Mapping of Science by Combined Co-Citation and Word Analysis. II: Dynamical Aspects

Robert R. Braam, Henk F. Moed, and Anthony F. J. van Raan
Center for Science and Technology Studies (CWTS), University of Leiden, Wassenaarseweg 52, P.O. Box 9555, 2300 RB Leiden, The Netherlands

Combined analysis of co-citation relations and words is explored to study time-dependent ("dynamical") aspects of scientific activities, as expressed in research publications. This approach, using words originating from publications citing documents in co-citation clusters, offers an additional and complementary possibility to identify and link specialty literature through time, compared to the exclusive use of citations. Analysis of co-citation relations is used to locate and link groups of publications that share a consensus concerning intellectual base literature. Analysis of word-profile similarity is used to identify and link publication groups that belong to the same subject-matter research specialty. Different types of "content-words" are analyzed, including indexing terms, classification codes, and words from title and abstract of publications. The developed methods and techniques are illustrated using data of a specialty in atomic and molecular physics. For this specialty, it is shown that, over a period of 10 years, continuity in intellectual base was at a lower level than continuity in topics of current research. This finding indicates that a series of interesting new contributions are made in course of time, without vast alteration in general topics of research. However, within this framework, a more detailed analysis based on timeplots of individual cited key-articles and of content-words reveals a change from more rapid succession of new empirical studies to more retrospective, and theoretically oriented studies in later years.

Introduction

Combined analysis of co-citation relations and word-profile similarities is explored to improve the capability of quantitative techniques to depict structural and dynamical aspects of scientific research. In our foregoing publication ("Mapping I," Braam et al., this issue) we emphasize the structural aspects ("local stability") of "science mapping," while in this article we focus on the analysis of dynamical aspects ("temporal stability") of scientific research.

Starting with a clustering of documents that often co-occur in the reference lists of publications (co-citation clustering), publications in the dataset are grouped on the base of (one or more) citations to these clustered documents. This "classification" of publications is believed to partition the dataset according to participation of publications in research specialties (Small & Griffith, 1974; Griffith et al., 1974; Small, 1977; Small & Crane, 1979). The prevalent idea that the "current research" publications of specialties are identified in this way, is based on theories of Price (1965) and Kuhn (1970). In particular the way researchers draw on earlier work, and their sharing of a set of "exemplars" (or "paradigm"), is considered to be reflected in the referencing practices of specialty members. In this article, the concept of a scientific "specialty" will be used in the sense of "focused attention to a set of related research problems, and concepts by a number of scientific researchers," irrespective of the intellectual and social background of the researchers involved. These scientific researchers constitute the formal communication of the specialty, i.e., publications in journals, proceedings, or books. Such "specialties" are, of course, only covered by co-citation analysis if a *shared intellectual focus on earlier literature* exists together with *shared interest in a set of related research problems and concepts*. Such congruence is presupposed in most co-citation studies, but is problemised in our research project.

In order to describe the research topics involved in groups of "current-research" publications (as classified by co-citation clustering), word-profiles are constructed. Word-profiles are lists of content-words related to publications, such as indexing terms and classification codes added to these publications by pro-

This study is part of a project financed by the Ministry of Education and Sciences, through the Netherlands Advisory Council for Science Policy (RAWB).

Received February 24, 1989; revised October 4, 1989; accepted October 10, 1989.

 CCC 0002-8231/91/040252-15$04.00

fessional indexing services, or words in title or abstract. Aggregation of word-profiles from publications citing (one or more) documents in a co-citation cluster, results in "group word-profiles" for the current research of specialties. These word-profiles can then be used to evaluate the degree of cognitive coherence within and similarity between these sets of "current research" publications ("Mapping I," Braam et al., 1990).

Analysis of word-profile similarities between co-citation based publication groups, makes it possible to identify *and* link (co-citation based) specialty literature not only within a given period, but also through time. Groups of "current research" publications identified by co-citation clusters of one given year, that are strongly related by word-profile similarity, can be seen as representing parts of the same, larger specialty. Between these groups, consensus concerning intellectual base literature is apparently not shared, or is relatively low. Publication groups identified by co-citation clusters from different years that have high word-profile similarity, can be considered—approximately—as different phases of the same specialty. If the clusters by which these publication groups are identified also share a high percentage of cited documents, this indicates stability in focus on earlier literature, i.e., stability in the intellectual base of a specialty. If stability in the intellectual base of a specialty is low, analysis of word-profile similarity is more useful than co-citation analysis in order to track down specialty literature through time. As there are no a priori reasons to assume stability in the intellectual base of specialties, nor to assume that different specialties cannot have highly identical intellectual bases, exclusive use of citations to track specialty literature through time is prone to errors.

It should be noted here, that the word-profiles based on indexing terms represent an ("external") indexer-viewpoint as to what research topics are involved in publications, whereas title and abstract words represent an ("internal") author-viewpoint.

In the present case, we evaluate the usefulness of word analysis in addition to prior grouping of publications based on co-citation analysis. It is, however, also possible to use an analysis of words in an independent fashion to group publications. In a related article we present results of a co-word clustering and compare these results to those of a co-citation cluster-based grouping of the same set of publications (Braam et al., 1989).

In this study, co-citation cluster-based publication groups of 10 successive years are analyzed for a specialty in atomic and molecular physics (viz. Rydberg Atoms and Molecules), in relation to their degree of cognitive resemblance (for the operationalization of this concept we again refer to our foregoing article Mapping I), using profiles of content-words representing an internal viewpoint, as well as profiles representing an external viewpoint. Further, time-plots of individual cited documents and content-words related to clusters are constructed and compared in order to study dynamic aspects in some more detail.

It will be shown that such additional word analysis provides an interesting possibility to improve interpretation of dynamic aspects of research in specialty literature delimited by co-citation analysis.

Data, Methods, and Techniques

Data

Publication data (including indexing terms, professional field-classification codes, titles, and abstracts), have been retrieved from INSPEC for 295 source publications from the period 1974–1986, based on a field expert's document file. References of these source publications have been retrieved from ISI/SCI (coverage of these publications by ISI was 93%). Data on these publications from both sources were then combined using dedicated software (Moed, 1988). The resulting database contains 273 source publications, with a total of 4225 different cited references of which 585 (13.8%) references were cited more than once, and 177 (4.2%) references were cited more than twice in one or more years. The numbers of different content "words" (either controlled terms, uncontrolled terms, classification codes, or title and abstract words) involved in these source publications is shown below together with the percentage of these terms, codes and words that are "unique" (i.e., occur exclusively in one publication) (Table I). Connectives, prepositions, pronouns, articles, and some verbs ("to be," and "to have") from the title and abstract of publications, were excluded because these words do not relate to the content of the publications involved.

The annual number of publications in the expert's document file was compared to an estimated* number of publications from this specialty, as contained in the INSPEC file. As shown in Figure 1, the expert's document file contains a subset of all publications in INSPEC identified as relevant to the "Rydberg" specialty. The discrepancy between the two lines in

TABLE I.

Content Words	All	Unique
controlled indexing terms	248	43%
uncontrolled indexing terms	2123	84%
classification codes	163	36%
title words	836	58%
abstract words	3423	52%

*This was done by using a set of controlled and uncontrolled terms, based on analysis of such terms in the expert's document file, to select documents in the INSPEC file, and inspection of the relevance of retrieved documents by the expert.

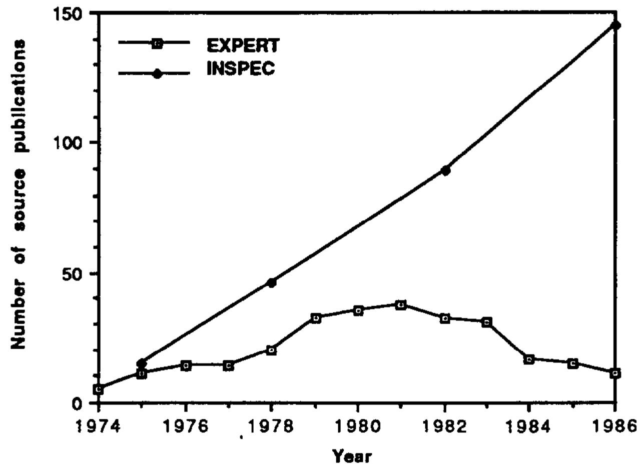


FIG. 1. Comparison of the number of Rydberg publications in an expert's document file and in the INSPEC online database.

Figure 1 reflects individual scope and interest of the expert as well as the expert's entrance and departure of the specialty. Our dataset, thus, represents the specialty as seen and appreciated by a single expert and not the "entire" specialty, though the selection is of course still a cognitively coherent collection of documents.

Methods and Techniques of Data Processing

A detailed account of co-citation analysis as developed and performed by ISI, can be found in the review articles of Small and colleagues on this subject (Small & Sweeney, 1985; Small et al., 1985). A description of our approach to the problem of choosing appropriate threshold levels, and of our approach to the description of research topics related to co-citation clusters was presented earlier (Braam et al. 1988). A fuller account of our method of combined co-citation and word analysis is given in our foregoing publication ("Mapping I," Braam et al. 1990).

As described in detail in Mapping I, analysis of word profile similarities between co-citation cluster based publication groups is used to evaluate variance between research topics involved in these publication groups. The possibility this method offers to track clusters through time, is in fact an addition to the method used by Small (1977), where clusters are followed through time by examining the number of key-articles (i.e., clustered highly cited documents) shared by co-citation clusters of successive years. These two methods, one using words related to publications citing co-citation clusters and the other using clustered cited documents, are complementary.

It will be clear, that our approach rests on the following two assumptions. First, that continuity in the use of specific content-words as subject-descriptors of citing publications of successive years, does indeed reflect stability in research topics in the current work of a specialty, as judged by indexers and authors respectively. Secondly, the assumption that persistence of key-articles in clusters of successive years indeed does reflect stability in the intellectual base of a specialty, as indicated by authors.

If these assumptions hold, *continuity in research topics* involved in the current work of a specialty, in as far as represented by the citing publications of co-citations clusters, can be established by determining the similarity of word-profiles, using the cosine formula (Jones & Furnas, 1987):

$$\mathrm{Sim}(A,B) = \frac{\sum_{i=1}^{n} W(Ai) * W(Bi)}{\sqrt{\sum_{k=1}^{n} (W(Ak))^2} * \sqrt{\sum_{j=1}^{n} (W(Bj))^2}} \quad (1)$$

where,

$W(Ai)$ = weight of term i for cluster A, in the boolean case $W(Ai) = 0$ (absence) or 1 (presence);
$W(Bj)$ = weight of term i for cluster B, in the boolean case $W(Bj) = 0$ (absence) or 1 (presence);
n = total number of terms for cluster A and cluster B together.

In order to study *dynamics of research topics* in more detail—in addition to examination of similarity values—a plot of the temporal development of individual words is made, i.e., we display which words emerge, persist, or disappear in the course of time.

The number of key-articles shared by clusters of successive periods (for example years), is determined, and expressed as an index value. For this purpose a "stability index" (Small, 1977) can be used. For clusters of successive years, the use of such index is essentially equivalent to the use of the Jaccard Index (Sneath & Sokal, 1973), which gives the cardinal number of the intersection of different clusters divided by the cardinal number of their union. In this study we used this latter index in order to determine *continuity in intellectual base*.

In order to study *changes in the intellectual base* in some more detail, a plot of the development of individual clustered key-articles is constructed.

Results and Discussion

Co-Citation Clustering

Co-citation clusters have been formed for the years 1976–1985, using a citation threshold of two citations per year (i.e., three citations or more are required for any year), and a co-citation strength threshold (using the cosine formula) equal to the median of this strength for each separate year.* The resulting time series of clusters is shown in Figure 2, together with links between clusters from different years based on shared key-articles (i.e., links indicating continuity in the intellectual base of research in this specialty). The clusters include for each year 80–100% of all yearly cited documents that reach the citation threshold. The number of publications citing documents included in the various clusters ranges between 11 and 27 per year. Assuming all source publications in the (expert-based) dataset to be relevant to the "Rydberg" specialty, the recall value of co-citation analysis varies in this case between 35% and 76% per year (58% over all years). Thus, if the above assumption is correct, co-citation clustering results are incomplete.

*The median co-citation strength values are in the range 0.30–0.50 for the years 1976–1985. This measure of central tendency was chosen for the sake of comparability of results for the different years. As the strength values were not normally distributed in all years, the median was preferred above the mean as a comparable measure of central tendency.

Although the expert's document file can be regarded as a cognitive unit, in most years more than one cluster is formed. Whether this split-up relates only to differences in intellectual base, or also to differences in research topics studied, will be discussed below. Further, not all clusters have links to older and/or younger clusters, while some clusters have lots of links. Most links concern the larger (size in terms of the number of clustered key-articles) clusters. In general, there seems to be a main line of continuity including the larger clusters for each year, accompanied by short sidelines with smaller clusters from only a few years in the past.

Description of Research Topics Related to Clusters

Research topics involved in the current work related to these co-citation clusters have been described using indexing terms (controlled and uncontrolled terms from INSPEC), and words from title and abstract of publications citing these co-citation clusters. Results are shown in Figures 3(a) and 3(b), excluding "unique"

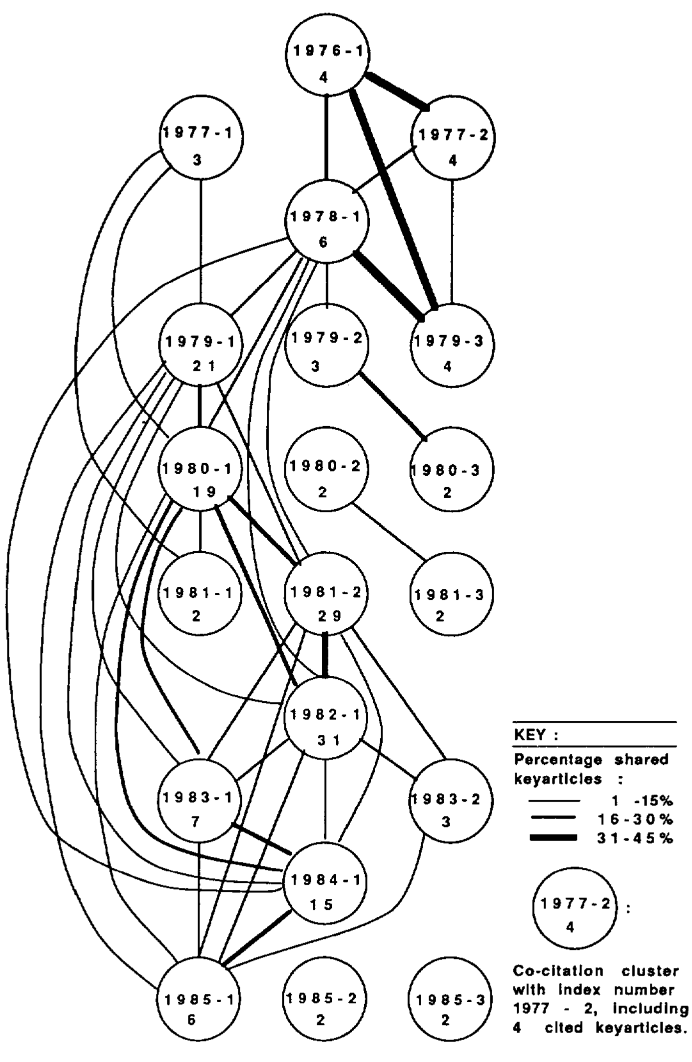


FIG. 2. Time series of Rydberg clusters for the period 1976–1985.

1976 - 1 (5/5)

CT: 5 ATOMIC EXCITED STATES
CT: 3 SODIUM
CT: 2 IONISATION OF ATOMS
UT: 2 NA
CT: 2 RUBIDIUM
CT: 2 STARK EFFECT

1977 - 1 (4/5)

CT: 4 ATOMIC EXCITED STATES
CT: 2 ATOMIC FLUORESCENCE
UT: 2 RYDBERG STATES

1978 - 1 (11/11)

CT: 8 ATOMIC EXCITED STATES
CT: 4 SODIUM
CT: 3 ATOMMOLECULE COLLISIONS
CT: 3 FIELD IONISATION (UT: 2)
CT: 3 IONISATION OF ATOMS
CT: 3 ATOMATOM COLLISIONS
CT: 2 ATOMIC SPECTRA
UT: 2 NA
CT: 2 OPTICAL PUMPING
UT: 2 RYDBERG STATES
CT: 2 XENON

1977 - 2 (6/7)

CT: 4 SODIUM
CT: 3 ATOMIC EXCITED STATES
CT: 3 ATOMIC FINE STRUCTURE
CT: 3 ATOMIC SPECTRA
UT: 3 NA
CT: 2 ATOMIC HYPERFINE STRUCTURE
CT: 2 FIELD IONISATION
UT: 2 HYPERFINE STRUCTURE
UT: 2 QUANTUM DEFECTS
UT: 2 RF SPECTROSCOPY
CT: 2 RUBIDIUM
UT: 2 RYDBERG STATES
CT: 2 STARK EFFECT
UT: 2 STARK SHIFTS

1979 - 1 (11/16)

CT: 7 ATOMATOM COLLISSIONS
CT: 7 ATOMIC EXCITED STATES
CT: 2 ARGON
CT: 2 ATOMIC SPECTRAL LINE BREADTH
CT: 2 FIELD IONISATION
CT: 2 HELIUM ATOMS
UT: 2 HIGH RYDBERG ATOMS
CT: 2 INERT GASES
CT: 2 IONISATION OF ATOMS
UT: 2 LCHANGING COLLISIONS
CT: 2 NEON
UT: 2 PRINCIPAL QUANTUM NUMBER
UT: 2 RB
CT: 2 RUBIDIUM
UT: 2 RYDBERG STATES
UT: 2 SELF BROADENING
CT: 2 SODIUM
UT: 2 THERMIONIC DETECTOR

1979 - 2 (3/5)

CT: 2 CAESIUM
CT: 2 SUPERRADIANCE

1979 - 3 (3/8)

CT: 2 ATOMIC EXCITED STATES
CT: 2 RUBIDIUM
UT: 2 RYDBERG LEVELS
CT: 2 STARK EFFECT

1980 - 1 (9/12)

CT: 8 ATOMIC EXCITED STATES
CT: 5 ATOMATOM COLLISIONS
CT: 5 RUBIDIUM
CT: 3 ATOMIC SPECTRAL LINE BREADTH
UT: 3 DSTATE
CT: 2 ATOMIC FLUORESCENCE
CT: 2 ATOMIC SPECTRA
CT: 2 ATOMMOLECULE COLLISIONS
UT: 2 RB
CT: 2 SODIUM
CT: 2 SPECTRAL LINE SHIFT
UT: 2 SSTATE
UT: 2 TRILEVEL ECHO TECHNIQUE
CT: 2 TWOPHOTON SPECTRA

1980 - 2 (3/3)

CT: 3 AB INITIO CALCULATIONS
CT: 3 MOLECULAR ELECTRON CORRELATIONS
CT: 3 MOLECULAR ELECTRONIC STATES
UT: 3 RYDBERG STATES
UT: 2 CONFIGURATION INTERACTION STUDY
CT: 2 MOLECULAR VIBRATION
CT: 2 OXYGEN
UT: 2 O2
UT: 2 POTENTIAL CURVES
CT: 2 POTENTIAL ENERGY CURVES
AND SURFACES

1980 - 3 (2/5)

1981 - 1 (3/3)

CT: 3 HYDROGEN NEUTRAL
MOLECULES
UT: 3 H2

1981 - 2 (21/21)

CT: 11 ATOMIC EXCITED STATES
CT: 7 SODIUM
CT: 6 STARK EFFECT
CT: 4 ATOMATOM COLLISIONS
CT: 4 ATOMIC SPECTRA
CT: 4 HELIUM ATOMS
CT: 4 XENON
CT: 3 ATOMIC INELASTIC COLISIONS
UT: 3 BLACK BODY RADIATION
CT: 3 IONISATIONS OF ATOMS
UT: 3 NA
UT: 3 RYDBERG ATOMS
CT: 2 ARGON
CT: 2 ATOMIC STRUCTURE
CT: 2 ATOMION COLLISIONS
CT: 2 ATOMMOLECULE COLLISIONS
UT: 2 COUPLED CHANNEL CALCULATIONS
CT: 2 FIELD IONISATION
CT: 2 HYDROGEN COMPOUNDS
CT: 2 HYDROGEN NEUTRAL ATOMS
UT: 2 NA+AR
UT: 2 NA+HE
UT: 2 RATE CONSTANTS
UT: 2 RB
CT: 2 RUBIDIUM
UT: 2 RYDBERG ATOM
UT: 2 RYDBERG LEVELS
UT: 2 STARK SHIFT
CT: 2 TIME RESOLVED SPECTRA
UT: 2 XE

1981 - 3 (3/3)

CT: 3 TWOPHOTON SPECTRA
CT: 2 MOLECULAR ELECTRONIC
STATES
CT: 2 MOLECULAR ROTATION

1982 - 1 (21/21)

CT: 11 ATOMIC EXCITED STATES
CT: 7 SODIUM
CT: 6 ATOMATOM COLLISIONS
UT: 6 NA
CT: 4 ATOMIC SPECTRA
CT: 4 ATOMIC SPECTRAL LINE BREADTH
UT: 4 RB
UT: 4 RYDBERG STATES
CT: 4 SPECTRAL LINE SHIFT
UT: 3 EXITED STATES
CT: 3 FIELD IONISATION
CT: 3 IONISATION OF ATOMS
CT: 3 IONISATION OF MOLECULES
CT: 3 IONISATION POTENTIAL
CT: 3 MOLECULAR ELECTRONIC STATES
CT: 3 RUBIDIUM
UT: 3 RYDBERG ATOMS
UT: 3 SELECTIVE FIELD IONISATION
UT: 2 ALKALI ATOMS
CT: 2 ALKALI METALS
CT: 2 ARGON
CT: 2 AUTOIONISATION
CT: 2 HELIUM ATOMS
UT: 2 IONIZATION POTENTIAL
CT: 2 MASERS
CT: 2 OPTICAL PUMPING
CT: 2 PHOTOIONISATION
CT: 2 POTASSIUM
UT: 2 QUENCHING CROSS SECTIONS
UT: 2 RADIATIVE PROPERTIES
UT: 2 RYDBERG LEVELS
UT: 2 SUPERRADIANCE THEORY
CT: 2 TWOPHOTON SPECTRA

1983 - 1 (5/6)

CT: 5 ATOMIC EXCITED STATES
CT: 4 ATOMATOM COLLISIONS
CT: 3 ATOMIC SPECTRAL LINE BREADTH
UT: 2 PRINCIPAL QUANTUM NUMBERS
CT: 2 RUBIDIUM
UT: 2 RYDBERG ATOMS

1983 - 2 (6/7)

CT: 4 ATOMIC EXCITED STATES
CT: 4 SODIUM
UT: 3 NA
CT: 2 ATOMIC BEAMS
UT: 2 RYDBERG ATOMS

1984 - 1 (6/6)

CT: 3 ATOMMOLECULE COLLISIONS
UT: 3 IMPULSE APPROXIMATION
UT: 3 SHORTRANGE INTERACTION
CT: 2 ATOMATOM COLLISIONS
CT: 2 ATOMIC EXITED STATES
CT: 2 ATOMIC INELASTIC COLLISIONS
UT: 2 QUASIELASTIC PROCESS
CT: 2 RADIATION QUENCHING
UT: 2 RB
CT: 2 RUBIDIUM

1985 - 1 (3/6)

CT: 3 ATOMIC EXCITED STATES
CT: 2 ATOMATOM COLLISIONS

1985 - 2 (3/3)

CT: 2 AUTOIONISATION
CT: 2 HYDROGEN NEUTRAL MOLECULES
UT: 2 H2
CT: 2 IONISATION OF MOLECULES
CT: 2 PHOTOIONISATION
CT: 2 SPECTRA OF DIATOMIC INORGANIC
MOLECULES

1985 - 3 (1/4)

Key :
1985 - 1 (3/6) : Co-citation cluster with index number 1985 - 1, and with 6 citing publications from which 3 are 'central' (i.e. these publications cite exclusively to one cluster).
Controlled terms (CT) and uncontrolled terms (UT); terms occurring only once per cluster not shown.

FIG. 3(a). Indexing terms aggregated over central citing publications for Rydberg clusters 1976–1985.

1976 - 1 (5/5)

TT: 5 STATES
TT: 4 EXCITED
TT: 3 ELECTRIC
TT: 3 HIGHLY
AT: 5 STATES
AT: 3 EXCITED
AT: 3 FOUND
AT: 3 LEVEL

1977 - 1 (4/5)

TT: 4 RYDBERG
TT: 4 STATES
TT: 2 DEPOPULATION
AT: 3 STATES
AT: 2 ATOM
AT: 2 CROSS
AT: 2 ELECTRON

1977 - 2 (6/7)

TT: 4 RYDBERG
TT: 4 STATES
TT: 2 MEASUREMENTS
TT: 2 RUBIDIUM
AT: 4 QUANTUM
AT: 4 STATES
AT: 4 STRUCTURE
AT: 4 USED

1978 - 1 (11/11)

TT: 7 RYDBERG
TT: 6 STATES
TT: 3 ATOMS
TT: 2 COLLISION
AT: 8 RYDBERG
AT: 5 STATES
AT: 4 ATOMIC
AT: 4 ELECTRON

1979 - 1 (11/16)

TT: 6 ATOMS
TT: 4 RYDBERG
TT: 3 COLLISIONS
TT: 3 STATES
AT: 7 ATOMS
AT: 7 CROSS
AT: 7 SECTIONS
AT: 7 STATES

1979 - 2 (3/5)

TT: 2 RYDBERG
AT: 2 OBTAINED
AT: 2 RYDBERG
AT: 2 USING

1979 - 3 (3/8)

TT: 2 ABSORPTION
TT: 2 LEVELS
TT: 2 RYDBERG
AT: 3 LEVELS
AT: 3 TRANSITIONS
AT: 2 ATOMIC
AT: 2 GROUND

1980 - 1 (9/12)

TT: 5 RYDBERG
TT: 4 STATES
TT: 3 ATOMS
TT: 3 BROADENING
AT: 6 MEASURED
AT: 5 CROSS
AT: 5 N
AT: 5 RYDBERG

1980 - 2 (3/3)

TT: 3 AB
TT: 3 CALCULATIONS
TT: 3 INITIO
TT: 3 RYDBERG
AT: 3 CALCULATIONS
AT: 3 CONFIGURATION
AT: 3 DIABATIC
AT: 3 INTERACTION

1980 - 3 (2/5)

TT: 2 RYDBERG
AT: 2 GHZ
AT: 2 MICROWAVE
AT: 2 MILLIMETER
AT: 2 OBSERVED

1981 - 1 (3/3)

TT: 2 CROSS
TT: 2 H2
TT: 2 SECTIONS
AT: 3 H2
AT: 2 ANGULAR
AT: 2 CROSS
AT: 2 DISCUSSED

1981 - 2 (21/21)

TT:16 RYDBERG
TT:12 ATOMS
TT: 9 COLLISIONS
TT: 4 STARK
AT:14 ATOMS
AT:14 RYDBERG
AT: 8 CALCULATIONS
AT: 8 COLLISIONS

1981 - 3 (3/3)

TT: 3 ABSORPTION
TT: 3 EV
TT: 3 REGION
TT: 3 SPECTROSCOPY
AT: 3 ABSORPTION
AT: 3 CONSTANTS
AT: 3 EV
AT: 3 ROTATIONAL

1982 - 1 (21/21)

TT: 11 RYDBERG
TT: 7 STATES
TT: 6 ATOMS
TT: 4 EXCITED
AT: 11 STATES
AT: 8 ATOMS
AT: 8 RYDBERG
AT: 7 MEASURED

1983 - 1 (5/6)

TT: 5 RYDBERG
TT: 3 ATOMS
TT: 3 COLLISIONS
TT: 2 BROADENING
AT: 4 RYDBERG
AT: 3 ACCOUNT
AT: 3 ATOM
AT: 3 COLLISIONS

1983 - 2 (6/7)

TT: 5 RYDBERG
TT: 2 ATOMS
TT: 2 IONIZATION
TT: 2 NA
AT: 5 RYDBERG
AT: 4 FIELD
AT: 3 IONIZATION
AT: 3 N

1984 - 1 (6/6)

TT: 5 RYDBERG
TT: 3 STATES
TT: 2 ATOMS
TT: 2 COLLISIONS
AT: 6 RYDBERG
AT: 5 SHOWN
AT: 4 CROSS
AT: 4 SECTIONS

1985 - 1 (3/6)

TT: 2 ATOMS
TT: 2 COLLISIONS
AT: 3 N
AT: 2 ATOMS
AT: 2 COLLISIONS
AT: 2 INTERACTION

1985 - 2 (3/3)

TT: 3 RYDBERG
TT: 2 AUTOIONIZING
TT: 2 STATES
AT: 3 ION
AT: 3 RYDBERG
AT: 3 STATES
AT: 2 AUTOIONIZING

1985 - 3 (1/4)

Key :
1985 - 1 (3/6) : Co-citation cluster with index number 1985 - 1, and with 6 citing publica from which 3 are 'central' (i.e. these publications cite exclusively to one c

Title words (TT) and abstract words (AT); only the four most frequent terms are shown, excluc words occurring only once per cluster ('unique terms').

FIG. 3(b). Title and abstract words aggregated over central citing publications for Rydberg clusters 1976–1985.

terms (terms occurring only once) in order to exclude isolated aspects of research, and using only so called "central" publications for each cluster (i.e., publications that cite exclusively to one single cluster). Thus we are able to maximize differences between clusters.

As for indexing terms, the frequency of uncontrolled terms is much lower than the frequency of controlled terms for most of the "Rydberg" clusters (Fig. 3(a)). This probably reflects "text-specificity" of uncontrolled terms (84% of these terms are "unique"), as contrasted by the more unified character of controlled terms (43% "unique" terms) that are selected by the indexer from a subject thesaurus. Both these types of indexing terms are mostly phrases of two or more words, and are often mutually exclusive. For instance, the frequently used controlled term "atomic excited states" does not occur in the uncontrolled terms (Fig. 3(a)). This applies to a large extent also to the *individual* constituent words of indexing terms. For instance, from the controlled term "atomic excited states" the words "atomic" and "excited" are not present in the uncontrolled vocabulary (Fig. 3(a)). In contrast, single words from title or abstract are *not* mutually exclusive. For instance, the word "Rydberg" occurs frequently in both title and abstract (Fig. 3(b)). It is interesting to note that this word, which, according to the opinion of an expert, indicates the central concept of the specialty, is frequently used in title and abstract of citing publications, does also frequently occur in the uncontrolled terms but is *not* present in the controlled vocabulary. This word, although very useful as an "internal" indicator of work in the "Rydberg" specialty over almost the entire 10-year period under consideration, apparently has little value as an "external" subject indicator according to the *Physics Abstracts* indexing service.

A typical aspect of abstract words is the occurrence of rather general words indicating the empirical nature of the research involved, e.g., "found," "observed," "measured," "discussed," "account," and "calculations" (Fig. 3(b)). Such words are of course not very specific, but as a scientific specialty may not be empirically oriented, and as publications in specialties that are empirically oriented may still deal with conceptual problems, these words still are informative for other researchers as content descriptors.

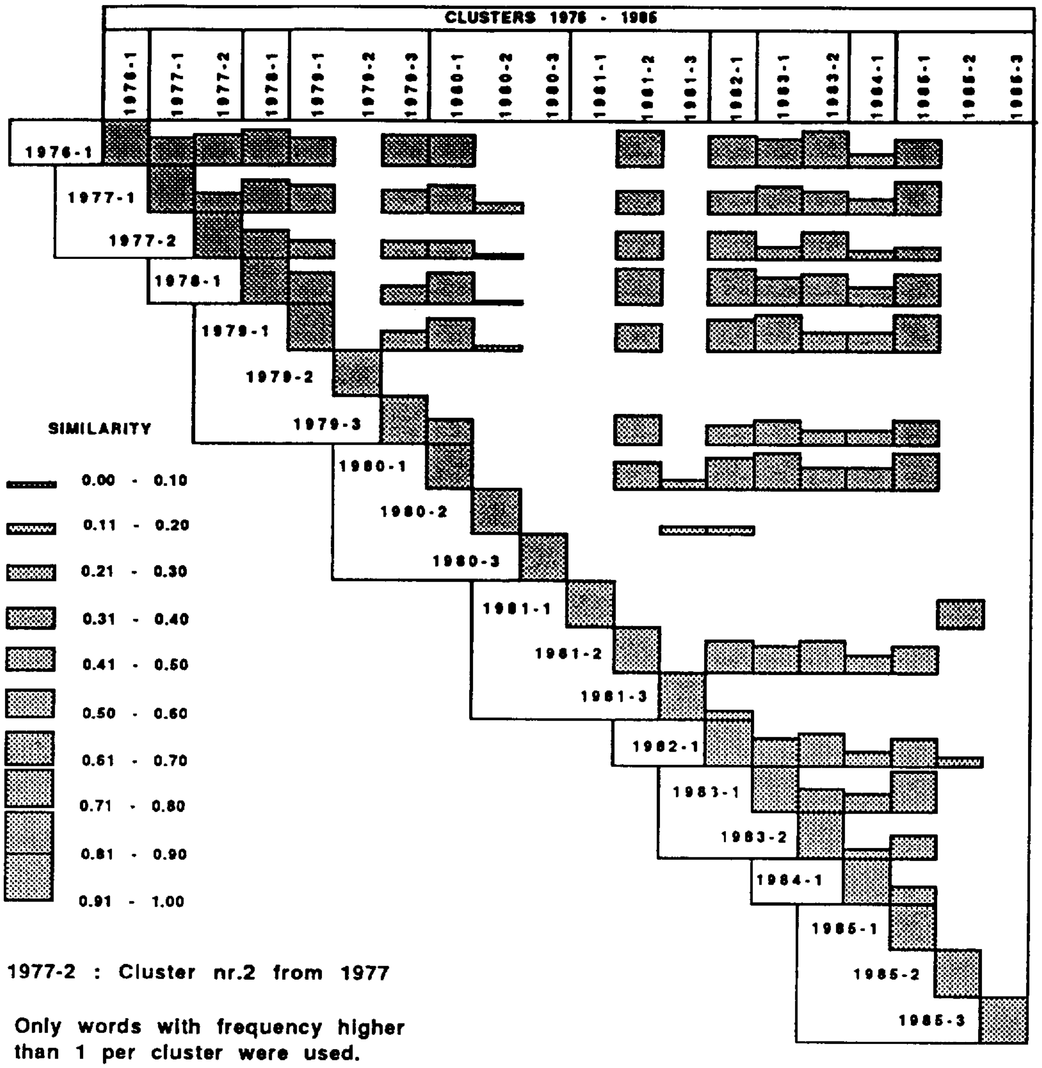
Word-Profile Similarity Relations between Clusters of the Same Year

The degree of word-profile similarity between clusters concerning *indexing terms* is given in Table 2. Word-profile similarity values concerning *title and abstract words* taken together are given in Table 3. As in this particular case-study all publications in the dataset are assumed to be related to one and the same "Rydberg" specialty, it may be asked why, for a number of years, still more than one co-citation cluster per year is formed. This yearly split-up may be related to incongruence in intellectual base only, or reflect also differences in topics of current research. Analysis of word-profile similarity concerning *indexing terms* indicates that different clusters of the years 1979 (1979-1 versus 1979-2, and 1979-2 versus 1979-3), 1980, 1981, and 1985 differ entirely in research topics, e.g., the clusters 1980-1, 1980-2, and 1980-3 have zero similarity one with another, i.e., no *indexing term* is shared by the word-profiles of these clusters (Table 2). However, if we inspect word-profile similarity concerning *title and abstract* words, all these clusters, except for the year 1985, are related, though at a relatively low level of similarity* (Table 3). For instance, clusters 1979-1, and 1979-2 have a (title and abstract) word-profile similarity to one another in the range 0.11–0.20. The words relating these clusters are, among others, "Rydberg," "interaction," "levels," "states," "electronic," "excitation," "ionization," and "quantum," all words indicating significant cognitive relations. Of the 1985 clusters, cluster 1985-3 is too small to have any word-profile at all, as it contains only one "central" publication. From the other two clusters, one (1985-1) deals with "atomic excited states" in relation to "collisions" and "interaction" of "atoms," and the other (1985-2) with "Rydberg" "states," "autoionisation," "photoionization," and "spectra of diatomic inorganic molecules" related to "H_2," and "hydrogen neutral molecules," (see Figs. 3(a) and 3(b)). Thus, except for the year 1985, in these cases authors and indexers apparently differ in opinion concerning the degree of content-relatedness of the cluster based publications groups, the authors indicating a weak linkage, whereas the indexers indicate isolation of these publication groups.

Word-profile similarity between different clusters of the years 1977, 1979 (1979-1 versus 1979-3), and 1983 is in the range of 0.31-0.40 for *indexing terms*, and between 0.11 and 0.40 for combined *title and abstract words* (see Figs. 3(a) and 3(b)). Thus, for these clusters both types of content words (indexing terms versus word from title and abstract) indicate a cognitive relation between the research topics involved in publications grouped by these clusters, though at a low to moderate similarity level (0.11 to 0.40). Content words differentiating between these clusters (see Figs. 3(a) and 3(b)) are, among others, "atomic fluorescence," "depopulation," "quantum," and "structure" (1977), "Rydberg *levels*," "Stark effect," "collisions," "states," and "absorption" (1979), "atom-atom collisions" and "atomic

*Low word-profile similarity between clusters may be the exclusive result of differences in length of their respective term lists, i.e., the terms of one cluster are fully included in the larger list of another cluster. In such cases, low similarity indicates a difference in scope concerning the spectra of problems, concepts, and methods that are involved in the current research related to these clusters.

TABLE 2. Word-profile similarity based on indexing terms (CT & UT) between RYDBERG clusters for the years 1976–1985.



spectral line breadth" in relation to "rubidium," versus "atomic beams" in relation to "Sodium" or "Na" (1983). Thus, in these cases author-viewpoints and indexer-viewpoints are congruent, both indicating cognitive relation between different clusters of the same year at a low to moderate level.

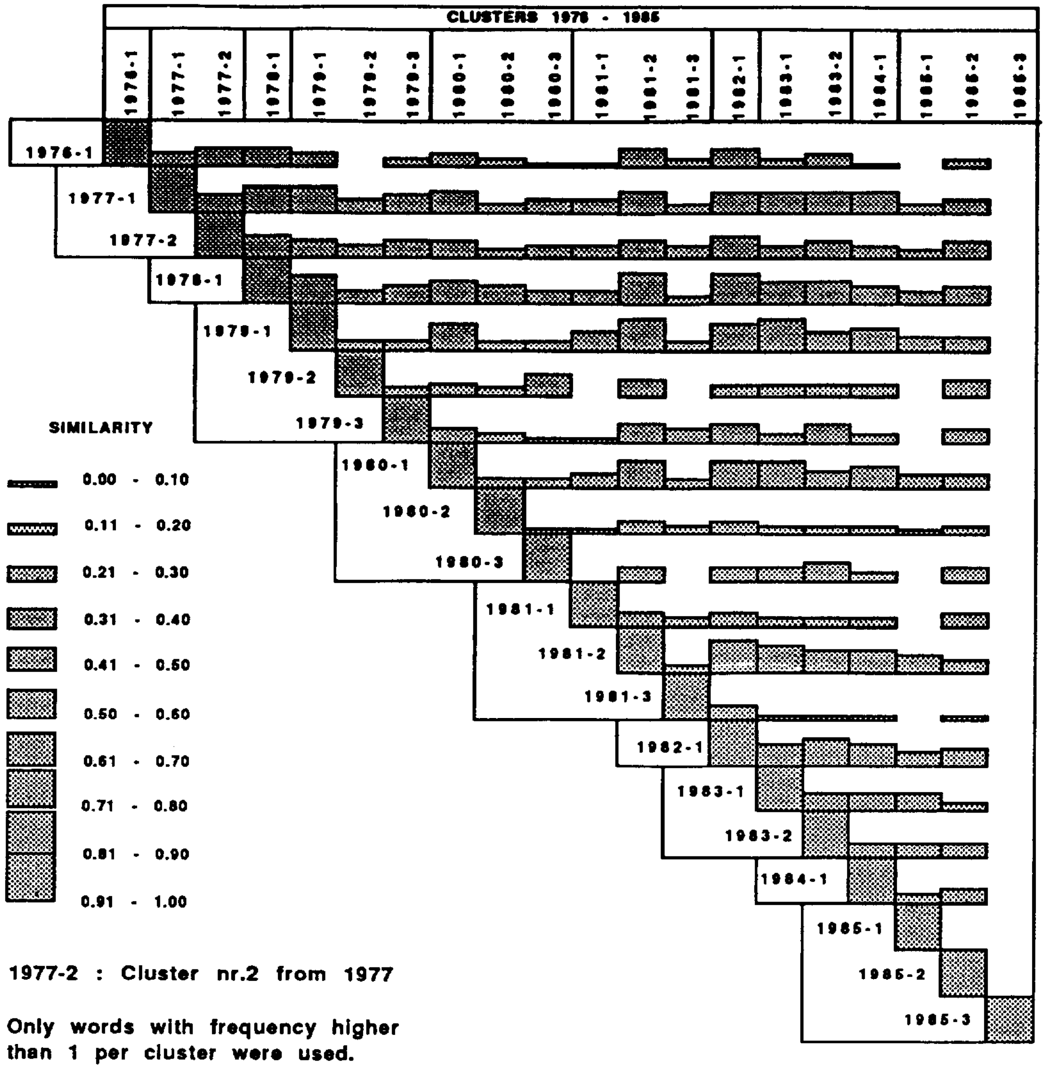
In general, we conclude from these results that the presence of more than one co-citation cluster in a number of years relates not only to low "consensus on intellectual base literature," but also to a considerable extent to differences in research topics studied in current research publications. These differences in research topics concern various experimental and/or theoretical aspects, the study of different atoms, the use of different (types) of instruments, etc.

Tracking "Rydberg" Publication Clusters through Time using Citations and Words

Focusing on the larger clusters (size in terms of the number of co-cited documents in clusters), continuity with respect to its intellectual base seems to dominate the specialty (as far as identified by the co-citation clustering) for the period under consideration (Fig. 2). This is indicated by the numerous links based on shared key-articles between clusters of different age. Also, a number of "sidelines," and isolated groups are indicated by clusters linked to few if any clusters of other years, e.g., clusters 1979-2 and 1980-3 indicate a "sideline," while clusters 1980-2 and 1981-3 (though linked to each other), and clusters 1985-2 and 1985-3 indicate isolated groups of publications. Cluster 1979-3 is a particular case, as it is strongly related by cited documents shared with older clusters, but lacks any relation to more recent clusters; all relations to more recent clusters are via cluster 1979-1, the largest cluster for this year.

Word-profile similarity relations based on indexing terms between clusters of different years indicate a high similarity in research topics between clusters of all years, not only between successive years, but also between clusters in a specific year and clusters of all other

TABLE 3. Word-profile similarity, based on title (TT) and abstract (AT) words, between RYDBERG clusters for the years 1976–1985.



years. However, not all clusters are involved, in fact, except for cluster 1979-3, clusters that are almost completely isolated with respect to clustered cited literature, are also isolated with respect to indexing terms (1979-2, 1980-2, 1980-3, 1981-3, 1985-2, 1985-3). Cluster 1979-3, has high similarity in indexing terms to clusters of all other years. Thus, the lack of relations to more recent clusters with respect to cited clustered documents, indicates a shift in "intellectual base" in the specialty (as given by cited papers), but *not* in topics of research, at least according to the opinion of indexers.

In general, more continuity seems present in research topics than in "intellectual base" (cited papers) if we look at indexing terms compared to cited literature involved in co-citation clusters (see Fig. 2 and Table 2). This clearly indicates that stability in topics of research, as judged by indexers, is not necessarily accompanied by an equal level of consensus on the importance or usefulness of earlier literature, as specified by authors. For instance, clusters 1976-1, 1977-1, and 1977-2 are related by citations to clusters of at most three other years, while relations concerning word-profiles based on indexing terms are present with all other years. An even more striking case is cluster 1979-3. This cluster is not related to any more recent cluster on the base of citations, with indexing terms however a relation is found to clusters of all more recent years (Fig. 2 and Table 2).

Results of word-profile similarity based on words from title and abstract of publications citing clusters, show even more continuity in topics of research than indicated by the above results for indexing terms (Table 3). The similarity values however are, on average, somewhat lower for title and abstract words than for indexing terms. Thus, authors indicate somewhat more, though weaker, linkages than indexers do.

A most interesting finding is that clusters isolated with respect to both citations *and* indexing terms, still appear to be related to almost all other clusters by title and abstract words. This suggests that these clusters still are cognitively related to the rest of the "Rydberg" specialty. Words linking these clusters to the other clusters are abstract words such as "Rydberg," "state," "excitation," "excited," and "state," as well as the title words "Rydberg," "state," and "absorption" (Fig. 3(b) clusters 1979-2, 1980-2, 1980-3, 1981-1, 1981-3, and 1985-2). The question arises, *why* these clusters are isolated with respect to indexing terms *and* clustered cited documents, while title and abstract words do indicate a relationship to the "Rydberg" specialty. Aspects of research topics differentiating these clusters from all other clusters (see Fig. 3(a) and 3(b)) are, among others, "caesium" (1979-2); "ab initio calculations," "potential energy curves," "molecular vibration," and "oxygen" (1980-2); "microwave," and "millimeter" (1980-3); "distributions," "resonances," "cross," "sections," and "H2" (1981-1); "molecular rotation," "isotopic," and "spectroscopy" (1981-3); "double," and "spectra of diatomic inorganic molecules" (cluster 1985-2). Two logically possible explanations for this discrepancy are: (1) title and abstract words do indicate subject similarity more directly, and, therefore, more adequately than indexing terms or shared cited key-articles do; or (2) these clusters indicate publications that are somewhat remote to the "Rydberg" specialty, as indicated by author citations and indexer opinion, but the authors try to establish a tighter link, in their titles and abstracts, maybe as it is a popular area of research. Assuming the cognitive coherence of the originally selected "Rydberg" publications in our dataset, the first option seems the most plausible. However, this discrepancy between citations and indexing terms, and title and abstract words as indicators of specialty membership is also related to differences in functioning of these "information carriers" in scientific research practices. Instead of trying to pick out the "best" indicator of specialty membership, it is perhaps more productive to try to gain some more insight in research practices by comparing the results of these different indicators.

Details of Dynamic Aspects

Not a single clustered cited document was cited over the whole 10-year period under consideration, while only a few cited documents occur in more than three successive years (i.e., GALLAGHER, TF, PHYSICAL REVIEW A, Volnr. 15, p. 1945, 1977, and DUCAS, TW, PHYS REV LETT, Volnr. 35, p. 366, 1975). No cited document occurs in more than six years (Table 4). Of the whole of clustered cited documents only 38% (37 out of 97 documents) occur in more than one year in a cluster. Newly introduced documents in clusters are generally young articles (46% are one year or younger, and 86% are three years or younger when appearing for the first time in a cluster, as far as concerns papers cited more than twice a year). From 1983 on, however, hardly any new articles are introduced. Thus there is a sudden change, in 1982, in the "normal" process of replacement of older by younger documents.

In contrast, a typical combination of (controlled and uncontrolled) indexing terms has been found that characterizes the specialty over almost the entire period under consideration, namely "Rydberg levels," "Rydberg states," or "Rydberg atoms" together with "atomic excited states," "atom-atom collisions," and some names of substances (Rb and Na) used in experiments (Table 5). The relative number of terms that occur in more than one year is much higher for controlled terms (49%) than for uncontrolled terms (15%). The few uncontrolled terms that do occur in more than one year, however, are very characteristic of the specialty: "Rydberg levels," "Rydberg states," and "Rydberg atoms" of "Na," "Rb," and "H2." Controlled terms, though less characteristic, establish a framework that makes the uncontrolled terms more understandable: "atomic excited states," "atom-atom collisions," and "ionization of atoms," indicate how the term "Rydberg" fits into more understandable terms and established concepts in physics.

Of all title words 36% (28 out of 78 different terms) occurs in more than one year in a cluster. Five words occur very frequently and give a good description of the general topic that relates the current work in the specialty of each year to the other current work of other years in the period 1976–1985: "Rydberg," "excited," "states," "atoms," and "collisions" (Table 6). A significantly larger number of abstract words occur in more than one year as compared to title words, both relative (48%) and in absolute numbers (169 out of 351). Also, the abstract words "Rydberg," "excited," "states," "atoms," and "collisions" occur very frequently in all or most of the years in the 1976–1985 period (Table 7). In addition, new words occur: "quantum," "cross," "sections," and "n." Further, a number of specific abstract words indicating the mathematically based experimental nature of the research in the specialty occur in many years: "experimental," "results," "observed," "measured," "calculations," "presented," and "discussed." No new indexing terms appear from 1983 on, and the same holds for title words, as far as concerns words that occur in at least two publications ("isolated words" excluded), while abstract words are not new from 1980 on. Thus, introduction of new aspects of research concentrates on the period before 1983, and diminished somewhat already since 1980.

Classification codes show a change, from 1980 on, in emphasis in the research from more experimental in the earlier years to more theoretical in later years, a pattern probably typical for experimental research in physics (Table 8).

In general, it seems a change occurred in the specialty from rapid succession of new empirical studies to more retrospective and theoretical work in later years.

TABLE 4. Time-plot of clustered cited documents[a] that appear in clusters from at least two different years.

CLUSTER CORE ARTICLES		PUBLICATION YEAR CITING PUBLICATIONS										
First author	Journal title, Volnr., First page, Pub. year	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	CUM. FREQ.
OMONT A	J PHYS-PARIS, 30, 1343, 1977				X	X		X	X	X		29
GALLAGHER TF	PHYSICAL REVIEW A, 15, 1945, 1977			X	X	X	X	X		X		28
FERMI E	NUOVO CIMENTO, 11, 157, 157, 1934				X	X		X		X	X	23
GALLAGHER TF	PHYSICAL REVIEW A, 16, 1098, 1977						X	X	X		X	18
JEYS TH	PHYSICAL REVIEW LETT, 44, 390, 1980						X	X	X			18
ALEKSEEV VA	SOVIET PHYS JETP, 22, 882, 1966				X	X		X	X			16
GALLAGHER TF	PHYSICAL REVIEW LETT, 35, 644, 1975		X		X	X	X					16
HICKMAN AP	PHYS REV A, 18, 1339, 1978				X	X	X		X	X		15
ZIMMERMAN ML	PHYSICAL REVIEW A, 20, 2251, 1979						X	X	X			15
DUCAS TW	PHYS REV LETT, 35, 366, 1975	X	X	X	X							14
HICKMAN AP	PHYS REV A, 19, 994, 1979					X	X		X	X		14
MATSUZAWA M	J PHYSICS B, 12,3743, 1979						X			X	X	13
STEBBINGS RF	PHYS REV A, 12, 1453, 1975	X		X	X							11
GOUNAND F	PHYSICAL REVIEW A, 15, 2212, 1977				X	X	X					11
HUGON M	J PHYS B, 12, 2707, 1979					X	X	X				11
DEPRUNELE E	J PHYSICS B, 12, 2511, 1979						X	X		X		10
GALLAGHER TF	PHYS REV LETT, 42, 835, 1979					X	X					10
HUGON M	J PHYS B ATOM MOL PH, 15, 2391, 1982								X	X	X	10
OLSON RE	PHYSICAL REVIEW A, 15, 631, 1977				X	X						9
BEITING E	J CHEM PHYS, 70, 3551, 1979					X	X	X				9
GALLAGHER TF	PHYSICAL REVIEW A, 19, 2161, 1979					X		X		X		9
KELLERT FG	J CHEM PHYS, 72, 3179, 1980						X	X				9
GALLAGHER TF	PHYSICAL REVIEW A, 17, 904, 1978				X		X					8
SMITH KA	PHYSICAL REVIEW LETT, 40, 1362, 1978				X		X					8
COOKE WE	PHYS REV A, 21, 588, 1980						X	X				8
MATSUZAWA M	J CHEM PHYS, 55, 2685, 1971				X					X		7
FABRE C	OPT COMMUN, 13, 393, 1975	X	X									7
FABRE C	PHYS REV A, 18, 229, 1978				X	X						7
GROSS M	PHYSICAL REVIEW LETT, 42, 835, 1979						X	X				7
BATES DR	PHIL T ROY SOC LON A, 242, 101, 1949			X	X							6
PERCIVAL IC	ADV ATOMIC MOLECULAR, 11, 1, 1975		X		X							6
LITTMAN MG	PHYS REV LETT, 36, 788, 1976			X	X							6
DEECH JS	J PHYS B, 10, L137,1977				X	X						6
DAMBURG RJ	J PHYSICS B, 12, 2637, 1979						X	X				6
GOUNAND F	J PHYS FRANCE, 40, 457, 1979						X	X				6
WEBER KH	OPTICS COMMUN, 31, 52, 1979							X	X			6
FIGGER H	OPT COMMUN, 33, 37, 1980					X	X					6
		1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	

[a]Clustered documents, cited more than twice a year. The fraction of these documents that occurs in at least two years is 38% over all years together (37 out of 97).

Conclusions

Co-citation analysis does not retrieve all publications contained in an expert's document file related to the "Rydberg" specialty, a specialty in atomic and molecular physics: a recall percentage of 58% over all years in the period 1976–1985 has been achieved (using a citation frequency of two citations per year and a co-citation strength threshold value equal to median of the annual distributions of these strength values).

In six of the 10 years under consideration more than one co-citation cluster was formed. Analysis of citations and indexing terms suggest that several cluster-based publication groups represent isolated aspects of research in the specialty, i.e., both intellectual base-literature and research topics seem to be different from the other publications of a particular year. Analysis of title and abstract words indicates, however, that these sets of publications are still cognitively related to the other publication groups of the specialty, but that their main emphasis is on topics that are somewhat peripheral to the specialty.

Analysis of linkages between clusters of different years points in the same direction. Many clusters of the "Rydberg" specialty are linked to clusters of other years by shared key-articles as well as by indexing terms and title and abstract words from publications citing these clusters (word-profiles). However, clusters isolated by cited key-articles are also isolated by indexing terms, though title and abstract words indicate a, relatively low, cognitive linkage to clusters of other years.

These differences between linkages based on indexing terms on the one hand, and title and abstract words on the other, relate to differences in view between indexers and authors on what research topics are involved in publications. Differences, though, are not very large, and are hypothesized to result from ignoring weak linkages by indexers, whereas authors may overemphasize

TABLE 5. Time-plot of nonunique indexing terms[a] occurring in citing publications of clusters from at least two different years.

INDEXING TERMS	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	CUM. FREQ.
CT: ATOMIC EXCITED STATES	X	X	X	X	X	X	X	X	X	X	73
CT: SODIUM	X	X	X	X	X	X	X	X			33
CT: ATOM-ATOM COLLISIONS			X	X	X	X	X	X	X	X	32
CT: RUBIDIUM	X	X		X	X	X	X	X	X		22
UT: NA	X	X	X			X	X	X			19
CT: ATOMIC SPECTRA		X	X		X	X	X				15
UT: RYDBERG STATES		X	X	X	X		X				15
CT: IONISATION OF ATOMS	X		X	X		X	X				13
CT: ATOMIC SPECT. LINEBR				X	X		X	X			12
CT: FIELD IONISATION		X	X	X		X	X				12
CT: STARK EFFECT	X	X		X		X					12
UT: RB				X	X	X	X		X		12
CT: ATOM-MOLECULE COLL.			X		X	X			X		10
UT: RYDBERG ATOMS						X	X	X			10
CT: HELIUM ATOMS				X		X	X				8
CT: MOLECULAR ELECTR. ST.					X	X	X				8
CT: TWOPHOTON SPECTRA					X	X	X				7
CT: ARGON				X		X	X				6
CT: SPECTRAL LINE SHIFT					X		X				6
CT: XENON			X			X					6
UT: RYDBERG LEVELS				X		X	X				6
CT: ATOMIC INELASTIC COL.						X			X		5
CT: HYDROGEN NEUTR. MOL.						X				X	5
CT: IONISATION OF MOLEC.							X			X	5
UT: H2						X				X	5
CT: ATOMIC FLUORESCENCE		X			X						4
CT: AUTOIONISATION							X			X	4
CT: OPTICAL PUMPING			X				X				4
CT: PHOTOIONISATION							X			X	4

[a]Controlled Terms (CT) and uncontrolled Terms (UT) from central citing publications, with frequency >1 per cluster. Fraction of these terms occurring in at least two years is 49% for CT's (23 out of 47) and 15% for UT's (6 out of 41).

these linkages somewhat (the "Rydberg" specialty was a popular area).

These results indicate that within the "Rydberg" specialty intellectual base literature is not shared by all researchers, but that there are several different "schools." Further, a number of cluster-based publication groups apparently have no strong relation to the main topics of the "Rydberg" specialty, but have a peripheral position.

These results do not simply reflect differences in specificity between cited references and indexing terms and keywords (or title and abstract words), as indicators of "specialty membership," a suggestion made by one of our referees. This is because different processes are involved behind citation and indexing practices or the practice of titling and abstracting. The former being related primarily to specification (for various reasons) of intellectual base, the latter directed primarily at specification (for various reasons) of research topics studied, as judged by indexers and authors. Indexers represent an information retrieval point of view, whereas authors represent an internal viewpoint. This difference in context has consequences for research topic description as authors probably want to gain a large audience, whereas indexers are more restrictive in order to optimize retrieval precision. It is, therefore, at least theoretically, more productive to compare results of these different data, and on this base try to gain some more insight in these processes and their relation.

The approach followed in this study, thus, enables an investigation of the continuity in the specialty, concerning its "intellectual base" in relation to its current research topics, as judged from different viewpoints. Continuity in "intellectual base" appeared to be at a lower level than continuity in topics of current research, the latter both to the opinion of authors as indexers. Continuity in topics of research was particularly reflected in title and abstract words, and thus by authors, but perhaps a little overemphasized.

TABLE 6. Time-plot of nonunique title words[a] occurring in citing publications of clusters from at least two different years.

TITLE WORDS	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	CUM. FREQ.
TT: RYDBERG		x	x	x	x	x	x	x	x	x	78
TT: STATES	x	x	x	x	x		x	x	x	x	43
TT: ATOMS			x	x	x	x	x	x	x	x	39
TT: COLLISIONS			x	x	x	x	x	x	x	x	27
TT: EXITED	x		x	x	x		x				14
TT: SODIUM	x	x			x	x	x				11
TT: HIGHLY	x			x	x		x				10
TT: IONIZATION				x		x	x	x			9
TT: LEVELS				x		x	x	x			9
TT: BROADENING					x		x	x			8
TT: ELECTRIC	x			x			x				8
TT: RUBIDIUM		x			x		x		x		8
TT: FIELD	x					x	x				7
TT: ABSORPTION				x		x					5
TT: COLLISIONAL					x		x				5
TT: LINES				x			x				5
TT: SPECTROCOPY		x				x					5
TT: STUDY					x	x					5
TT: TWO-PHOTON				x		x					5
TT: AUTOIONIZING							x			x	4
TT: BETWEEN				x	x						4
TT: COLLISION			x				x				4
TT: DOPPLER-FREE				x			x				4
TT: EXCITATION				x		x					4
TT: HIGH			x					x			4
TT: INTERACTION					x	x					4
TT: NA			x					x			4
TT: SHIFTS					x	x					4
	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	

[a]Title words (TT) occurring in central citing publications, with frequency >1 per cluster. Fraction of these words occurring in at least two years is 36% (28 out of 78).

It is concluded that in several years different "schools" are present in the specialty, that a series of interesting new contributions are made in course of time that changed the base for further research, but that general topics of research did not change in the specialty. Further, by inspection of timeplots of individual keyarticles and content-words, it was found that dynamics changed from rapid progression of new empirical studies in earlier years to more reflective and retrospective theoretical work in later years.

The present study starts with a co-citation clustering, and superimposes a word analysis. This means that no independent comparison was made between a structuring based on citations and one based on words. Such comparison has been made in a follow up study, the results of which are presented elsewhere (Braam et al., 1989).

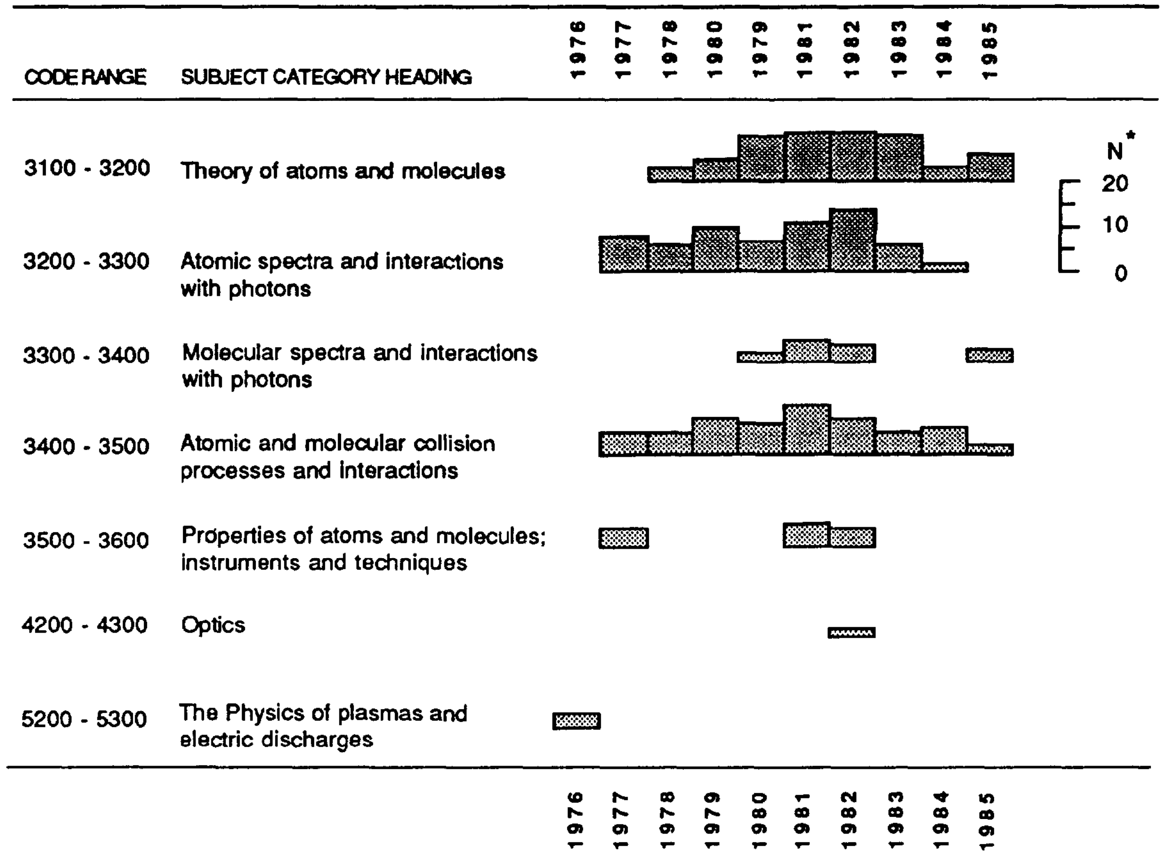
In general, the approach of combining co-citation and word analysis seems to provide a useful quantitative approach to study aspects of the dynamics of scientific research, as reflected in the formal literature at the level of research specialties. Compared to the exclusive use of citations, a more complete picture of dynamics of research can be established using such combined approach, including information on intellectual base as well as on judgments about topics of research from an "internal" (authors) and an "external" (indexers) viewpoint.

TABLE 7. Time-plot of nonunique abstract words[a] occurring in citing publications of clusters from at least two different years.

ABSTRACT WORDS	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	CUM. FREQ.
AT: RYDBERG		X	X	X	X	X	X	X	X	X	71
AT: STATES	X	X	X	X	X	X	X	X	X	X	64
AT: ATOMS			X	X	X	X	X	X	X	X	43
AT: CROSS		X	X	X	X	X	X	X	X		34
AT: QUANTUM		X	X	X	X	X	X	X			32
AT: STATE		X	X	X	X	X	X	X			32
AT: USING		X	X	X	X	X	X	X	X	X	32
AT: LEVELS		X	X	X	X	X	X	X			31
AT: N		X	X	X	X	X	X	X		X	30
AT: RESULTS		X	X	X	X	X	X	X		X	30
AT: SECTIONS		X	X	X	X	X	X	X	X		30
AT: OBSERVED	X	X	X	X	X	X	X	X			29
AT: FIELD		X	X	X	X	X	X	X			27
AT: INTERACTION			X	X	X	X	X	X	X	X	27
AT: COLLISIONS			X	X	X	X	X	X	X	X	26
AT: IONIZATION			X	X		X	X	X	X		24
AT: MEASURED				X	X	X	X	X	X		24
AT: EXCITED	X	X	X	X	X	X	X				23
AT: EXPERIMENTAL		X	X	X	X	X	X	X	X		23
AT: TRANSITIONS		X	X	X	X	X	X		X		21
AT: DISCUSSED			X	X	X	X	X	X			20
AT: USED		X		X		X	X				20
AT: ATOM		X		X		X	X	X			18
AT: ELECTRON		X	X	X		X		X	X		18
AT: PRESENTED			X	X	X	X	X				18
AT: AGREEMENT				X	X	X	X	X	X		17
AT: CALCULATIONS	X			X	X	X		X			17
AT: ELECTRIC	X		X	X		X	X	X			17
AT: ENERGY			X	X	X	X	X				17
AT: EXCITATION		X	X	X		X	X				17
AT: ATOMIC			X	X		X		X	X		16
AT: HIGH		X		X	X	X	X	X			16
AT: VALUES		X			X	X	X	X			16

[a]Abstract words (AT) occurring in central citing publications, with frequency >1 per cluster. Fraction of these words occurring in at least two years is 48% (169 out of 351). Only words with frequency above 15 over all years are shown.

TABLE 8. Time-plot of *Physics Abstracts* subject-classification code categories[a] occurring in Rydberg co-citation clusters for the years 1976–1985.



[a]Only central citing publications considered, different codes that are in the same range were counted as one for each publication, different clusters for each year taken together.

References

Braam, R. R., Moed, H. F., & Van Raan, A. F. J. (1988). Mapping of science: Critical elaboration and new approaches, a case study in agricultural related biochemistry. In L. Egghe and R. Rousseau (Eds.), *Informetrics 87/88* (pp. 15–22). Amsterdam: Elsevier Science Publishers. And also, in a more extended version: Braam, R. R., Moed, H. F., & Van Raan, A.F. J. (1987, November). *Mapping of science: Critical elaboration and new approaches, a case study in agricultural biochemistry*. Research Report to the Netherlands Advisory Council for Science Policy (RAWB), Leiden.

Braam, R. R., H. F. Moed, & van Raan, A. F. J. (1989). Comparison and combination of co-citation and co-word clustering. In A. F. J. van Raan, A. J. Nederhof, and H. F. Moed (Eds.), *Science and technology indicators: Their use in science policy and their role in science studies* (pp. 307–337). Leiden: DSWO Press, University of Leiden.

Braam, R. R., Moed, H. F., & Van Raan, A. F. J. (1990). Mapping of science by combined co-citation and word analysis: I structural aspects. *JASIS*,

De Solla Price, D. (1965). Networks of scientific papers. *Science*, *149*, 510–515.

Griffith, B. C., Small, H. G., Stonehill, J. A., & Dey, S. (1974). The structure of scientific literatures II: Toward a macro and microstructure for science. *Science Studies*, *4*, 339–365.

Jones, W. P., & Furnas, G.W. (1987). Pictures of Relevance: A Geometric Analysis of Similarity Measures. *Journal of the American Association for Information Science*, 38, 420–442.

Moed, H. F. (1988). The Use of Online Databases for Bibliometric Analysis. In L. Egghe and R. Rousseau (Eds.), *Informetrics 87/88* (pp. 133–146). Amsterdam: Elsevier Science Publishers.

Kuhn, T. S. (1970). *The structure of scientific revolutions*, Second Enlarged Edition. Chicago: University of Chicago Press.

Small, H. G., & Griffith, B. C. (1974). The structure of scientific literatures I: Identifying and graphing specialties. *Science Studies*, *4*, 17–40.

Small, H. G. (1977). A co-citation model of a scientific specialty: A longitudinal study of collagen research. *Social Studies of Science*, *7*, 139–166.

Small, H. G., & Crane, D. (1979). Specialties and disciplines in science and social science: An examination of their structure using citation indexes. *Scientometrics*, *1*, 445–461.

Small, H., & E. Sweeney (1985). Clustering the science citation index using co-citations: I. A comparison of methods. *Scientometrics*, *7*, 393–409.

Small, H., Sweeney, E., & Greenlee, E. (1985). Clustering the science citation index using co-citations: II. Mapping science. *Scientometrics*, *8*, 321–340.

Sneath, P. H. A., & Sokal, R. R. (1973). *Numerical taxonomy*. San Francisco: Freeman.