

1965 International Conference on
Computational Linguistics

MODELS OF LEXICAL DECAY

D. Kleinecke

TECHNICAL MILITARY PLANNING OPERATION
GENERAL ELECTRIC COMPANY
735 State Street (P. O. Drawer QQ)
SANTA BARBARA, CALIFORNIA 93102



Abstract

Lexical decay is the phenomenon underlying the dating techniques known as "glottochronology" and "lexicostatistics." Much of the controversial nature of work in this field is the result of extremely imprecise foundations and lack of attention to the underlying statistical and semantic models.

A satisfactory semantic model can be found in the concept of semantic atom. Notwithstanding a number of philosophical objections, the semantic atom is an operationally feasible support for a lexicon which is a semantic subset of all possible meanings and at the same time, exhausts the vocabulary of a language. Lexical decay is the process by which the lexical item covering an atom is replaced by another lexical item.

Exponential lexical preservation is, in this model, directly analogous to decay phenomena in nuclear physics. Consistency requires that the decay process involved in exponentially preserved vocabularies be a Poisson process. This shows how to form test vocabularies for dating and proves that presently used vocabularies are not correctly formed.

Dialectation studies show that historically diverging populations must be modelled by correlated Poisson processes. Definitive statistical treatment of these questions is not possible at this time, but much desirable research can be indicated.

Introduction

This paper is an attempt to establish the method of dating by lexical decay upon an adequate theoretical foundation. The method discussed is that invented by Swadesh (1) over a decade ago and usually known as glottochronology or lexicostatistics. In the intervening years it has been widely applied, but often to the accompaniment of much confusion and controversy. It seems that much of the confusion can be removed by a rigorous treatment of the phenomenological model and careful application of statistics. The controversy can be removed only by the completion of a sufficient number of supporting studies. Rigorous formulation permits us to pinpoint what studies are needed and what conclusions are being sought.

Granting (as not everyone seems willing to do) that the basic fact of "uniform" lexical decay occurs, the problem to be attacked is that of correctly formulating models for lexical decay and of correctly deriving statistical consequences from these models. In what follows, we will construct a set of models which seem to fit the needs of the method of dating by lexical decay. Our approach is strictly pragmatic, that is, we construct the model we need without concerning ourselves about its a priori reasonableness. Later we try to assemble some arguments which justify the model. In no sense is this an approach for first principles.

The analogy between lexical decay and the decay phenomena of nuclear physics has been often noted and dismissed. In the present paper, we insist that this analogy is much more than an analogy; it is, on the first level, an identity. The only alternative to this hypothesis seems to be a kind of mystic faith that the decay occurs but without palpable manipulable principles. The burden of the proof that the identity is false lies with the doubter and we will make no further demonstration of its validity.

Decay phenomena in nuclear physics are governed by relatively simple, well understood principles. To apply these results to lexical decay we first establish the concepts of a semantic atom and a set of independent semantic atoms. The observed fact of exponential decay of vocabulary then is accounted for by assuming that the lexical item covering an atom decays according to a Poisson process. Generally speaking, the converse of this is also true, and only a Poisson process would produce exponential decay. From these considerations, we can draw many conclusions about how to and how not to construct test vocabularies for dating purposes.

With this model in hand, we can draw conclusions of a statistical nature. For example, we can develop formulas for the proper method of dating the split between three or more languages and for good estimators in more complex situations.

We can construct an imprecise heuristic model for the dynamic semantics underlying the Poisson process. So long as the first order theory is adequate, this is much in the nature of a curiosity. It seems, however, that first order theory is not adequate. Actually, such a conclusion is really premature because the kind of verification studies needed have not been made. Assuming the pessimistic conclusion, we have to construct second (or higher) order theories to account for the inadequacies of first order theory. At the moment, we have no useful results in this direction--the problem merges into the problem of dialectation. Probably the most important service we can render is to indicate exactly what kind of detailed studies are needed.

Semantic Atoms

It is very easy to raise objections of a philosophical nature to the concept of a semantic atom. In this paper we will simply ignore

these objections and define semantic atom in an operational way. There are also operational difficulties, but these seem to be surmountable.

A semantic atom is a completely defined unit concept sufficiently specified to remove all ambiguity. For example, in an anthropological context, we might have "sun, as pointed at by a male anthropologist at high noon in the middle of summer on an average day among a group of young men with plenty to drink". The kind of subtleties needed to complete the definition reminds one of Korzybskian General Semantics, but the intention is not the same. We seek to remove ambiguity but we must have a non-unique concept--one that is always present.

Certainly there has been little use of semantic atoms anywhere in the past. Those interested in semantics for its own self will reject them as useless or meaningless; lexographers deal in more generalized concepts. It would be hard to argue that they have general utility, but they are precisely what is needed for studying lexical decay.

Each semantic atom, in any speech at any time, is assumed to be covered by some lexical item. That is, there is some word whose meaning includes that of the atom. Thus, vocabularies can be formed over any set of semantic atoms by listing the covering lexical item for each atom. The kind of decay being studied is that where the covering lexical item is replaced by another item. The replaced word only rarely immediately disappears from the language as a whole, but it has disappeared from the semantic atom.

An independent set of semantic atoms is a set of atoms all of which differ among themselves enough to make the decay at any atom completely independent of that at any other atom. Thus, only one from sets of words, like numerals or pronouns, with habitual interrelations

can appear in the set. Independent sets are useful because in them, problems of inter-atom correlations need not be considered.

Before passing on, we should say a few words as to the practical use of semantic atoms. There does not seem to be any doubt that the collectors of vocabularies want to work with semantic atoms-- even if their results are completely unsuccessful. In an entry "dog = hund" they would like to say that there is a semantic atom and its cover in English is "dog", in German, "hund". The pitfalls of this sort of thing are well-known. Some care in defining atoms might make it feasible if we require not complete identity of the English and German semantics, but rather the existence of some concrete concept where both the English and German words are appropriate. Clearly this much weaker requirement will be easier to satisfy, so we adopt it.

We conclude that, with adequate precautions, semantic atoms can be operationally feasible even if true rigor is impossible. In the case of little-known languages, there is much more chance for error. We should encourage collectors of vocabularies to improve the precision of their definitions so that the atom in question can be identified.

Decay Process

We assume that lexical decay, for a set of independent semantic atoms, is a Poisson process. That is, it satisfies three conditions:

1. Each atom decays independently of all the other atoms.
2. Each atom decays independently of its history of earlier decay.
3. There is a constant λ such that for each atom the probability of one decay in a short time interval Δt is $\lambda \Delta t$, and the probability of more than one decay is negligible.

It is rather easy to deduce that for longer time intervals t , the probability of not decaying is $\exp(-\lambda t)$, and if there are N atoms, the expected number of undecayed atoms after time t is $N \exp(-\lambda t)$.

This formula is the usual formula for lexical decay. It should be pointed out that it was tested, statistically, in the first publication by Swadesh, and it failed to pass. The difficulty is probably due to the word list used which is not an independent set of atoms.

If we examine the assumptions made so far, we see that any list of semantic atoms can be used if they are: (1) independent; and (2) assured of existence throughout the time in question. There is no satisfactory a priori basis for assuming that some kinds of semantic atoms decay at different rates than other kinds, and it is doubtful if enough historical evidence can be collected to make such a conclusion statistically significant.

The question whether λ is a universal constant, a constant within any one language but possibly differing between languages, or a variable, is easier to discuss. So far, indications are that λ is about equal to $1/5000$ years. Now this means that over the span of most historic evidence, $\exp(-\lambda t)$ will be greater than about 0.60. There is a great deal of scatter to be expected in the results because $N \exp(-\lambda t)$ is an expectation, not an exact prediction.

There have been a number of studies of the exponent of exponential decay. All of them are too superficial to be conclusive (2). An adequate study in any one language would have to meet several criteria which make it into a major research effort. A set of independent semantic atoms must be selected--selected prior to detailed study--and no atoms, however difficult, dropped without complete explanations (3). Then the history of each atom must be traced through the historical record to locate the lexical item covering the atom at

each point in time. In reporting the study, all of this should be fully documented in detail. Each instance of decay can then be recognized and tallied. Statistical tests should be applied to see whether or not the model is satisfied and to estimate λ . For example, if there are 100 semantic items¹, there should be about one decay every 50 years uniformly spread through time. These things can be checked statistically. We hope that scholars will undertake definitive studies of this type for as many cases as possible (4).

Until the results of the kind of research just mentioned are available, the status of λ is unsure. We anticipate it will be recognized as a universal constant.

There remains the problem of making a Poisson process a reasonable assumption. In other words, we need to describe some sort of mechanism which makes words slip off semantic atoms independently of how long they have been covering the atom, and at a constant rate λ per unit time, at least over short time intervals. Incidentally, since λ is on the order of 1/5000 years, 50 years is a short time interval. Since the speakers of normal languages are not historians, the independence from history seems easy to accept.

The constant rate is harder to accept. First of all we have to account for an identical figure in populations, literate and illiterate, and between a handful of speakers and half a billion speakers. The decay effect must be independent of the number of speakers, hence it must be operative at the level of the single isolated speaker. This is satisfactory since, by and large, the amount of speech reaching an individual does not seem to have changed much throughout history and does not vary much between cultures at the present day.

But why does a speaker decide to change an occasional lexical item--about 1% in his lifetime--and maintain the rest. The only

hypothesis we have been able to construct is that all words are always under pressure--perhaps from several semantic "directions" at the same time. Most atoms resist change most of the time, but some set of accidents (all very real events at the sociological and psychological levels, but random accidents in our context) weakens a few, and the lexicon decays. In other words, there is a constant dynamic movement among secondary and incidental covers of the semantic atom which threaten the principal cover. Usually the threatening lexical items recede, but occasionally, in a random way, about once every five thousand years the principal cover is displaced and a lexical decay occurs.

The hypothetical mechanism advanced to explain lexical decay can be checked against history by case studies of semantic atoms. Each atom should show time periods when the principal word was nearly displaced. During these periods it is difficult to decide whether the old word or a new word is the principal cover. Usually the new word will pass away again, but sometimes it will displace the old word. A very tentative guess based on a casual examination of one hundred current English words suggests there are about four very heavily threatened words per hundred. Since we can expect about one word to be decaying at this moment, we conclude that about three out of four times the old word survives. All of this needs to be verified or disproven in detailed studies.

Decay Statistics

The statistical consequences of the model--the first order model described above--need to be explored. We cannot handle all possible situations, but the following examples should provide an adequate demonstration of technique so that any other problems which occur can be solved in the same manner.

First, let us consider N languages deviating independently from a common parent which is not known to us. The following discussion is a bit more cumbersome than some alternative approaches, but it generalizes more easily.

Let α be any set of the N languages and let $P(\alpha)$ be the probability that the given semantic atom is covered by the original lexical item in exactly the languages of set α . New covering words are assumed to be different in each of the innovating languages. $P(\alpha)$ is a function of time and satisfies the following differential equation:

$$\frac{d}{dt} P(\alpha) = - \sum_{i \in \alpha} \lambda P(\alpha) + \sum_{j \notin \alpha} \lambda P(\alpha \oplus j)$$

where i and j are languages, \in and \notin mean "belongs to" and "does not belong to" respectively, and \oplus is the union of α and the set containing only the language j .

Let $|\alpha|$ denote the number of languages in α . If $|\alpha| = N$, the equation is easy to solve:

$$P(\alpha) = \exp(-\lambda t)^{|\alpha|}, \quad |\alpha| = N.$$

If a few cases-- $|\alpha| = N - 1$, $|\alpha| = N - 2$, etc.--are solved, we are lead to hypothesize that

$$P(\alpha) = \exp(-\lambda t)^{|\alpha|} (1 - \exp(-\lambda t))^{N - |\alpha|}.$$

This can be proven by induction on $|\alpha|$ from $|\alpha| = N$ downward since $|\alpha \oplus j| = |\alpha| + 1$. Then

$$\frac{d}{dt} P(\alpha) = -|\alpha| \lambda P(\alpha) + \lambda (N - |\alpha|) \exp(-\lambda t)^{|\alpha|+1} (1 - \exp(-\lambda t))^{N-|\alpha|-1}$$

so that

$$\frac{d}{dt} \left(P(\alpha) \exp(\lambda t)^{|\alpha|} \right) = (N - |\alpha|) \lambda \exp(-\lambda t) (1 - \exp(-\lambda t))^{N-|\alpha|-1};$$

$$P(\alpha) \exp(\lambda t)^{|\alpha|} = (1 - \exp(-\lambda t))^{N-|\alpha|};$$

and the hypothesis is proven by induction.

Thus, $P(\alpha)$ depends only on the value of $|\alpha| = n$. We can recognize $P(n)$ for $n=2, 3, \dots, N$ but $P(0)$ and $P(1)$ cannot be distinguished so we combine these into P' which is obtained by

$$\begin{aligned} P' &= 1 - \frac{N(N-1)}{2} P(2) - \dots - \frac{N!}{n! N-n!} P(n) - \dots - P(N) \\ &= 1 - \left((1 - \exp(-\lambda t)) + \exp(-\lambda t) \right)^N - P(0) - NP(1); \end{aligned}$$

since there are $N! / n! (N-n)!$ sets with $|\alpha| = n$. Thus,

$$P' = (1 - \exp(-\lambda t))^{N-1} (1 + (N-1) \exp(-\lambda t)).$$

Now suppose that from K semantic atoms we observe that k_N atoms are covered by the same word in all languages, and k_{N-1} in all but one, and so on to k_2 , and there are k' atoms differently covered in all languages. The probability of this occurring is

$$P(N)^{k_n} P(N-1)^{k_{N-1}} \dots P(2)^{k_2} P^{k'}$$

$$= x^{(2k_2 + 3k_3 + \dots + Nk_N)} (1-x)^{NK - k' - (2k_2 + \dots + Nk_N)} (1 + (N-1)x)^{k'}$$

where $x = \exp(-\lambda t)$. A maximum likelihood estimate for x seems to be the best single value we can assign to x . This is obtained by setting the (logarithmic) derivative of probability to zero so that

$$0 = \frac{A - k'}{x} - \frac{NK - A}{1 - x} + \frac{(N-1)k'}{1 + (N-1)x}$$

where $A = k' + 2k_2 + 3k_3 + \dots + Nk_N$. Or

$$N(N-1)Kx^2 - ((N-1)A - NK + k')x - (A - k') = 0.$$

If $N=2$, $x^2 = k_2/K$, which is the well-known formula for the separation between two languages. For general N , x is the solution of the quadratic equation given above. Note that the answer depends on the statistic A which does not usually appear in discussions of lexical dating.

An even more general difference between this treatment and usual treatment by pairs is found in the use made of the number of all the languages containing a certain lexical item as the cover of a semantic atom. This kind of count is almost never made in the literature on dating problems.

Another case which constantly recurs in practice is that of three languages; 1, 2 and 3, say. The pair 1 and 2 are more closely

related than language 3 is to either 1 or 2 . Suppose t is time from the common ancestor of 1, 2 and 3 to 3, and t' the time from the common ancestor of 1 and 2 to 1 or 2 . Let $x = \exp(-\lambda t)$, $x' = \exp(-\lambda t')$ so that x/x' is the probability associated with the time from the common ancestor of 1, 2 and 3 to that of 1 and 2 .

We might observe any of five situations concerning the cover of a semantic atom. It may be the same in all (1, 2, 3); or in any pair (1, 2), (1, 3) or (2, 3), or different in each. The probability of each of these events is

$$x_{123} = x \frac{x}{x'} x^{12} = x^2 x' ,$$

$$x_{13} = x_{23} = x \frac{x}{x'} x' (1 - x') = x^2 (1 - x') ,$$

$$x_{12} = \left(1 - x + x \left(1 - \frac{x}{x'} \right) \right) x^{12} = x' (x' - x^2) ,$$

$$\begin{aligned} x' &= 1 - x^2 x' - 2x^2 (1 - x') - x' (x' - x^2) = 1 - x^{12} - 2x^2 + 2x^2 x' \\ &= (1 - x') (1 + x' - 2x^2) . \end{aligned}$$

Supposing k_{123} , k_{13} , k_{23} , k_{12} and k' of each of these is observed when K atoms are considered. The total probability is

$$x^{2(k_{13} + k_{23} + k_{123})} x' (k_{12} + k_{123}) (1 - x')^{k' + k_{13} + k_{23}} (x' - x^2)^{k_{12}} (1 + x' - 2x^2)^{k'}$$

Maximum likelihood estimates for x and x' are gotten from the equations obtained by setting the (logarithmic) derivatives by x and x' to zero separately.

$$0 = \frac{2(k_{12} + k_{13} + k_{123})}{x} - \frac{2k_{12}x}{x' - x^2} - \frac{4k'x}{1 + x' - 2x^2} ,$$

$$0 = \frac{k_{12} + k_{123}}{x'} - \frac{k' + k_{13} + k_{23}}{1 - x'} + \frac{k_{12}}{x' - x^2} + \frac{k'}{1 + x' - 2x^2} .$$

These equations are best solved numerically for given values of k_{12} , k_{13} , k_{23} and k' .

The methodology is straight-forward and there is no need to multiply examples. In every case we obtain new formulas based on maximum likelihood estimators. Another area in which these methods could also be utilized is in the construction of significance tests and confidence bands. With this basis, most of the machinery of modern statistics would be available for use.

Criticism of First Order Theory

As we explained in discussing semantic atoms, we feel there is no adequate observational data to which to apply these formulas for a conclusive test of their value. We have made a few experimental applications using the unsatisfactory data available in the literature.

Numerically, the time estimates we obtained, which we will not quote here, do not differ a great deal from those obtained by considering pairs alone. This is to be expected if the phenomena are at all consistent. The value in the formulas derived above lies in the fact that they correctly combine the data from several pairs.

The first-order method does have one very important difficulty which appears almost immediately if we try to treat more than three languages. This difficulty is in the family tree of the languages.

In the entire first-order development, we have implicitly used the concept of a tree. Languages go together as a "common ancestor"

until some point in time when they divide and become two separate languages. The tree is the first-order model of dialectation--it is known to be inadequate, at least in many situations. In spite of a century or so of studies, we simply do not understand how dialectation occurs. More study is greatly needed, especially in the construction of higher-order models, but the problem lies outside the scope of this paper.

The difficulty with the tree rises in decay studies because only splitting is compatible with our statistical model. We have no alternative to constructing a family tree if we wish to apply the method outlined above. However, it seems to be easy to find examples which do not allow a tree to be constructed. Consider four languages; A, B, C and D. Suppose one semantic atom has the same cover in A and B, and another different cover in C and D. And at the same time, some other atom has one cover in A and C, and a different cover in B and D. We cannot fit this data into any family tree.

A little more specifically in the Romance languages, we find that the same innovation with respect to Latin is shared by several or all the later languages. Some of this can be explained by the colloquial versus learned speech theory, but no family tree can be constructed to explain all the combinations of innovations. If we had an adequate explanation of the phenomena involved in these shared innovations, it is quite possible that we could assume Romance was the direct descendant of Imperial Latin without going back to Plautus or thereabouts, as seems to be required by the first order theory.

A tentative beginning in this direction can be made by a second-order theory based on the dynamic model of lexical influence.

Second-Order Lexical Decay

The imprecise model of semantic pressures we formed to explain lexical decay suggests the following second-order model.

For each semantic atom, we consider not only a covering lexical item as before, but also a potential covering item. The potential cover is the source of pressure against the cover. When the cover decays, it is replaced by the potential cover. Naturally we also assume that the potential cover decays and is replaced by a new potential cover. In the interest of simplicity and because we have no numerical data, we will assume both decays have the same constant λ .

First, let us consider a single language. The situation at an atom can be of four types: (I) both the original cover and potential cover remain; (II) the original cover remains, but the potential cover has decayed; (III) the original cover has decayed and the potential cover has replaced it; (IV) the cover is now neither the original nor the potential cover.

Let P_I and P_{II} be the probability of the first two situations. Then

$$\frac{d}{dt} P_I = -2\lambda P_I ,$$

$$\frac{\partial}{\partial t} P_{II} = -\lambda P_{II} + \lambda P_I ;$$

so that

$$P_I = \exp(-2\lambda t) ,$$

$$P_{II} = \exp(-\lambda t) (1 - \exp(-\lambda t)) .$$

The original cover remains in these two cases only so that the probability of it remaining is

$$P_I + P_{II} = \exp(-\lambda t)$$

which is exactly the same as in first-order theory.

When the second-order theory is applied to N languages, the results are quite complicated. We divide the languages into four sets $\alpha, \beta, \gamma, \delta$ depending on which situation holds in the language; in set α , situation I holds, and so on. Then we have the basic differential equation

$$\begin{aligned} \frac{\partial}{\partial t} P(\alpha, \beta, \gamma, \delta) = & - \sum_{i \in \alpha} 2\lambda P(\alpha, \beta, \gamma, \delta) \left(\sum_{j \in \beta} + \sum_{k \in \gamma} \right) \lambda P(\alpha, \beta, \gamma, \delta) \\ & + \sum_{j \in \beta} \lambda P(\alpha \oplus j, \beta \ominus j, \gamma, \delta) + \sum_{k \in \gamma} \lambda P(\alpha \oplus k, \beta, \gamma \ominus j, \delta) \\ & + \sum_{\ell \in \delta} \lambda P(\alpha, \beta \oplus \ell, \gamma, \delta \ominus \ell) + \sum_{\ell \in \delta} \lambda P(\alpha, \beta, \gamma \oplus \ell, \delta \ominus \ell) \end{aligned}$$

which has the solution

$$P(\alpha, \beta, \gamma, \delta) = [\exp(-\lambda t)]^{2|\alpha| + |\beta| + |\gamma|} [1 - \exp(-\lambda t)]^{|\beta| + |\gamma| + 2|\delta|}$$

We have no way of recognizing the condition of the potential cover, so sets α and β should be combined into a set η and

$$P(\eta, \gamma, \delta) = [\exp(-\lambda t)]^{|\eta| + |\gamma|} [1 - \exp(-\lambda t)]^{|\gamma| + 2|\delta|} .$$

Before we can actually apply the maximum likelihood technique to languages without known ancestors, we have to make some further combinations because sets with $|\eta| = 1$ can not be distinguished from those with $|\eta| = 0$ or those with $|\gamma| = 1$ from those with $|\gamma| = 0$. Moreover, we cannot distinguish original covers from potential covers so that two sets η and γ must be combined with the same sets in the reverse order.

The general case is very complicated, so we restrict ourselves to two languages. We then observe that the covers are either the same or different. If they are the same, we have either $|\eta| = 2$ and $|\gamma| = |\delta| = 0$, or $|\gamma| = 2$ and $|\eta| = |\delta| = 0$. Thus, the probability is

$$\begin{aligned} & [\exp(-\lambda t)]^2 + [\exp(-\lambda t)]^2 [1 - \exp(-\lambda t)]^2 \\ & = \exp(-2\lambda t) [1 + (1 - \exp(-\lambda t))^2] \end{aligned}$$

which differs from the first order theory by the term in the square bracket.

The simplest case where the second-order theory is really required is that of four languages. We will illustrate the results by one expression. If k_{22} words are covered by two items both in two languages, k_4 words by one item in all languages, k_3 by one item in three languages, k_2 by one item in two languages, and k' by no common items, then the expression to be solved for maximum likelihood is

$$\frac{4k_{22} + 4k_4 + 3k_3 + 2k_2}{p} - \frac{2k_{22} + k_3 + 3k_2 + 5k'}{1 - p}$$

$$\begin{aligned}
& - \frac{4k_4(1-p)^3}{2-4p+6p^2-4p^3+p^4} + \frac{k_3(3-4p+4p^3)}{2+3p-2p^2+p^4} \\
& + \frac{k_2(4-2p-3p^2)}{2+4p-p^2-p^3} + \frac{k'(5+6p+9p^2)}{1+5p+3p^2+3p^3}
\end{aligned}$$

where $p = \exp(-\lambda t)$.

This second-order theory is not satisfactory not only because it leads to very complex formulas, but it also seems to be qualitatively inadequate. The formula for splitting between two languages is not greatly modified except for very long times, and the change does not seem to be enough to account for data showing short times of division. It is hard to tell whether the formula for several languages including the quantity k_{22} is any help--so far we have no striking results to quote from its use.

A second-order theory where potential cover decayed at a different rate than the original cover might correct some of these defects, but we have no evidence upon which to estimate the decay rate in this case. It is much likely that a more elaborate mechanism must be postulated--it need not lead to more elaborate results. The model must be based on a kind of dialectation study which seems to be absent as yet from the literature.

Conclusion

We have derived a number of formulas relating to the estimation of time depths by observations of lexical decay. The methods used can be applied to obtain many more similar formulas as required in studies of actual data.

All of these formulas are based on models of lexical decay using the concept of semantic atoms and their lexical covers. Lexical decay

is identified with a change in lexical cover. If the semantic atoms are sufficiently independent, the decay is a Poisson process.

Probably the most important practical conclusion is the result that any set of semantic atoms can be used to evaluate lexical decay provided the set is made up of atoms:

1. far enough removed in meaning from one another to assure independence,
2. which represent concepts assured to have been in existence throughout the time period being studied.

End Notes

- (1) See Robert B. Lees, "The Basis of Glottochronology" *Language*, 29.113-27 (1953).
- (2) There is no outstanding study of this problem. Attempts to "improve" the test vocabulary by limiting it to meanings which have behaved well in earlier studies are methodologically disastrous because they bias the value of λ .
- (3) This requirement is also intended to remove bias from the estimate of λ .
- (4) This is a matter of classical philological research independent of statistical syntheses made from the results.