


Does Science Progress? A Statistical Approach to Postmodern Theories of Knowledge

Daniel L. Chen

- Motivation
- Research Question
- Empirical Method
- Data
- Key Results
- Findings and Interpretations
- Next Steps

Motivation



Science and Culture Wars

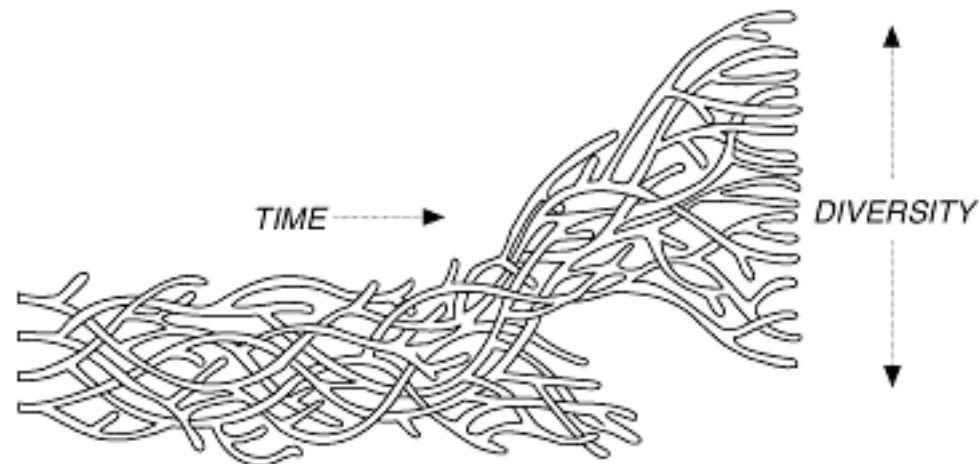
- **Thomas Kuhn's *Structure of Scientific Revolutions* (1962) proposes a theory of knowledge - science undergoes periodic paradigm shifts**
 - **Postmodernists - no truth in science ("Science as Solidarity," Rorty 1986)**
 - "Coming to Blows Over How Valid Science Really Is," *New York Times*, July 22, 2001.
 - "The Foundations of Science," *The Economist*, October 25, 2001.
 - "A Philosopher's Call to End All Paradigms," *The Chronicle of Higher Education*, Sept 15, 2000.
 - **Logical, gradual, progressive**
 - Sokal, Alan, "A Physicist Experiments with Cultural Studies," *Lingua Franca*, July 1996.
- **Economic Theory**
 - **Laws governing science and knowledge accumulation**
 - why some societies are technologically creative
 - economic incentives scientists face (Dasgupta, David)
 - Knowledge Production Function (Jones, Weitzman)
 - **Kuhn theory an untested theory of interest to economists (Stigler, Rosenberg)**
 - Microinnovation vs. macroinnovation (Mokyr)

Motivation



“The evolution of a science is a fascinating area for study: subtle, complex, but surely obeying laws which can eventually be discovered (Stigler 1982)”

- **Previous tests - archival, rhetoric, subjective, anecdotal**
 - Jones - age at which people come up with breakthrough innovations
 - Crane - #s of papers in two research subspecialties
- **Citation Tree (Graph)**
 - Each node is a paper connected to all the papers that cite it and it cites
 - Evolutionary biology
 - species tree has branches that get chopped off during times of mass extinction
 - subsequent branches flower off a small part of the previous tree



Research Question



Can we empirically detect paradigm shifts?

- **What is a paradigm shift?**
 - Theories, methods, concepts; entire worldview
 - Incommensurability - mutually unintelligible, not objectively comparable
- **Aristotelian vs. Newtonian physics (Kuhn's leading example)**
 - Aristotle: matter does not exist separate from qualities (position or solidity)
 - Wherever there's position, there is matter;
 - There is matter everywhere in space
 - Space without matter (vacuum) is impossible
- **Different religions**

Research Question



Conflicts arise when people do not acknowledge other perspectives can exist

- Rich data on transmission of scientific ideas (compared to those of religious or political ideas)
- Apt context to study incommensurability of perspectives

- Kuhn's theory obviously correct, obviously false

- **Economics/econometrics**
 - Networks
 - Uncover stylized facts to inform new theory and science policy

Empirical Method



Mass Extinctions

- Many branches getting sparse within short time period
- Science: Linguistics (Chomsky), Geology (Plate Tectonics) vs.
- Non-Science: Literary Criticism
 - If identical:
 - paradigm shifts not detectable
 - fads in non-sciences look the same as paradigm shifts in the sciences-
 - If different: help calibrate how parameters and speed of extinction
- Note: vast majority of scientific citations are positive (over 97%) (Hull 1998)
 - Looking for mass extinctions, not particular influential papers

Data



ISI Web of Science

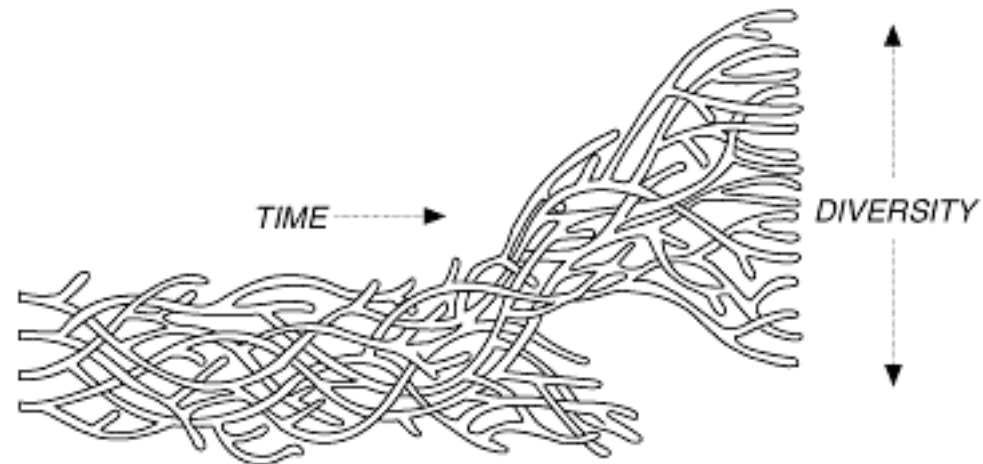
- **Linguistics: 1956-2001, 49,972 articles**
 - **Intersection of Linguistic Society of America list and ISI availability**
 - **70% downloaded (42 of 59 journals, including top 5)**
 - **Anthropological Linguistics, Journal of Linguistics, Language, Lingua, Natural Language & Linguistic Theory**
- **Geology: 1945-2001, 10,458 articles**
 - **2 journals: Journal of Geology and Geological Society of America Bulletin**
- **Comparative Literature: 1975-2001, 6,892 articles**
 - **3 journals: Publication of the Modern Language Association of America, Modern Philology, Modern Language Notes**
- **Patent Citation: 1975-1999, 2,088,795 patents**

Cluster Analysis



Method

- Many branches getting sparse within short time period
 - Van Dongen (2000) - last year of any bush = time of death
 - Cluster algorithm - random walks on graph: disconnected clusters are unlikely to have random walks beginning in one cluster reaching the other
- Comparative Literature
- Linguistics
- Geology
- Vary cluster size
 - Robustness
 - Smoothness patterns

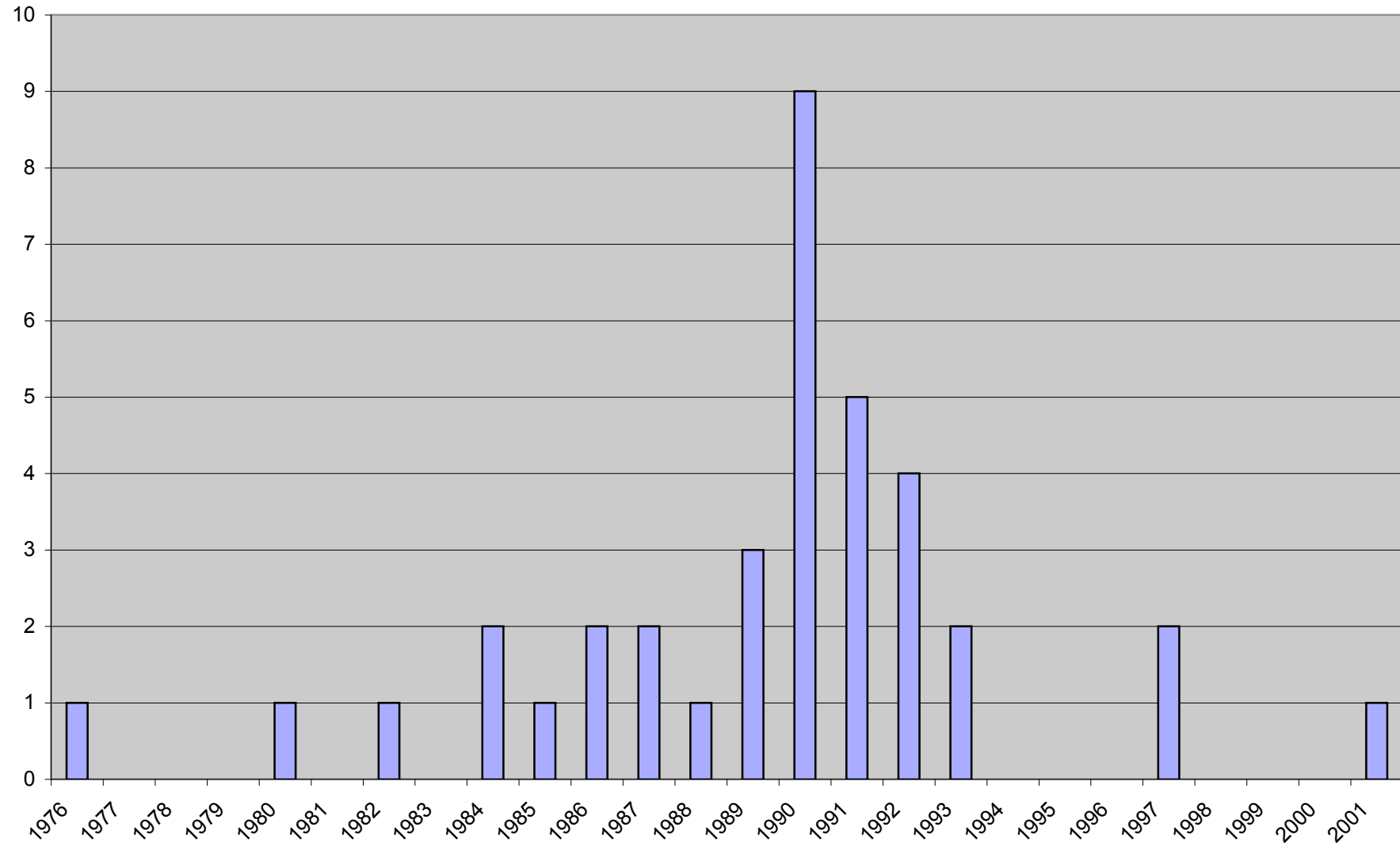


Cluster Analysis



Extinctions in Comparative Literature

Comparative Literature

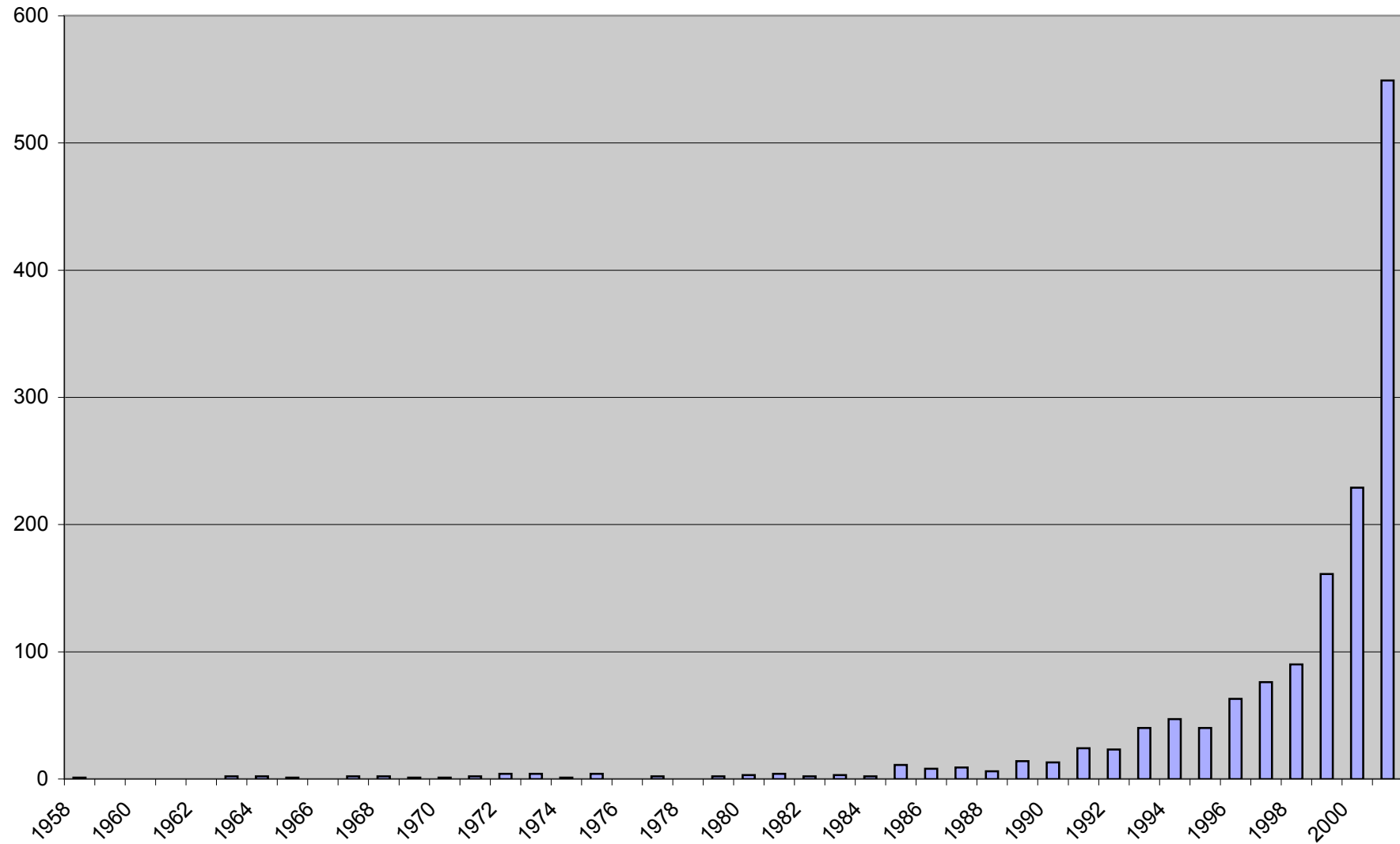


Cluster Analysis



Extinctions in Linguistics

Linguistics

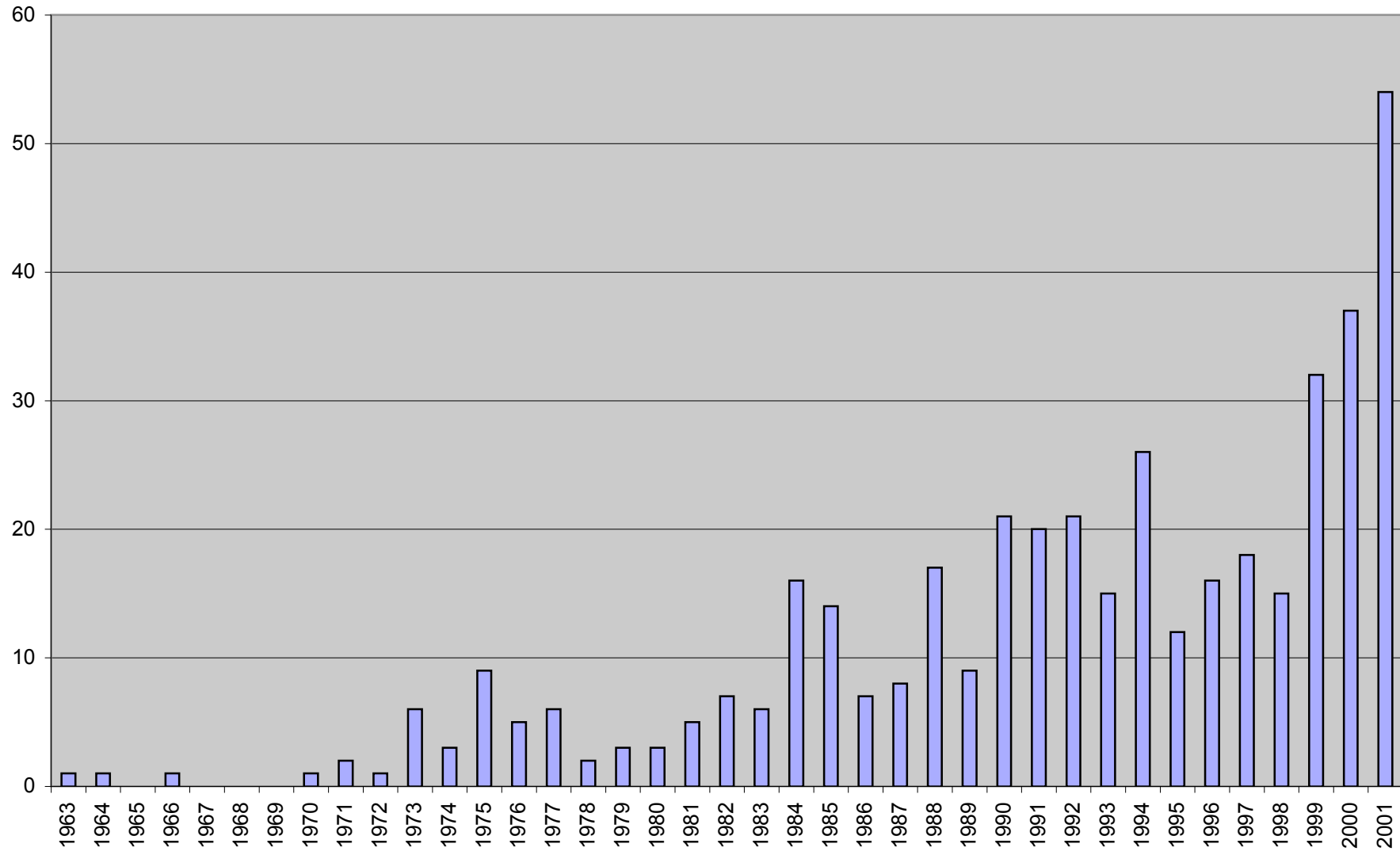


Cluster Analysis



Extinctions in Geology

Geology

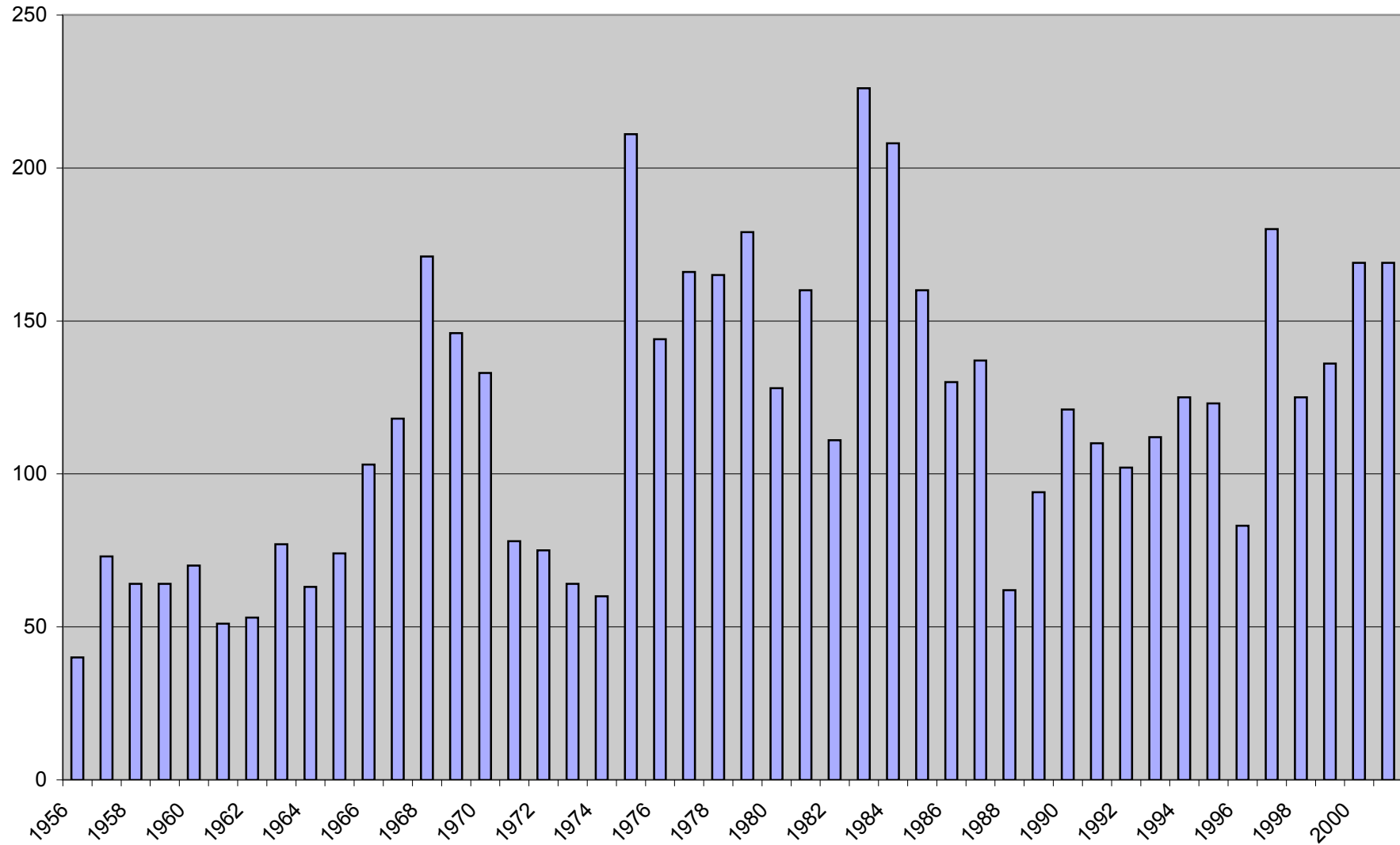


Cluster Analysis



Extinctions in Linguistics (small clusters)

Small Clusters

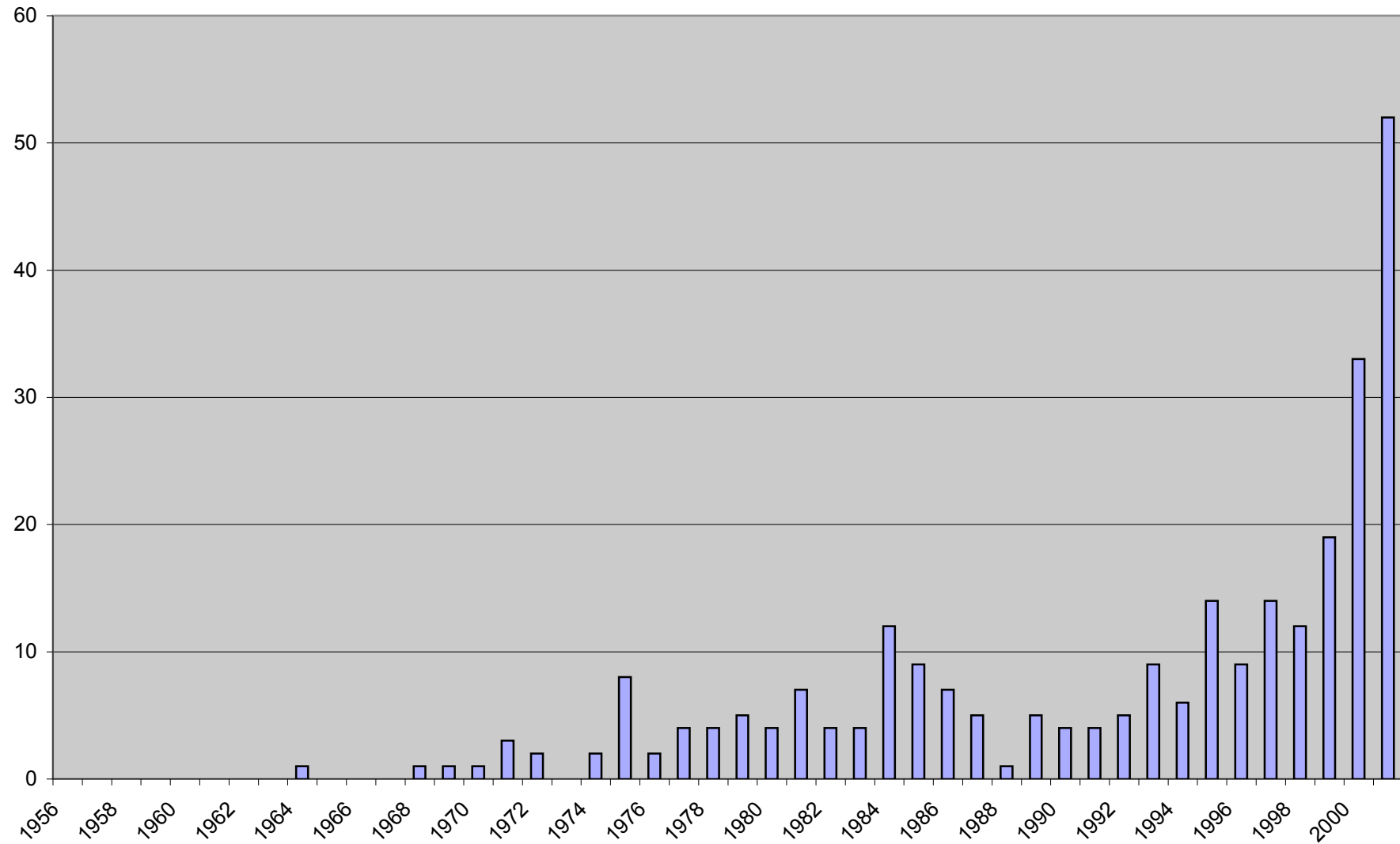


Cluster Analysis



Extinctions in Linguistics (medium clusters)

Medium Clusters

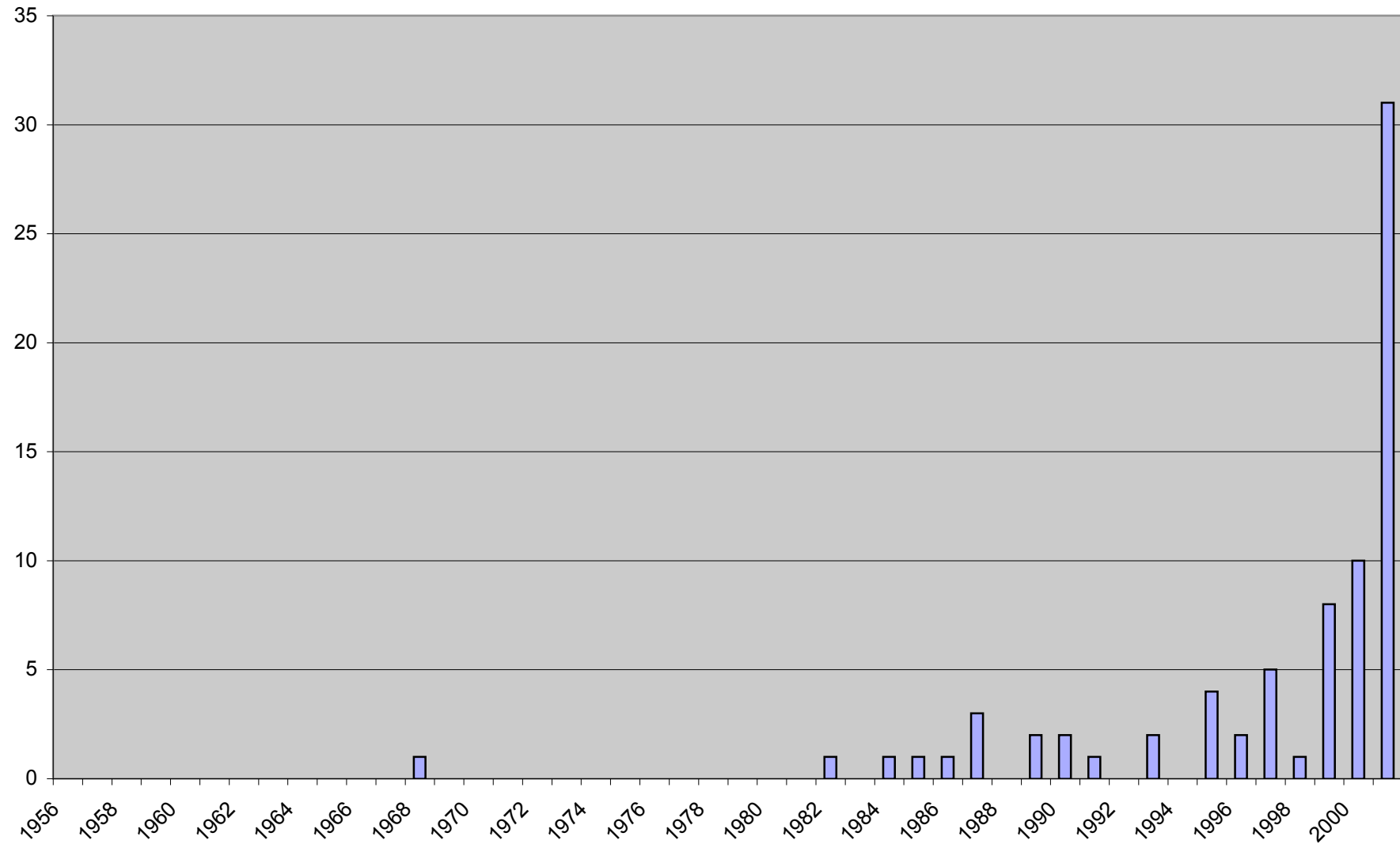


Cluster Analysis



Extinctions in Linguistics (large clusters)

Large Clusters

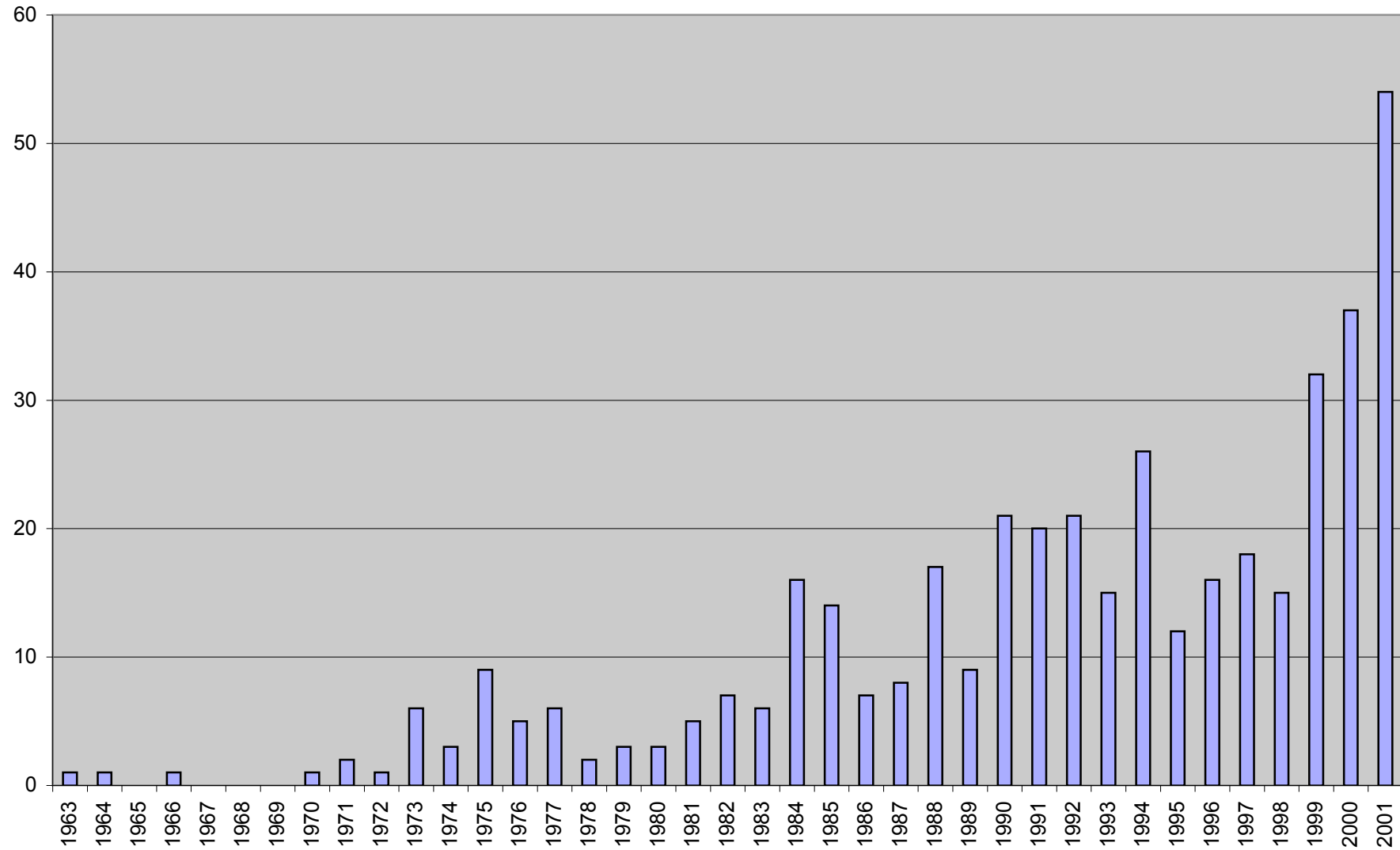


Cluster Analysis



Extinctions in Geology

Medium Clusters

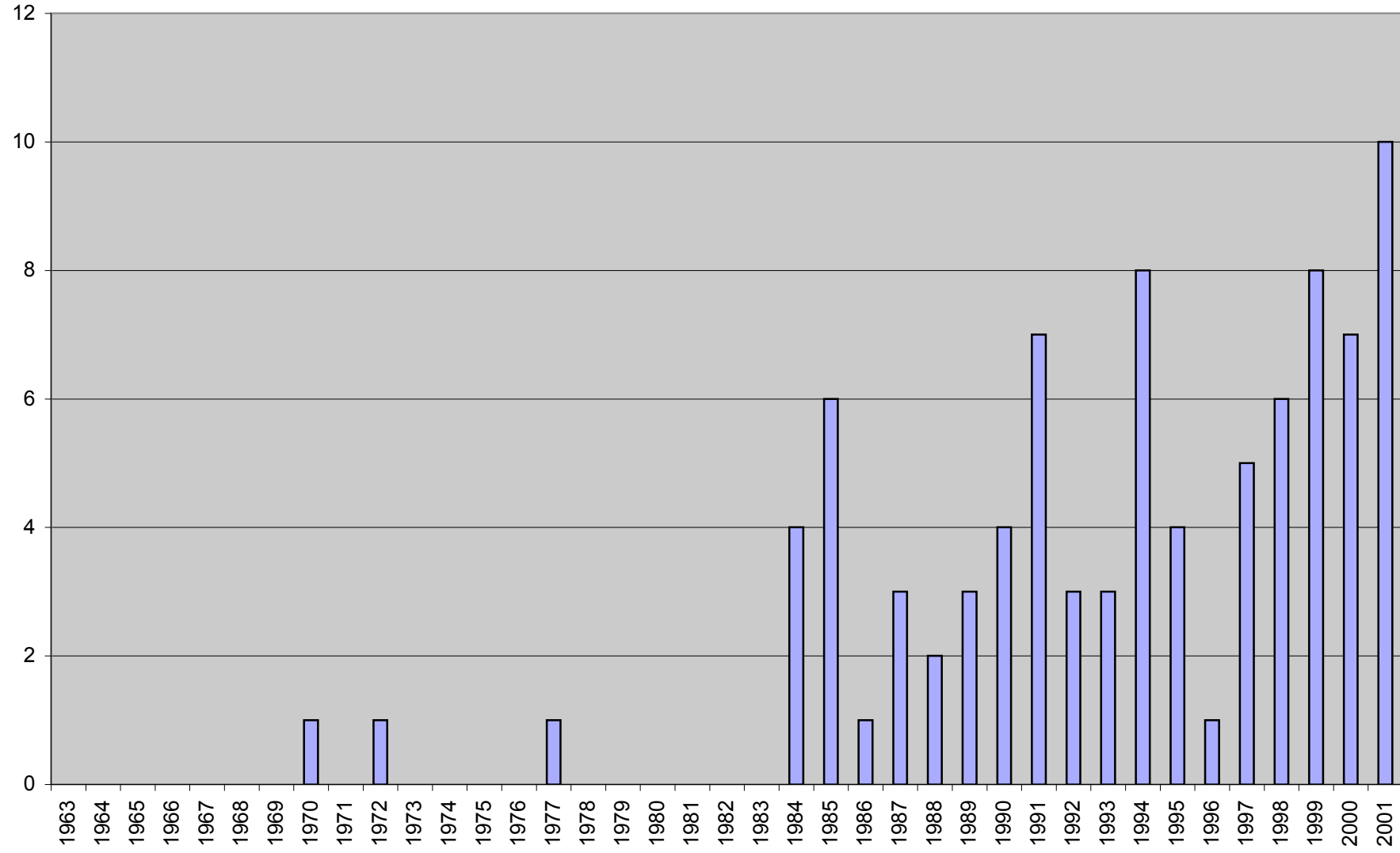


Cluster Analysis



Extinctions in Geology

Large Clusters

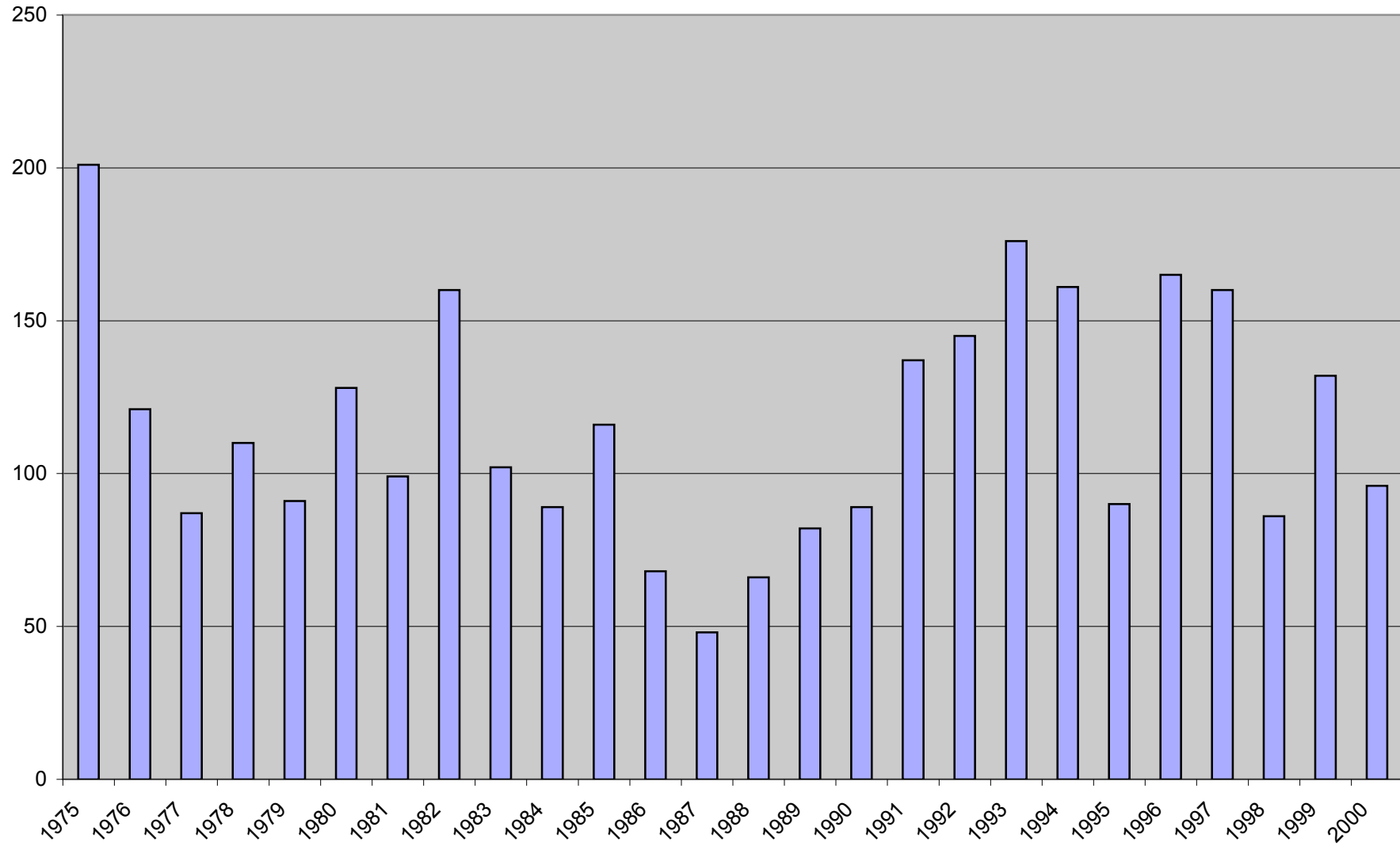


Cluster Analysis



Extinctions in Comparative Literature

Small Clusters

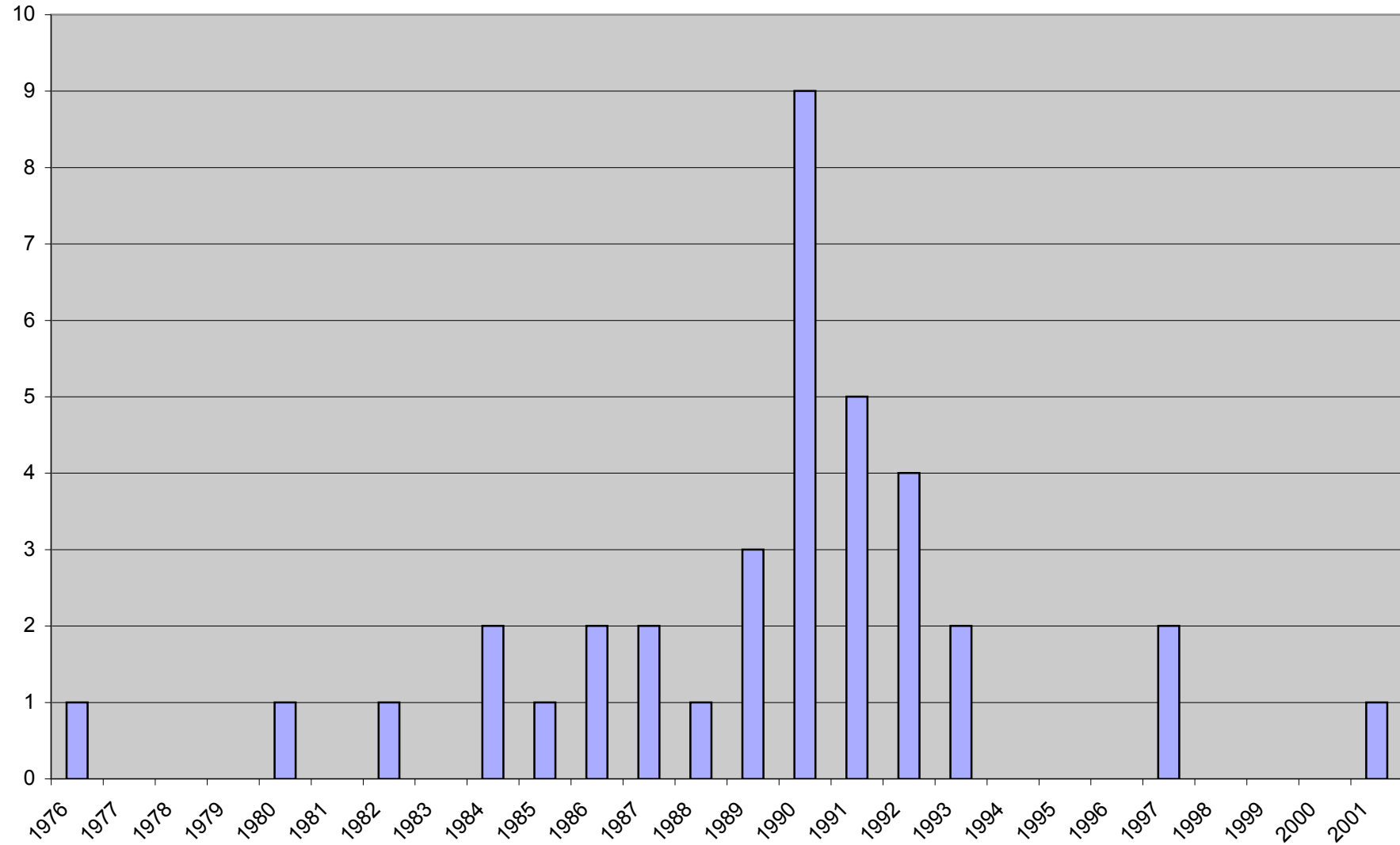


Cluster Analysis



Extinctions in Comparative Literature

Medium Clusters

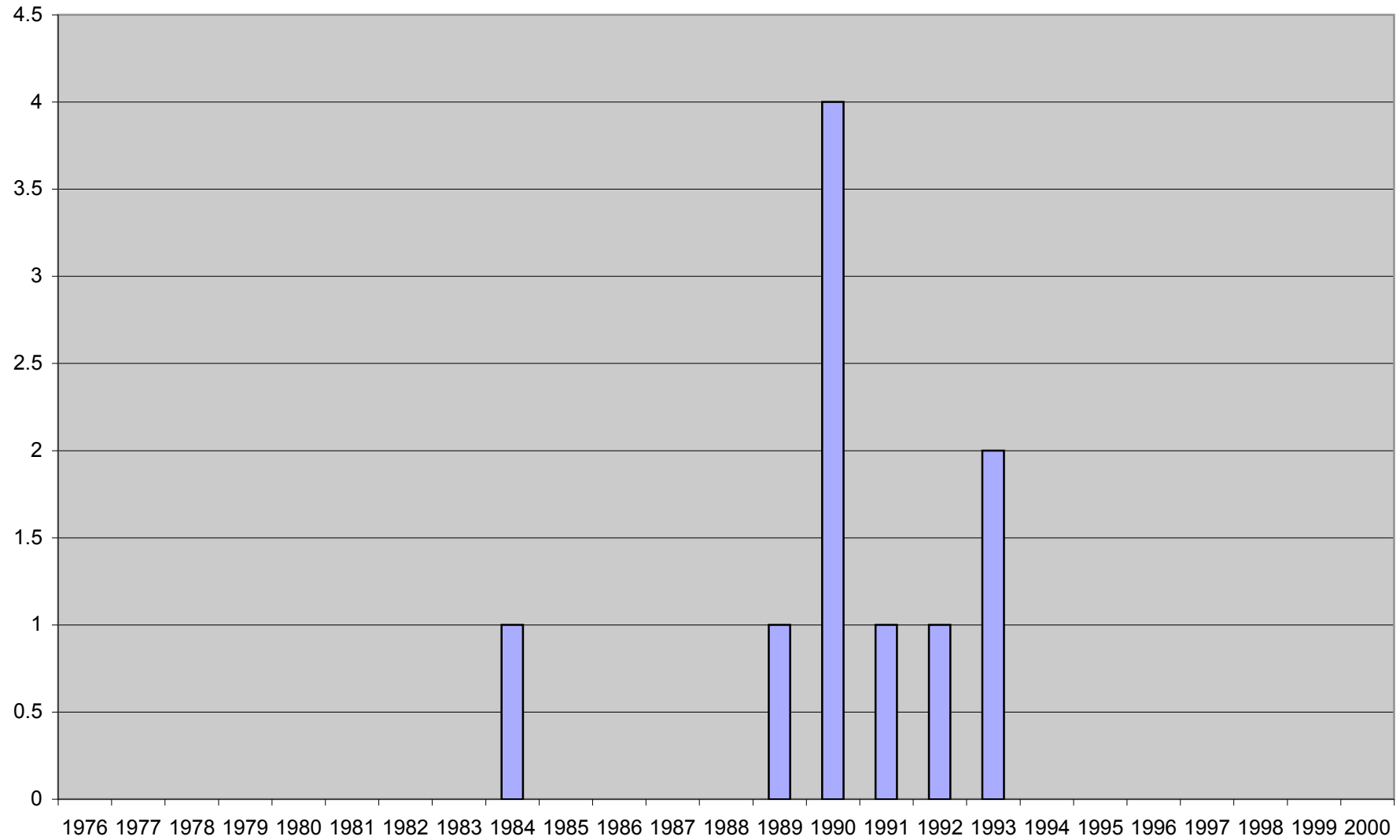


Cluster Analysis



Extinctions in Comparative Literature

Large Clusters



Fossil Imprint



Mass Extinctions

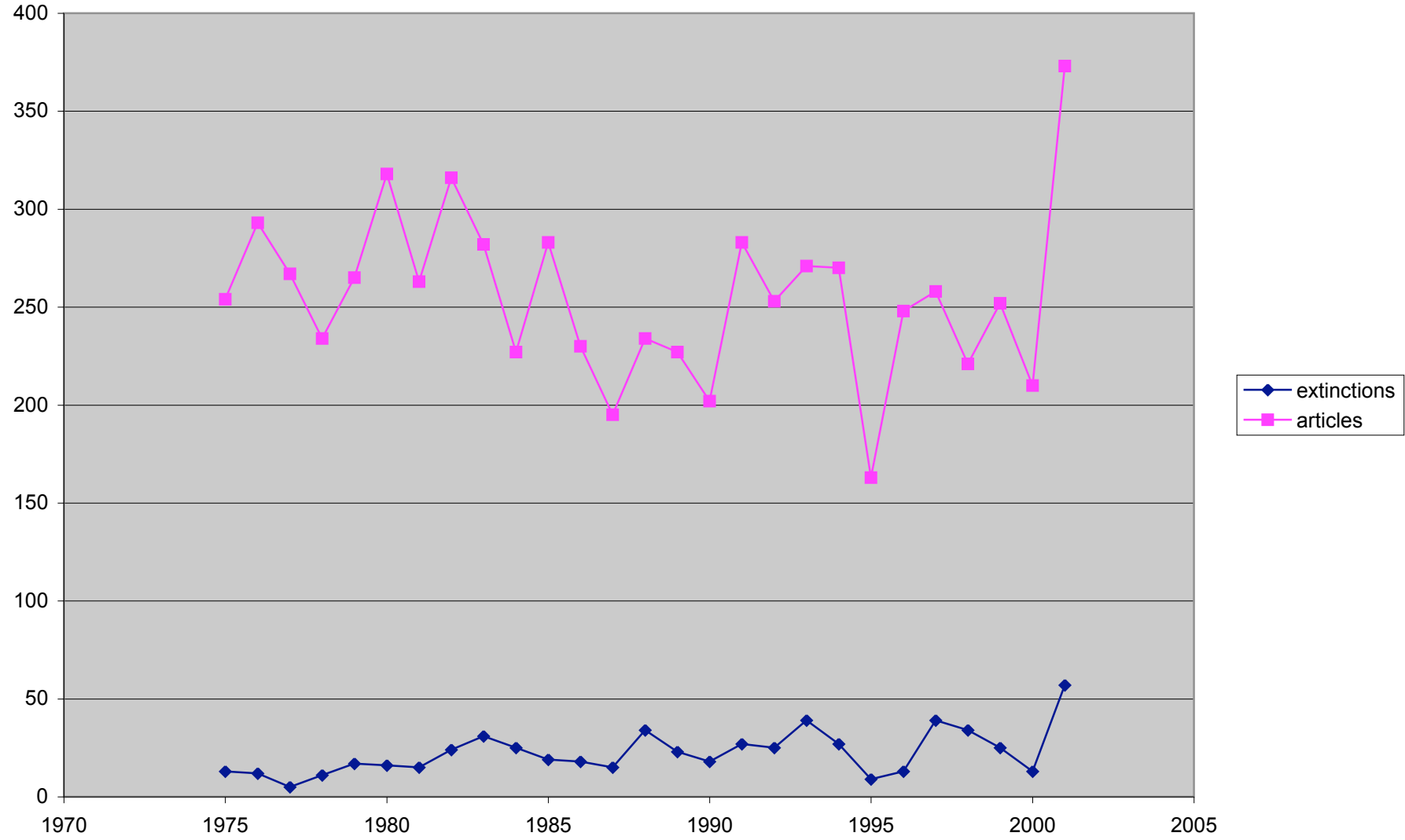
- Each citation as fossil imprint - time of extinction is last time it is cited
- Look for periods when extinctions rise above a background rate of extinction
- Comparative Literature - constant extinctions
- Linguistics - increasing extinctions
- Geology - increasing extinctions, despite level or decreasing # articles
- Patents - increasing extinctions

Fossil Imprint



Extinctions in Comparative Literature

Comparative Literature

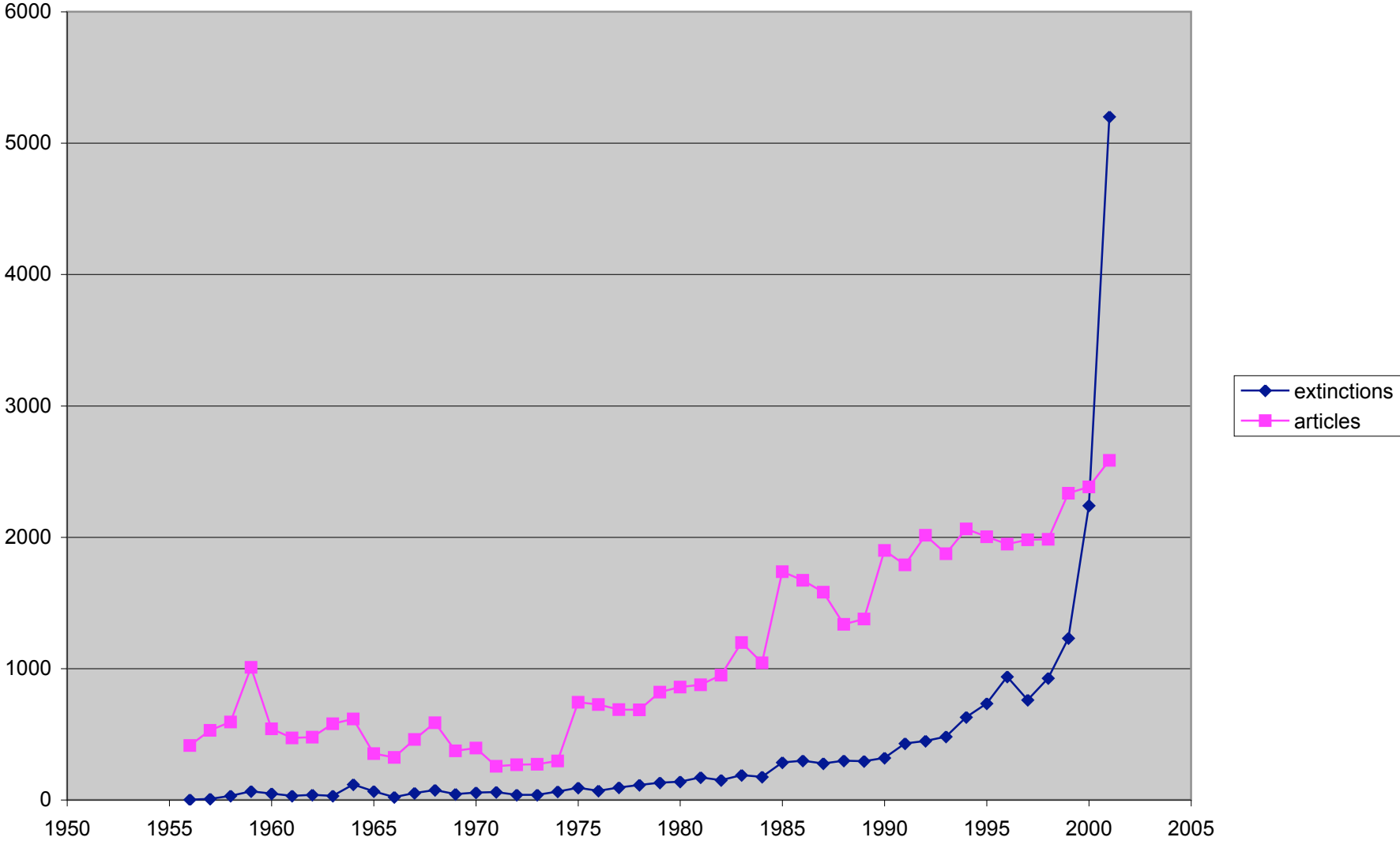


Fossil Imprint



Extinctions in Linguistics

Linguistics

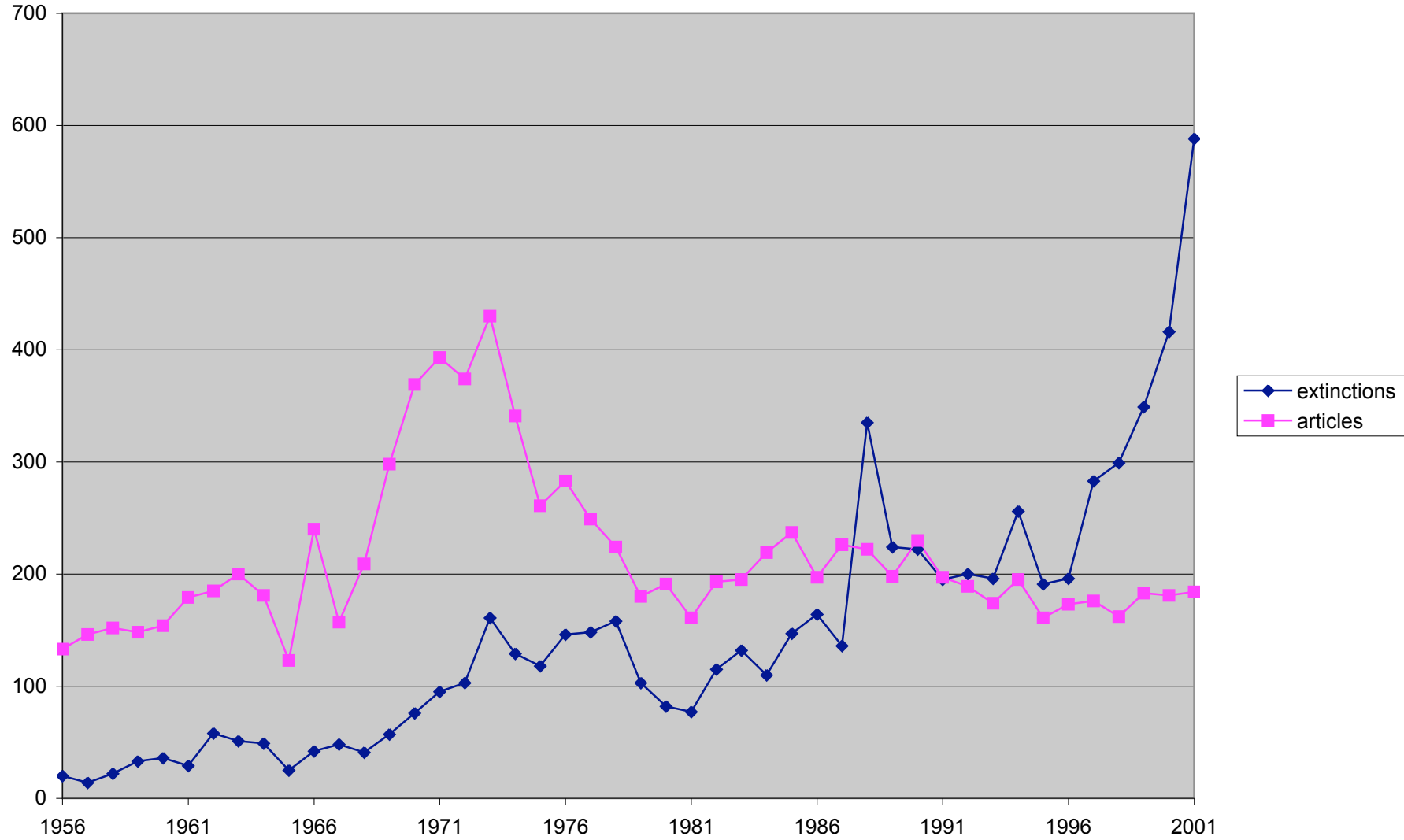


Fossil Imprint



Extinctions in Geology

Geology

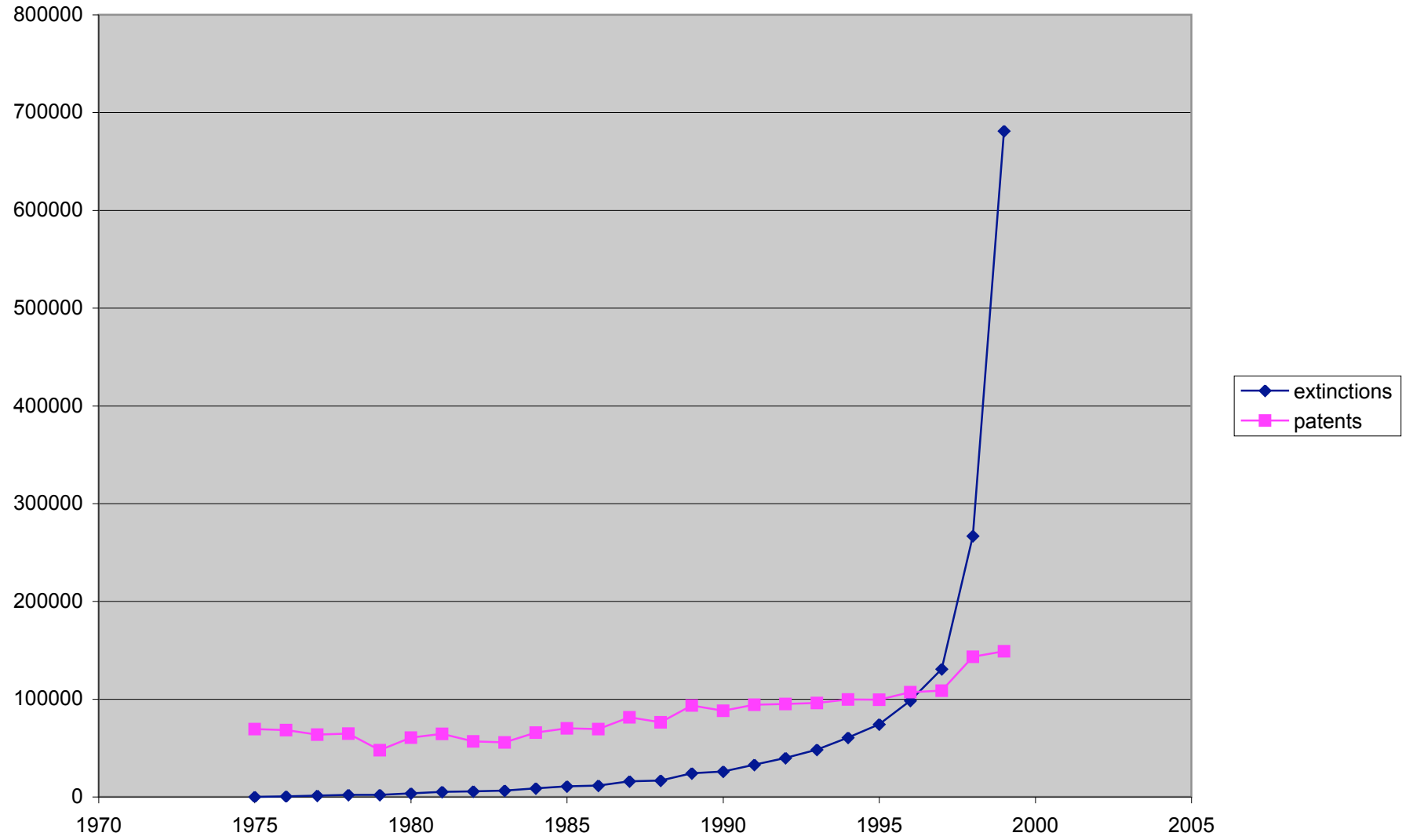


Fossil Imprint



Extinctions in Patent Citations

Patents



Findings and Interpretations



Literary Criticism accumulates differently from Science

- **Comparative Literature undergoes fads frequently**
 - **Possibly 1990 - deconstructionism (Derrida), postcolonial studies**
 - **Few articles are cited, extinctions common**
- **Linguistics**
 - **Chomsky 1964?**
 - **Branches end in 2001 / increasing extinctions**
 - **Shrink cluster size - looks more like comparative literature**
- **Geology**
 - **Similar to Linguistics, despite level or decreasing # of articles**
- **Patents**
 - **Similar to Linguistics and Geology**

Next Steps



Methods

- Big Bang
 - Birth of clusters
 - Each article' s earliest progenitor
 - Look for periods when there are unusually many births / progenitors
- Epidemics
 - Take into account the network structure, formally
 - Time to decay / mutation - citations of citations; later citations

Next Steps



Theories

- Evolutionary Epidemiology
 - Multiple paradigms coexisting
 - Branch length; Process of becoming sparse / extinct; Specialization (speciation)
- Knowledge Production Function
 - Branches arise from splitting or cross-fertilization (recombinant growth; Weitzman)?
 - branching growth - probability of extinction vs. no limits to growth
 - Likelihood of intellectual impact a function of distance between old ideas combined (Jacobs)?
 - When is interdisciplinary work productive?
 - Tradeoff in inputs to scientific progress
 - Can we detect lack of progress from branching structure? (grant \$, affirmative action)
 - Do legal areas subject to greater efficiency pressures progress more (Posner)?

Next Steps



Policy

- Linking biographical data
 - Affirmative Action - Ethnic, gender, ideological diversity stimulate knowledge?

- Legal data
 - Westlaw
 - Shephard's: Positive/negative citations

- Statistical Applications
 - Social interactions, spatial models, communication, genealogies, epidemics, email, internet