The Paraphrase Search Assistant: Terminological Feedback for Iterative Information Seeking

Peter G. Anick
Compaq Computer Corporation
40 Old Bolton Road
Stow, Massachusetts
978-496-9173
Peter.Anick@compaq.com

Suresh Tipirneni
Compaq Computer Corporation
40 Old Bolton Road
Stow, Massachusetts
978-496-8328
Suresh.Tipirneni@compaq.com

ABSTRACT

We present a new linguistic approach to the construction of terminological feedback for use in interactive query refinement. The method exploits the tendency for key domain concepts within result sets to participate in families of semantically related lexical compounds. We outline an algorithm for computing a ranked list of result set "themes" and describe a web application, the *Paraphrase Search Assistant*, designed to make use of the theme extraction algorithm to support a recognition-based, iterative information seeking dialog.

Keywords

Query reformulation, data visualization, terminological feedback

1. INTRODUCTION

Henninger and Belkin have argued that online information retrieval ought to be perceived as a problem solving process:

"Information retrieval systems must not only provide efficient retrieval, but must also support the user in describing a problem that s/he does not understand well. The process is not only one of providing a good query language, but also one of supporting an iterative dialog model. …The user is engaged in a problem-solving session in which the problem to be solved, that of finding relevant information, evolves and is refined through the process of seeing the results of the intermediate queries." [17]

A number of approaches have been proposed for assisting the information seeker in this problem solving process. These include techniques for helping the user more effectively explore the contents of a database or result list, via document clustering, visualization techniques, tilebars, category feedback and the like (e.g., [23], [16]). A separate set of techniques have been developed for assisting in query reformulation, including on-line thesauri, key term extraction, and relevance feedback (e.g., [28], [30], [15]). There is some evidence from recent studies comparing "opaque" and "penetrable" relevance feedback suggesting that allowing end-users to interact directly with terminological feedback improves not only actual (measured) retrieval performance but also perceived performance, trust in the system, and subjective usability [19].

Unfortunately, choosing an appropriate set of terms to present to a user as suggestions for query refinement remains a difficult task. Thesauri do not exist for most databases and statistical approaches such as relevance feedback require users to assess the relevance of articles in a result list, a task many find difficult [11]. What is more, there are typically hundreds of terms that are potentially relevant to an information need; without some structure imposed upon the terms, it would be impossible for a user to inspect more than a handful.

In this paper, we present a new approach to the generation and presentation of terminological feedback, based on the hypothesis that key concepts within a document collection are more likely than other terms to participate in a wide variety of semantically related lexical compounds. We show how this notion of "lexical dispersion" can be exploited to generate *thematically organized* terminological feedback from result lists at query-time. Such structured feedback, by drawing attention to the implicit themes and relationships underlying families of domain-specific lexical collocations, facilitates the process of iterative information seeking in a number of important ways:

- It helps the user recognize aspects and dimensions of the problem space
- It helps the user assess the results of a search
- It provides a source of terminology for recognition-driven query reformulation.

SIGIR '99 8/99 Berkley, CA USA

In section 2, we motivate the "lexical dispersion hypothesis" and describe the algorithm we have developed to exploit lexical dispersion for the purpose of identifying and structuring key terminology. In section 3, we present the web-based interface built for the *Paraphrase Search Assistant*, illustrating how search tactics making use of the feedback are integrated into the system/user dialog. In section 4, we discuss our findings from a pilot release of *Paraphrase* for searching intranet web sites. Section 5 considers the advantages and drawbacks of this approach. In section 6, we review related work and in section 7 we present our conclusions and directions for future research.

2. THE LEXICAL DISPERSION HYPOTHESIS

2.1 Linguistic motivation

In English, as in other languages, new concepts are often expressed not as new single words, but as concatenations of existing nouns and adjectives. This is especially noticeable in technical language, where long chains of nouns are not uncommon ("database management system", "byte code interpreter") and it appears to be one of the primary characteristics of sublanguages such as medical briefs and naval messages, which have domain specific grammars for combining terms into terse expressions of complex concepts (e.g., [18], [20]). Multi-word terms also permeate everyday language. Noun compounds are regularly used to encode ontological relationships -- an "oak tree", for example, **isa** tree. They may encode many other kinds of relationships as well [13]. In "tree rings", rings are a property of the tree. In "tree roots", the roots are a part of the tree. One would therefore expect that a set of documents dealing with trees would contain many different collocations containing the word "tree", (e.g., tree diseases, tree bark, tree sap, tree roots, pine tree, coniferous tree) since such compounds linguistically serve to identify subordinate categories, attributes, and other relationships within the domain of trees. Similarly, other salient concepts related to trees would likely be expressed as nominal compounds within such a collection. For example, we might find compounds incorporating the concept "forest" (rain forest, forest fire) or specializations of types of trees (white birch, silver birch).

These observations suggest that a statistical analysis of the terms appearing within noun compounds in a document collection might serve to expose some of the main topical threads running through the collection. Specifically, we define the *lexical dispersion hypothesis* as the proposition that a word's *lexical dispersion* – the number of *different* compounds that a word appears in within a given document set - can be used as a diagnostic for automatically identifying key concepts, or "facets" of that document set.

2.2 Corpus-wide lexical dispersion

We have tested the lexical dispersion hypothesis on a number of text corpora, including the Financial Times corpus used in TREC, a computer troubleshooting web site, a corporate newsfeed, and the results of a world wide web query on "jazz music." For each corpus, we constructed a list of all sequences of 2-4 terms matching the pattern { *?adjective noun+* } and computed the dispersion of each word appearing within the compounds. In every case, we found lexical dispersion to be extremely high (in the thousands) for those terms with greatest dispersion and moderately high (> ten) for thousands of other words. As expected, the terms with highest lexical dispersion did tend to represent key concepts or themes of their respective domains. The top three terms for the Financial Times database, for example, were the following (listed with their dispersion and a sample of the compounds containing each term):

market [16,368] (stock market, bond market, gilt market, market share)

group [16,220] (insurance group, industrial group, Britten Group, Pegasus Group)

company [14,572] (public company, car company, company formation, company plan)

For the jazz music database, the top terms included:

music [20,826] (world music, sheet music, music director, Zulu choral music)

jazz [11,403] (dixieland jazz, jazz ensemble, Antibes Jazz Festival, Yerba Buena Jazz Band)

band [8,937] (swing band, band leader, Robert Cray Band, miserable garage rock band)

As might be expected, the terms with the highest dispersion also tended to have very high document frequencies within their respective collections; therefore, they had extremely low discrimination value as search terms when used by themselves. Their compounds, however, tended to be not only highly germane to the subject area but also much more discriminatory as search terms. These compounds provided a potpourri of attributes, specializations, related concepts, etc., for the more generic terms they contained.

The most encouraging result of this study of lexical dispersion was the finding that across a variety of databases, the number of terms with even a moderate degree of dispersion was very high. This raised the possibility that dispersion levels for words might be sufficiently high even within relatively small (but focused) subsets of a collection to serve to identify key themes within result lists returned for users' queries. We tested this using dozens of queries, broad and narrow, on each of the collections and found that, like the databases as a whole, result sets almost

invariably contained families of lexical compounds reflecting categories relevant to the query topic. During the course of these experiments, we made a number of pragmatic refinements to the term extraction algorithm, which we describe in the next section.

2.3 Term extraction from result lists

We will refer to the thematic terms identified by our algorithm as *facets* and to their associated compounds as *values*.[1] While the lexical dispersion hypothesis, as presented above, provides a simple criterion for capturing such terms, our actual implementation has evolved to take into account a number of practical factors, many of which we are continuing to explore.

The algorithm consists of two stages. At index time, each document is parsed to extract instances of lexical compounds. Non-capitalized head nouns are reduced to their non-inflected form and any compounds containing heads or modifiers from a generic list of "noise" terms are eliminated.[2] At query-time, the compounds appearing in the first n documents of a ranked result list are used to compute the lexical dispersion of each term occurring within these compounds. These terms are then sorted by dispersion and the top m (in our case, 40) are added to a list of candidate facets.

From this set of candidates, we wish to choose approximately 15 for display to the user as feedback categories. We have experimented with a number of strategies for reranking the candidate facets. The simplest strategy is to use the raw dispersion by which the terms are already ranked. However, since dispersion is a measure of the number of *different* compounds a term appears in, it is possible for one or two long articles to contribute to the dispersion of a term that otherwise occurs relatively infrequently across the result set. One way to favor terms that appear in many different articles is to rerank the list using each term's *spread*, the number of result list articles in which the term appears in at least one compound. Our current implementation uses a variant of *spread*-based ranking which ranks terms by spread if their dispersion exceeds 5 compounds, then ranks the remainder by their dispersion. This variant tends to favor facets which apply across the result set, while demoting those with small dispersions, regardless of their frequency of occurrence.

Another parameter that can greatly affect the quality of the extracted facets is the number of result set documents used to compute them. This number needs to be large enough that the key concepts are likely to be reflected statistically within lexical compounds in the subset. On the other hand, for many natural language queries (in which terms are ORed by default), the density of truly relevant articles tends to fall off significantly after the top 20 or 30 articles. For example, the natural language query "cold fusion", applied to our Financial Times database, produces over 3000 hits, of which only ten actually contain the phrase "cold fusion". In this case, facets generated from a result set of even 100 documents are consequently more likely to reflect the more prevalent concepts of "ice cream" and "cold war", rather than the intended topic of "cold fusion." Fortunately, we have found that the set of facets generated from the top 50 documents for a query is often similar to that generated from a much larger set in those cases where there are many documents that truly match a query. As a result, we have implemented a two-phase approach to extracting facets and their values. The top 50 documents are used to generate a set of facets. Then the top 100 documents are used to extract a more extensive set of compounds for each previously chosen facet.

2.4 Examples

In this section we show some examples to illustrate the nature of the facets selected by the algorithm described above. The first example comes from the query "neural network" applied to a corporate newsfeed that includes patent abstracts. The top ten facets produced are:

corporation, neural, network, system, data, time, sequence, rule, architecture, vector

The *system* facet includes such values as handwriting recognition system, cooling system, and fraud management system, types of applications to which neural networks have been applied. Similarly, *data* includes EEG data, customer data, and historical data, providing yet another cut for classifying neural networks. Values for the *corporation* facet include Seiko Epson Corporation, Xerox Corporation, and Digital Equipment Corporation, all patent assignees. This facet illustrates how naming conventions (e.g., company names using the term "Corporation") help the algorithm capture some types of proper names[3]. Other facets, such as *neural* and *network*, provide attributes and types of neural nets (weighted sum network, predictive network) as well as some "noise" compounds that really have more to do with other senses of the individual terms (neural tissue, cable network).

Our second example comes from the query "water pollution" applied to the Financial Times database:

[1] Our use of these terms is borrowed from Boolean information retrieval systems, in which clusters of terminological variants and specializations within a query or domain model are often known as "facets" of the query or domain.

[2] These include quantitative nouns and adjectives, such as cardinal or ordinal numbers, words like "many", "some", "amount", temporal nouns like "year", and qualitative adjectives, such as "significant" and "reasonable."

[3] Personal names (without titles) are one notable exception.

authority, company, environment, environmental, group, management, pollution, public, waste, water

As a group, these terms capture a wide range of dimensions relevant to the topic - companies involved (Ecuadorian oil company, Icelandic company), pollutants (organic waste, nuclear waste), environmental issues (environmental protection, environmental impact, environmental regulation), relevant authorities (water authority, National Rivers Authority), management issues (waste management, river management), aspects relating to the public (public health issue, public pressure), and types of pollution (agricultural pollution, mine water pollution). By scanning these facets and values, a searcher can get a good feel for the issues covered within the Financial Times articles.

These two examples are fairly representative of the quality of the feedback generated by our algorithm. There is a fair degree of noise (e.g., phrases that are off-topic or misparsed) as well as significant gaps (e.g., categories whose members are primarily single word terms), compared to categories that might be constructed by hand. Nevertheless, the majority of the automatically extracted "facets" do represent useful dimensions for (1) reasoning about the domain topics(s), (2) assessing the contents of the result set, and (3) choosing a strategy for refining the search.

3. THE PARAPHRASE SEARCH ASSISTANT

The *Paraphrase Search Assistant* is a web-based application designed to make use of the term extraction algorithm described above to support a recognition-based, iterative searching strategy. The *Paraphrase Search Assistant* is implemented on top of the AltaVista search engine [1], which processes natural language queries and returns ranked result lists.

3.1 User interface

The interface, shown in figure 1, consists of a single web page divided into three frames. The top frame contains a text box for submitting queries. The large frame in the lower right is used for displaying the result list. Additionally, whenever a user enters a query, facets are displayed as an alphabetized list of select boxes in the lower left hand frame.

Figure 1 shows the *Paraphrase* screen after the user has entered the query "renewable resources". At this point, the user may "browse" the facets by clicking on the select boxes. These expand to reveal up to 40 phrases containing the facet term, sorted by their frequency of occurrence in the first 100 articles of the result set. Unlike most previous approaches to interactive query reformulation (e.g., [12], [6]), in which a user is asked to choose all relevant terms for mass inclusion in a query, the model of iterative search presented by the *Paraphrase* interface is more in keeping with the interpretation of a facet's values as a set of alternatives for exploring the document space. We wanted to make it easy for users to iterate through a set of alternative specializations, one at a time, while maintaining their original query context.

To do so, we divided the search expression into two distinct conceptual components – a primary "search topic" and a "focus." The user's initial search expression is what defines the search topic; this is the expression for which the set of facets is computed and presented, as described above. When the user clicks on a facet value, it is added to the focus. Immediately, the system constructs a new query from the concatenation of (1) the original search topic terms, (2) the full facet phrase (as a single contiguous term), and (3) the individual components of the facet phrase. Due to AltaVista's ranking algorithm, which grants rarer terms higher weights, documents containing the full phrase are more likely to bubble to the top of the result list. By also including the components of the phrase in the query, we enhance recall as well; component terms need not appear adjacently to contribute to the result.

Figure 2 shows the screen after the user has selected the term "tidal power" from the *power* select box. Note that the original search topic in the top frame's text box remains unchanged but "tidal power" is now listed as the focus above the result list. Each time the user selects a different facet value (e.g., "wind power") as the focus, the previous focus value is superseded and a new result set is computed. The facets, however, are not recomputed, leaving the user with a stable context from which to choose other query refinements. In this way, the user can conveniently explore the database by iterating through facet phrases one at a time, reshuffling the ranked result list with each mouse click.

3.2 Other search tactics

Besides the ability to "narrow" a query expression through the selection of a facet value, the *Paraphrase Search Assistant* provides direct support for two other search tactics [4]. These tactics are presented within a "suggestions" section offered at the bottom of the result list frame.

First, a user may choose to do a new search containing only the focus term. Clicking on this option has the effect of *replacing* the current search topic with the current focus term. The new query is run and a new set of facets is computed from the results. This option is particularly useful in those cases where a user may want to shift contexts completely based on the recognition of a new topic of interest within the terminological feedback. Secondly, a user may choose to actually append the focus term to the current query expression. Selecting this option places the combined query into the search topic box and recomputes the

facets for the combined query, thereby giving the user the opportunity to further refine the combined query (i.e., *drill down*) by the selection of additional facet values.

4. USER BEHAVIOR

The *Paraphrase Search Assistant* is currently available to Compaq intranet users over several intranet web sites, including a 175,000 article computer troubleshooting database (STARS), a newsfeed, and several corporate information sites (NSIS). No training is provided beyond the "suggestions" appearing on the web page itself. By capturing sequences of user selections in session logs, we have begun to study how *Paraphrase*'s features are utilized in actual user-driven search situations.

From a sample of 712 sessions involving query reformulation[4] captured over a three-month period, the logs indicate that facet expansion (viewing the terms inside a facet's select box) was employed in roughly 70% of the sessions, while feedback terms were actually selected for refinement in roughly 60% of sessions. These figures indicate a fairly high uptake, considering the non-laboratory setting and lack of any explicit training on the interface. Nearly a quarter of the sessions involved three or more facet selections.

Users conducted some portions of their searches by iteratively selecting different values for the same facet category. Of those cases in which a facet selection followed another facet selection within the same query context, we find that NSIS users selected a value from the same category 67% of the time, while STARS users selected from the same category 33% of the time. Around 20% of facet value selections were for a facet whose name was one of the query terms. We estimate that for 70% of user queries, one of the query terms appears in the displayed facet list. Given this likelihood, the 20% uptake is surprisingly low, showing that users are not necessarily looking for phrases that are directly related (as attributes or subtypes) to their query terms.

The logs show a preference for choosing facet values for which the facet name plays the role of the head in the chosen phrase (60% of selections); however, the data confirms that both roles are important to users. The percentages may reflect the actual ratio of head and modifier roles in the data, which we have not measured.

Many users, such as those searching the STARS computer troubleshooting database, were performing high precision searches under time pressure (to resolve a call). A relatively low ratio of fewer than three facet expansions for each facet value selected suggests that users tended to hone in quickly on those facets relevant to their needs. The *drill down* and *replace* operations were used, but relatively rarely, in only 8% and 7% of sessions, respectively.

A transcript of a typical session is shown below, illustrating how the facets help to map the user's vocabulary to alternative expressions found in the database. The user, after reviewing the initial result list, expands the facet corresponding to the query term itself ("banking") and selects a phrase. The sixth article in the new result list is examined. The user then expands a different facet ("services") and finds the phrase "financial services". The top article is selected from the new result list, terminating the session.

Query: banking
Facet expansion: banking
Term selection: "banking industry"
Document selection: #6
Facet expansion: services
Term selection: "financial services"
Document selection: #1

The following sequence is an excerpt from a long session in which the user, apparently researching drivers for the DE500, uses facets repeatedly to focus the result list on specific subtopics while keeping the initial query constant.

Query: de500
Facet expansion: driver
Term selection: "de500 driver"
Facet expansion: version
Term selection: "version"
Facet expansion: adapter
Term selection: "de500 ethernet adapter"
Document selection: #7
Facet expansion: drivers
Term selection: "de500 drivers"
Document selection: #14
Facet expansion: network
Term selection: "dc21x4 network card driver"
Document selection: #14

5. DISCUSSION

The model of retrieval supported by the *Paraphrase Search Assistant* is consistent with the problem solving approach argued for by Henninger and Belkin [17]. Information seeking is conducted as a sequence of small focused searches, each one building upon knowledge gained from the previous results. *Paraphrase* contributes to the dialog both by assisting the user in the interpretation of the results and by providing recognition-based tactics for proceeding with the next step of the search. As

[4] We define a session as a temporally contiguous sequence of search operations following the input of an initial query expression. Query reformulation is defined as any modification to the initial query, including the selection of a facet.

one library information specialist commented after experimenting with the system, "The themes (facets) give people the context of what's there. It gives them prompts. In most search engines, you don't get that. You just guess."

Certainly, an attractive feature of using lexical dispersion to capture relevant terminology is that the terminology so retrieved is by default organized into categories and these categories have recognizable names. The *Paraphrase* interface can display up to 40 values for each of the 15 facets presented to the user, yielding a possible total of 600 feedback terms. It would be daunting for a user to comb through an unstructured list containing so many items. While it is true that the "natural language" categories produced by our algorithm are rarely conceptually homogeneous sets, users are familiar enough with how phrases are constructed to understand why particular items appear in each list. Indeed, it is often possible to guess the meaning of an unknown term or proper name by its collocations within the phrase. We have found that browsing the categorized feedback is conducive to the serendipitous discovery of unanticipated relationships between query concepts and feedback terms.

There are drawbacks as well. Since feedback terms are generated by an analysis of the result set, if the result set does not contain meaningful clusters of documents, the phrasal feedback is likely to be equally diffuse. In such cases, phrases associated with the query terms themselves are usually the only recourse, as they tend to reveal the set of possible interpretations for the query terms and therefore provide the best opportunity for focusing the query[5]. Another drawback is the fact that certain kinds of terms relevant to a discourse topic may simply not often appear in lexical compounds. A query on "species extinction" is more likely to return a list of animal subtypes (*wild animal*, *exotic animal*, *rare animal*) than a list of specific endangered species.

Results from our pilot study have been very encouraging, in that searchers appear able to grasp and utilize the phrasal feedback without the need for explicit training on the use of the interface.

6. RELATED WORK

The PLEXUS [31], OAKASSIST [7], and HIBROWSE [22] interfaces demonstrated how manually constructed knowledge organized into facets could be employed in order to help users explore a problem space and formulate their queries. The use of co-occurrence statistics to automate the discovery of semantic relationships among terms has a long history in Information Retrieval [27] and Computational Linguistics [9]. Ruge [25] and Grefenstette [14] applied corpus linguistic techniques to the automatic construction of thesauri for use in information retrieval. Strzalkowski [29] utilized a broader notion of dispersion in a formula to compare the specificity of semantically related terms for the automatic construction of a lexical domain map. Nakagawa [21] used the heuristic that simple Japanese nouns that are part of many compound nouns make good index words to automate the extraction of index terms for computer manuals. The REALIST hyperterm system [24] and AI-STARS user interface [2] made on-line databases of phrases available to end-users for interactive query reformulation. Bates [5] integrated a thesaurus into a search interface to make clusters of related terms available when forming Boolean queries. Cooper and Byrd's Lexical Navigation system [10] provided extensive interactive browsing and query formulation capabilities through a network of terminology derived from automatic corpus analysis. Bruza's Hyperindex Search Engine [8] allowed users to browse through a lattice of search terms composed of noun phrases of various levels of specificity. Anick and Vaithyanathan's interface to a clustered document database [3] employed cluster descriptor terms as "facets" which expanded to sets of phrases containing the facet terms.

[5] This is precisely the motivation behind Bruza's "hyperindex" interface [8].

7. CONCLUSIONS

We have presented a new approach to the construction of terminological feedback for use in iterative query refinement, exploiting the tendency for key domain concepts within result lists to participate in families of lexical compounds. By presenting these families as groups of specialized topics relating to more general facets, the *Paraphrase Search Assistant* in effect creates a dynamic browsable table of contents from which the user may better assess the current search results as well as select the next step in his/her search. Query logs from new users of the *Paraphrase* system show that searchers are willing and able to make use of the feedback so presented without the need for explicit training on the use of the interface.

Future work will be directed at refining the term extraction algorithm to improve the quality of feedback across a range of result set conditions. We will be continuing to monitor user behavior to better understand how to facilitate search tactics appropriate to different database domains and user needs. We are also exploring the possibility of using syntactic contexts other than noun phrases to ferret out key terms and relationships that do not appear in noun compounds.

8. REFERENCES

[1] *The AltaVista Search Revolution*. Digital Equipment Corp., 1996.

[2] Anick, Peter G., Jeffrey D. Brennan, Rex. A. Flynn, David R. Hanssen, Bryan Alvey, and Jeffrey M. Robbins, "A Direct Manipulation Interface for Boolean Information Retrieval via Natural Language Query", in *Proceedings of SIGIR'90*, pp. 135-150, 1990.

[3] Anick, Peter and Shiv Vaithyanathan, "Exploiting Clustering and Phrases for Context-based Information Retrieval", in *Proceedings of SIGIR'97*, pp.314-323, 1997.

[4] Bates, Marcia J., "Information Search Tactics", *JASIS*, 30:4, pp. 205-214, 1979.

[5] Bates, Marcia J., "Design for a Subject Search Interface and Online Thesaurus for a Very Large Records Management Database", in *Proceedings of JASIS'90*, pp. 20-28, 1990.

[6] Beaulieu, Micheline, Thien Do, Alex Payne and Susan Jones, "ENQUIRE Okapi Project", *British Library Research and Innovation Report 17*, Jan. 1997.

[7] Borgman, C. L., D. O. Case, and C. T. Meadow, "The Design and Evaluation of a Front-End Interface for Energy Researchers", *JASIS*, vol. 40, pp. 99-109, 1989.

[8] Bruza, P.D. and S. Dennis, "Query Reformulation on the Internet: Empirical Data and the Hyperindex Search Engine", in *RIAO'97*, pp. 488-499, 1997.

[9] Church, Kenneth Ward and Patrick Hanks, "Word Association Norms, Mutual Information, and Lexicography", *Computational Linguistics*, 16:1,1990, pp. 22-29.

[10] Cooper, James W. and Roy J. Byrd, "Lexical navigation: Visually prompted query expansion and refinement", in *Digital Libraries'97*.

[11] Croft, W. Bruce. "What Do People Want from Information Retrieval? (The Top 10 Research Issues for Companies that Use and Sell IR Systems)". *D-Lib Magazine*, Nov, 1995.

[12] Efthimiadis, Efthimis N., "End-users' Understanding of Thesaural Knowledge Structures in Interactive Query Expansion", *Advances in Knowledge Organization*, vol. 4, pp. 295-303, 1994.

[13] Finin, Timothy W., "The Interpretation of Nominal Compounds in Discourse", *Technical Report MS-CIS-1982-3*, University of Pennsylvania, 1982.

[14] Grefenstette, Gregory, "Use of Syntactic Context to Produce Term Association Lists for Text Retrieval", in *Proceedings of SIGIR'92*, pp.89-97, 1992.

[15] Harman, Donna, "Towards Interactive Query Expansion", in *Proceedings of SIGIR'88*, pp. 321-331, 1988.

[16] Hearst, Marti, "Using Categories to Provide Context for Full-Text Retrieval Results", in *Proceedings of RIAO'94*, pp.115-130, 1994.

[17] Henninger, Scott and Nicholas Belkin, "Interface Issues and Interaction Strategies for Information Retrieval Systems", in *Proceedings of the Human Factors in Computing Systems Conference (CHI'96)*, ACM Press, New York, 1996.

[18] Jacquemin, Christian and Jean Royaute, "Retrieving Terms and their Variants in a Lexicalized Unification-Based Framework", in *Proceedings of SIGIR'94*, pp. 132-141, 1994.

[19] Koenemann, J., and Belkin, N. J., "A Case for Interaction: a Study of Interactive Information Retrieval Behavior and Effectiveness", in *Proceedings of the Human Factors in Computing Systems Conference (CHI'96)*, ACM Press, New York, 1996.

[20] Marsh, Elaine, "A Computational Analysis of Complex Noun Phrases in Navy Messages", in *Proceedings of COLING'84*, pp. 505-508, 1984.

[21] Nakagawa, Hiroshi, "Extraction of Index Words from Manuals", in *Proceedings of RIAO'97*, pp.598-611.

[22] Pollitt, Steven A. The key role of classification and indexing in view-based searching. In Proceedings of IFLA, Copenhagen, Sept. 1997.

[23] Rao, Ramana et al, "Rich Interaction in the Digital Library", *Communications of the ACM*, 38:4, pp. 29-39, 1995.

[24] Ruge, G, C. Schwarz and G. Thurmair, "A Hyperterm System Based on Natural Language Processing", *International Forum on Information and Documentation*, 15(3):3-8, 1990.

[25] Ruge, Gerda, "Experiments on Linguistically based Term Associations", in *Proceedings of RIAO'91*, pp. 528-545, 1991.

[26] Salton, Gerard, *Automatic Text Processing – The Transformation, Analysis and Retrieval of Information by Computer*, Addison-Wesley, Reading, MA, 1989.

[27] Spark Jones, Karen, *Automatic Keyword Classification for Information Retrieval*, Butterworths, London, 1971.

[28] Srinivasan, Padmini, "Query Expansion and MEDLINE", *Information Processing & Management*, vol. 32, no. 4, pp. 431-443, 1996.

[29] Strzalkowsi, T., "Building a Lexical Domain Map from Text Corpora", in *Proceedings of COLING'94*, pp. 604-610, 1994.

[30] Tseng, Yuen-Hsien, "Multilingual Keyword Extraction for Term Suggestion", In *Proceedings of SIGIR'98*, pp. 377-378, 1998.

[31] Vickery, A and H. M. Brooks, "PLEXUS – The Expert System for Referral", *Information Processing & Management*, 23(2):99-117, 1987.

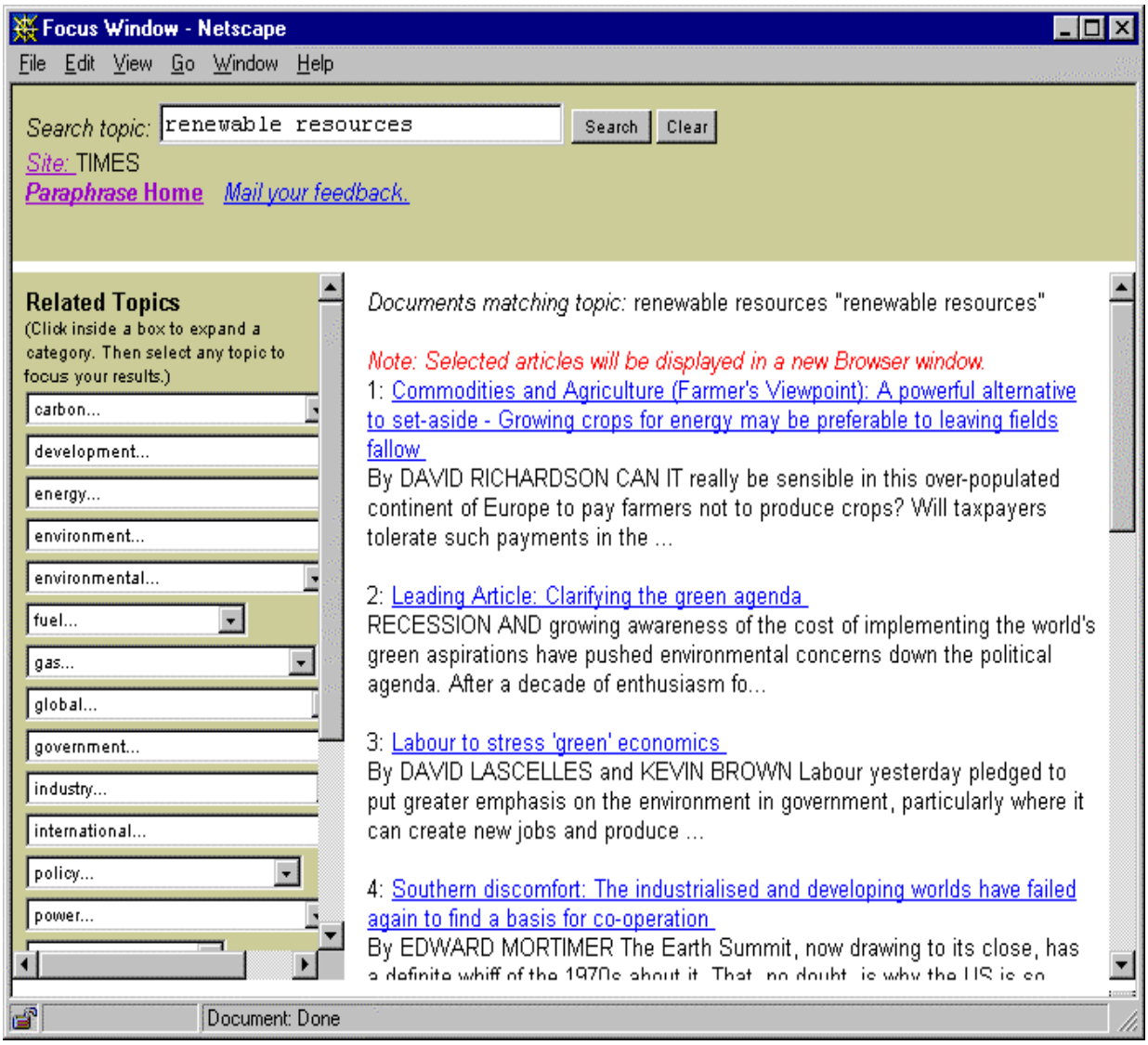

Figure 1. *Paraphrase* screen after entering the query "renewable resources".

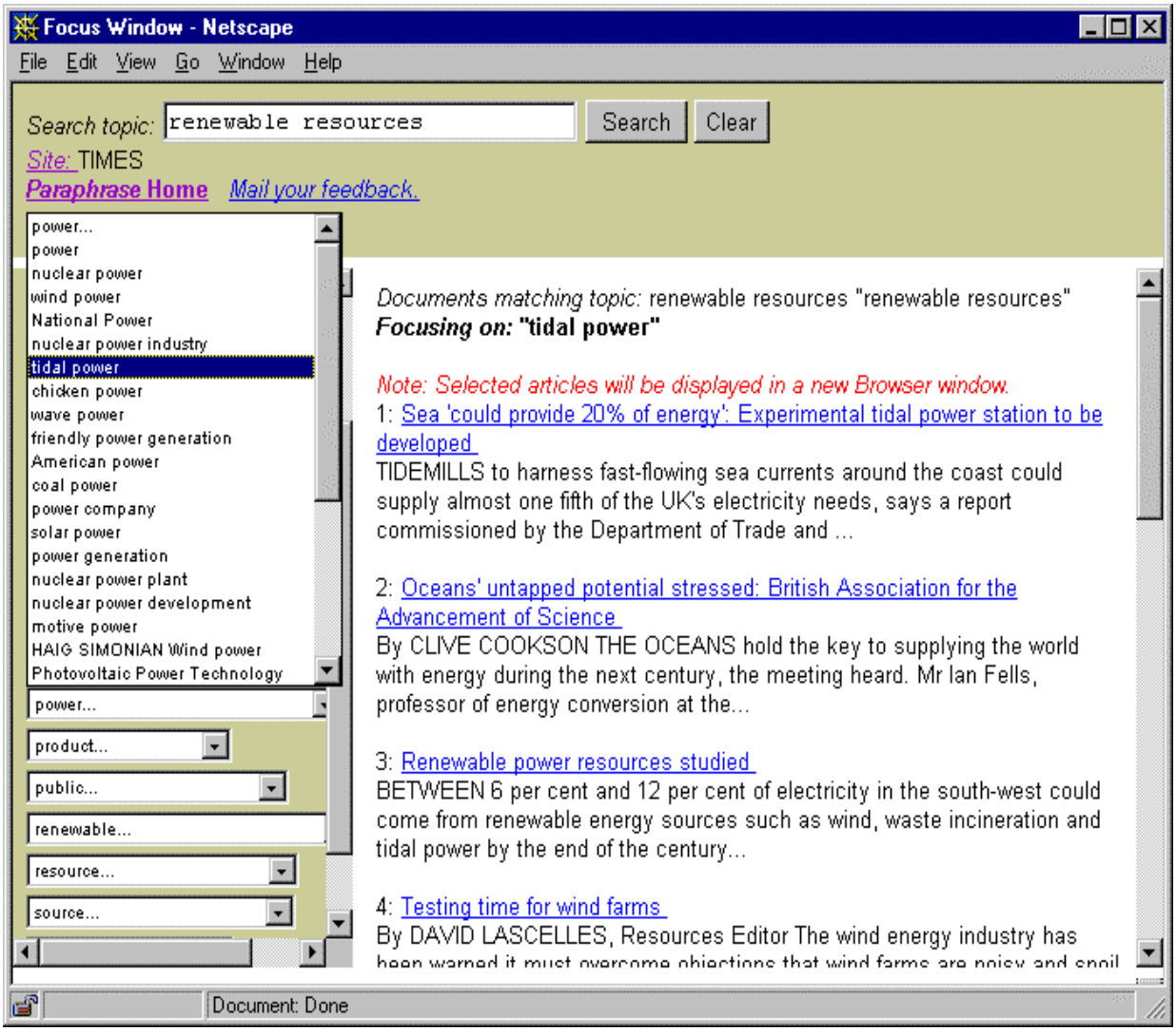

Figure 2. *Paraphrase* screen after selecting the term "tidal power" from the facet "power".