, MA: MIT Press.
- FRIEDMAN, W. J. (1993). Memory for the time of past events. *Psychological Bulletin*, **113**, 44-66.
- FRIEDMAN, W. J., & WILKINS, A. J. (1985). Scale effects in memory for the time of events. *Memory & Cognition*, **13**, 168-175.
- HÖLSCHER, J. (1989). Geschichte und Vergessen [History and forgetting]. *Historische Zeitschrift*, **249**, 1-17.
- HUTTENLOCHER, J., HEDGES, L. V., & BRADBURN, N. M. (1990). Reports of elapsed time: Bounding and rounding process in estimation. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **16**, 196-213.
- JANSEN, C. J. M., & POLLMANN, M. M. W. (1995). Ronde getallen [Round numbers]. *Tijdschrift voor nascholing en onderzoek van het reken-wiskun-deonderwijs*, **14**, 31-35.
- JANSEN, C. J. M., & POLLMANN, M. M. W. (1997). *On round numbers*. Manuscript submitted for publication.
- KINTSCH, W. (1977). *Memory and cognition*. New York: Wiley.
- LINTON, M. (1975). Memory for real-world events. In D. A. Norman & D. E. Rumelhart (Eds.), *Explorations in cognition* (pp. 376-404). San Francisco: Freeman.
- MICHON, J. A. (1986). J. T. Fraser's "Levels of temporality" as cognitive representations. In J. T. Fraser, N. Lawrence, & F. C. Haber (Eds.), *Time, science and society in China and the West: The study of time V* (pp. 51-66). Amherst, MA: University of Massachusetts Press.
- NEISSER, U., WINOGRAD, E., BERGMAN, E. T., SCHREIBER, C. A., PALMER, S. E., & WELDON, M. S. (1996). Remembering the earthquake: Direct experience vs. hearing the news. *Memory*, **4**, 337-357.
- POLLMANN, M. M. W. (1997). *Newspapers and the historical past: An empirical inquiry into the presence of the past in newspaper texts*. Manuscript submitted for publication.
- RICHARDSON-KLAVEHN, A., & BJÖRK, R. A. (1988). Measures of memory. *Annual Review of Psychology*, **39**, 475-543.
- SIGURD, B. (1988). Round numbers. *Language in Society*, **17**, 243-252.
- WAGENAAR, W. A. (1986). My memory: A study of autobiographical memory over six years. *Cognitive Psychology*, **18**, 225-252.

(Manuscript received September 5, 1996;
revision accepted for publication January 2, 1997.)