Automatic Summarization

Automatic Summarization

Ani Nenkova

University of Pennsylvania
USA
nenkova@seas.upenn.edu

Kathleen McKeown

Columbia University
USA
kathy@cs.columbia.edu

now
the essence of knowledge

Boston – Delft

Foundations and Trends® in Information Retrieval

Published, sold and distributed by:
now Publishers Inc.
PO Box 1024
Hanover, MA 02339
USA
Tel. +1-781-985-4510
www.nowpublishers.com
sales@nowpublishers.com

Outside North America:
now Publishers Inc.
PO Box 179
2600 AD Delft
The Netherlands
Tel. +31-6-51115274

The preferred citation for this publication is A. Nenkova and K. McKeown, Automatic Summarization, Foundation and Trends® in Information Retrieval, vol 5, nos 2–3, pp 103–233, 2011

ISBN: 978-1-60198-470-8

Foundations and Trends® in Information Retrieval
Volume 5 Issues 2–3, 2011

Editorial Scope

Foundations and Trends® in Information Retrieval will publish survey and tutorial articles in the following topics:

- Applications of IR
- Architectures for IR
- Collaborative filtering and recommender systems
- Cross-lingual and multilingual IR
- Distributed IR and federated search
- Evaluation issues and test collections for IR
- Formal models and language models for IR
- IR on mobile platforms
- Indexing and retrieval of structured documents
- Information categorization and clustering
- Information extraction
- Information filtering and routing
- Metasearch, rank aggregation and data fusion
- Natural language processing for IR
- Performance issues for IR systems, including algorithms, data structures, optimization techniques, and scalability
- Question answering
- Summarization of single documents, multiple documents, and corpora
- Text mining
- Topic detection and tracking
- Usability, interactivity, and visualization issues in IR
- User modelling and user studies for IR
- Web search

Information for Librarians
Foundations and Trends® in Information Retrieval, 2011, Volume 5, 5 issues. ISSN paper version 1554-0669. ISSN online version 1554-0677. Also available as a combined paper and online subscription.

Foundations and Trends® in
Information Retrieval
Vol. 5, Nos. 2–3 (2011) 103–233

DOI: 10.1561/1500000015

Automatic Summarization

Ani Nenkova[1] and Kathleen McKeown[2]

[1] *University of Pennsylvania, USA, nenkova@seas.upenn.edu*
[2] *Columbia University, USA, kathy@cs.columbia.edu*

Abstract

It has now been 50 years since the publication of Luhn's seminal paper on automatic summarization. During these years the practical need for automatic summarization has become increasingly urgent and numerous papers have been published on the topic. As a result, it has become harder to find a single reference that gives an overview of past efforts or a complete view of summarization tasks and necessary system components. This article attempts to fill this void by providing a comprehensive overview of research in summarization, including the more traditional efforts in sentence extraction as well as the most novel recent approaches for determining important content, for domain and genre specific summarization and for evaluation of summarization. We also discuss the challenges that remain open, in particular the need for language generation and deeper semantic understanding of language that would be necessary for future advances in the field.

We would like to thank the anonymous reviewers, our students and Noemie Elhadad, Hongyan Jing, Julia Hirschberg, Annie Louis, Smaranda Muresan and Dragomir Radev for their helpful feedback. This paper was supported in part by the U.S. National Science Foundation (NSF) under IIS-05-34871 and CAREER 09-53445. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NSF.

Contents

1

Introduction

Today's world is all about information, most of it online. The World Wide Web contains billions of documents and is growing at an exponential pace. Tools that provide timely access to, and digest of, various sources are necessary in order to alleviate the information overload people are facing. These concerns have sparked interest in the development of automatic summarization systems. Such systems are designed to take a single article, a cluster of news articles, a broadcast news show, or an email thread as input, and produce a concise and fluent summary of the most important information. Recent years have seen the development of numerous summarization applications for news, email threads, lay and professional medical information, scientific articles, spontaneous dialogues, voicemail, broadcast news and video, and meeting recordings. These systems, imperfect as they are, have already been shown to help users and to enhance other automatic applications and interfaces.

1.1 Types of Summaries

There are several distinctions typically made in summarization and here we define terminology that is often mentioned in the summarization literature.

Extractive summaries (extracts) are produced by concatenating several sentences taken exactly as they appear in the materials being summarized. *Abstractive summaries (abstracts)*, are written to convey the main information in the input and may reuse phrases or clauses from it, but the summaries are overall expressed in the words of the summary author.

Early work in summarization dealt with *single document summarization* where systems produced a summary of *one* document, whether a news story, scientific article, broadcast show, or lecture. As research progressed, a new type of summarization task emerged: *multi-document summarization*. Multi-document summarization was motivated by use cases on the web. Given the large amount of redundancy on the web, summarization was often more useful if it could provide a brief digest of many documents on the same topic or the same event. In the first deployed online systems, multi-document summarization was applied to clusters of news articles on the same event and used to produce online browsing pages of current events [130, 171].[1] A short one-paragraph summary is produced for each cluster of documents pertaining to a given news event, and links in the summary allow the user to directly inspect the original document where a given piece of information appeared. Other links provide access to all articles in the cluster, facilitating the browsing of news. User-driven clusters were also produced by collecting search engine results returned for a query or by finding articles similar to an example document the user has flagged as being of interest [173].

Summaries have also been distinguished by their content. A summary that enables the reader to determine about-ness has often been called an *indicative summary*, while one that can be read in place of the document has been called an *informative summary* [52]. An indicative summary may provide characteristics such as length, writing style, etc., while an informative summary will include facts that are reported in the input document(s).

[1] %%%%http://lada.si.umich.edu:8080/clair/nie1/nie.cgi, %%%%http://newsblaster.columbia.edu.

Most research in summarization deals with producing a short, paragraph-length summary. At the same time, a specific application or user need might call for a *keyword summary*, which consists of a set of indicative words or phrases mentioned in the input, or *headline summarization* in which the input document(s) is summarized by a single sentence.

Much of the work to date has been in the context of *generic summarization*. Generic summarization makes few assumptions about the audience or the goal for generating the summary. Typically, it is assumed that the audience is a general one: anyone may end up reading the summary. Furthermore, no assumptions are made about the genre or domain of the materials that need to be summarized. In this setting, importance of information is determined only with respect to the content of the input alone. It is further assumed that the summary will help the reader quickly determine what the document is about, possibly avoiding reading the document itself.

In contrast, in *query focused summarization*, the goal is to summarize only the information in the input document(s) that is relevant to a specific user query. For example, in the context of information retrieval, given a query issued by the user and a set of relevant documents retrieved by the search engine, a summary of each document could make it easier for the user to determine which document is relevant. To generate a useful summary in this context, an automatic summarizer needs to take the query into account as well as the document. The summarizer tries to find information within the document that is relevant to the query or in some cases, may indicate how much information in the document relates to the query. Producing snippets for search engines is a particularly useful query focused application [207, 213]. Researchers have also considered cases where the query is an open-ended question, with many different facts possibly being relevant as a response. A request for a biography is one example of an open-ended question as there are many different facts about a person that could be included, but are not necessarily required.

Update summarization addresses another goal that users may have. It is multi-document summarization that is sensitive to time; a summary must convey the important development of an event beyond what the user has already seen.

The contrast between generic, query-focused, and update summarization is suggestive of other issues raised by Sparck Jones in her 1998 call to arms [194]. Sparck Jones argued that summarization should not be done in a vacuum, but rather should be viewed as part of a larger context where, at the least, considerations such as the purpose of summarization (or task which it is part of), the reader for which it is intended, and the genre which is being summarized, are taken into account. She argued that generic summarization was unnecessary and in fact, wrong-headed. Of course, if we look at both sides of the question, we see that those who write newspaper articles do so in much the same spirit in which generic summaries are produced: the audience is a general one and the task is always the same. Nonetheless, her arguments are good ones as they force the system developer to think about other constraints on the summarization process and they raise the possibility of a range of tasks other than to simply condense content.

1.2 How do Summarization Systems Work?

Summarization systems take one or more documents as input and attempt to produce a concise and fluent summary of the most important information in the input. Finding the most important information presupposes the ability to understand the semantics of written or spoken documents. Writing a concise and fluent summary requires the capability to reorganize, modify and merge information expressed in different sentences in the input. Full interpretation of documents and generation of abstracts is often difficult for people,[2] and is certainly beyond the state of the art for automatic summarization.

How then do current automatic summarizers get around this conundrum? Most current systems avoid full interpretation of the input and generation of fluent output. The current state of the art in the vast majority of the cases relies on sentence extraction. The extractive approach to summarization focuses research on one key question: how

[2] For discussion of professional summarization, see [114].

can a system determine which sentences are important? Over the years, the field has seen advances in the sophistication of language processing and machine learning techniques that determine importance.

At the same time, there have been recent advances in the field which move toward semantic interpretation and generation of summary language. Semantic interpretation tends to be done for specialized summarization. For example, systems that produce biographical summaries or summaries of medical documents tend to use extraction of information rather than extraction of sentences. Research on generation for summarization uses a new form of generation, *text-to-text generation* and focuses on editing input text to better fit the needs of the summary.

1.2.1 Early Methods for Sentence Extraction

Most traditional approaches to summarization deal exclusively with the task of identifying important content, usually at the sentence level. The very first work on automatic summarization, done by Luhn [111] in the 1950s, set the tradition for sentence extraction.

His approach was implemented to work on technical papers and magazine articles. Luhn put forward a simple idea that shaped much of later research, namely that some words in a document are descriptive of its content, and the sentences that convey the most important information in the document are the ones that contain many such descriptive words close to each other that. He also suggested using frequency of occurrence in order to find which words are descriptive of the topic of the document; words that occur often in the document are likely to be the main topic of this document. Luhn brought up two caveats: some of the *most* common words in a technical paper or a magazine article, and in fact in any type of document, are not at all descriptive of its content. Common function words such as determiners, prepositions and pronouns do not have much value in telling us what the document is about. So he used a predefined list, called a *stop word list*, consisting of such words to remove them from consideration. Another class of words that do not appear in the stop word list but still cannot be indicative of the topic of a document are words common for a particular domain. For example, the word "cell" in a scientific paper in cell biology is not likely to give us much idea about what the paper is about. Finally, words that appear in the document only a few times are not informative either. Luhn used empirically determined high and low frequency thresholds for identifying descriptive words, with the high thresholds filtering out words that occur very frequently throughout the article and the low thresholds filtering out words that occur too infrequently. The remaining words are the descriptive words, indicative of the content that is important. Sentences characterized by high density of descriptive words, measured as clusters of five consecutive words by Luhn, are the most important ones and should be included in the summary.

In the next section we discuss how later work in sentence extraction adopted a similar view of finding important information but refined the ideas of using raw frequency by proposing weights for words, such as TF*IDF, in order to circumvent the need for coming up with arbitrary thresholds in determining which words are descriptive of a document. Later, statistical tests on word distributions were proposed to decide which words are topic words and which are not. Other approaches abandoned the idea of using words as the unit of operation, and used word frequency indirectly to model the similarity between sentences and derive measures of sentence importance from these relationships. We present these approaches in greater detail in Section 2, as they have proven to be highly effective, relatively robust to genre and domain, and are often referenced in work on automatic summarization.

There are some obvious problems with Luhn's approach. The same concept can be referred to using different words: consider for example "approach", "method" "algorithm", and "it". Different words may indicate a topic when they appear together; for example "hurricane", "damage", "casualties", "relief" evoke a natural disaster scenario. The same word can appear in different morphological variants — "show", "showing", "showed" — and counts of words as they appear in the text will not account for these types of repetition. In fact, Luhn was aware of these problems and he employed a rough approximation to morphological analysis, collapsing words that are similar except for the last six letters, to somewhat address this problem. After our presentation of word frequency-driven approaches in Section 2.1, we briefly discuss work based on the use of coreference systems and knowledge sources that perform input analysis and interpretation. These methods can better address these challenges and are discussed in Section 3. In essence these are still frequency approaches, but counting is performed in a more intelligent manner. Such approaches incur more processing overhead, which is often undesirable for practical purposes, but comes closer to the ideal of developing systems that are in fact interpreting the input before producing a summary.

Edmundson's [52] work was the foundation of several other trends in summarization research which eventually led to machine learning approaches in summarization. He expanded on Luhn's approach by proposing that multiple features may indicate sentence importance. He used a linear combination of features to weight sentences in a scientific article. His features were: (1) number of times a word appears in the article, (2) the number of words in the sentence that also appear in the title of the article, or in section headings, (3) position of the sentence in the article and in the section, (4) the number of sentence words matching a pre-compiled list of cue words such as "In sum". A compelling aspect of Edmundson's work that foreshadows today's empirically based approaches, was the creation of a document/extractive summary corpus. He used the corpus both to determine weights on the four features and to do evaluation. His results interestingly suggest that word frequency is the least important of the four classes of features, for his specific task and corpus. His other features take advantage of knowledge of the domain and genre of the input to the summarizer. We discuss such domain dependent approaches, which make use of domain-dependent knowledge sources and of specific domain characteristics, for summarization of scientific articles, medical information and email in Section 5.

In other relatively early and seminal work, Paice [164, 165] shifted the research focus toward the need for language generation techniques in summarization. He focused on the problem in extractive summarization of accidentally selecting sentences that contain unresolved references to sentences not included in the summary or not explicitly included in the original document. The problem can arise not only because of the presence of a pronouns but also because of a wide variety of other phrases (exophora) such as "Our investigations have shown this to be true." and "There are three distinct methods to be considered." Paice built an extractive summarizer which uses the presence of phrases from a list that he compiled, such as "The main goal of our paper ... ", to determine an initial set of seed sentences that should be selected. Then an aggregation procedure adds sentences preceding or following the seed until all exophora are resolved. Paice also suggested modifying sentences to resolve exophora when the reference can be found but did not implement an actual system for doing this. Paice's research was the first to point out the problem of accidentally including exophora in extractive summaries, but the solution of simply adding more sentences until the antecedent is found is not satisfactory and much later research on using language generation for summarization has revisited the problem as we discuss in Section 4.

1.2.2 Non-extractive Approaches

The current state of the art in the vast majority of the cases completely ignores issues of language generation and relies on sentence extraction, producing *extractive summaries* composed of important sentences taken verbatim from the input. The sole emphasis in such systems is to identify the important sentences that should appear in the summary. Meanwhile, the development of automatic methods for language generation and text quality has become somewhat independent subfields of research motivated but not directly linked to the field of summarization. Below we briefly introduce some of the main areas of research that are needed for enhancing current summarization systems.

Sentence ordering. This is the problem of taking several sentences, such as those deemed to be important by an extractive summarizer, and presenting them in the most coherent order.

Below we reproduce an example from [9], that shows two different orderings of the same content. The first example is one rated as poor by readers, and the second is one rated as good. The examples make it clear that the order of presentation makes a big difference for the overall quality of the summary and that certain orderings may pose problems for the reader trying to understand the gist of the presented information.

Summary 1; rated poor

P1 Thousands of people have attended a ceremony in Nairobi commemorating the first anniversary of the deadly bombings attacks against U.S. Embassies in Kenya and Tanzania.
P2 Saudi dissident Osama bin Laden, accused of masterminding the attacks, and nine others are still at large.
P3 President Clinton said, The intended victims of this vicious crime stood for everything that is right about our country and the world.
P4 U.S. federal prosecutors have charged 17 people in the bombings.
P5 Albright said that the mourning continues.
P6 Kenyans are observing a national day of mourning in honor of the 215 people who died there.

Summary 2; rated good

P1 Thousands of people have attended a ceremony in Nairobi commemorating the first anniversary of the deadly bombings attacks against U.S. Embassies in Kenya and Tanzania. Kenyans are observing a national day of mourning in honor of the 215 people who died there.
P2 Saudi dissident Osama bin Laden, accused of masterminding the attacks, and nine others are still at large. U.S. federal prosecutors have charged 17 people in the bombings.
P3 President Clinton said, "The intended victims of this vicious crime stood for everything that is right about our country and the world". Albright said that the mourning continues.

Sentence revision. Sentence revision was historically the first language generation task attempted in the context of summarization [89, 116, 146, 147, 162]. Sentence revision involves re-using text collected from the input to the summarizer, but parts of the final summary are automatically modified by substituting some expressions with other more appropriate expressions, given the context of the new summary. Types of revisions proposed by early researchers include ELIMINATION of unnecessary parts of the sentences, COMBINATION of information originally expressed in different sentences and SUBSTITUTION of a pronoun with a more descriptive noun phrase where the context of the summary requires this [116]. Given that implementation of these revision operations can be quite complex, researchers in the field eventually established largely non-overlapping sub-fields of research, each concentrating on only one type of revision.

Sentence fusion. Sentence fusion is the task of taking two sentences that contain some overlapping information, but that also have fragments that are different. The goal is to produce a sentence that conveys the information that is common between the two sentences, or a single sentence that contains all information in the two sentences, but without redundancy.

Here we reproduce two examples of fusion from [96]. The first one conveys only the information that is common to two different sentences, A and B, in the input documents to be summarized (intersection), while the second combines all the information for the two sentences (union).

Sentence A Post-traumatic stress disorder (PTSD) is a psychological disorder which is classified as an anxiety disorder in the DSM-IV.

Sentence B Post-traumatic stress disorder (abbrev. PTSD) is a psychological disorder caused by a mental trauma (also called psychotrauma) that can develop after exposure to a terrifying event.

Fusion 1 Post-traumatic stress disorder (PTSD) is a psychological disorder.

Fusion 2 Post-traumatic stress disorder (PTSD) is a psychological disorder, which is classified as an anxiety disorder in the DSM-IV, caused by a mental trauma (also called psychotrauma) that can develop after exposure to a terrifying event.

Sentence compression. Researchers interested in sentence compression were motivated by the observation that human summaries often contain parts of sentences from the original documents which are being summarized, but some portions of the sentence are removed to make it more concise.

Below we reproduce two examples from the Ziff-Davis corpus of sentence compression performed by a person, alongside the original sentence from the document. It is clear from the examples that compression not only shortens the original sentences but also makes them much easier to read.

Comp1 The Reverse Engineer Tool is priced from $8,000 for a single user to $90,000 for a multiuser project site.

Orig1 The Reverse Engineer Tool is available now and is priced on a site-licensing basis, ranging from $8,000 for a single user to $90,000 for a multiuser project site.

Comp2 Design recovery tools read existing code and translate it into definitions and structured diagrams.

Orig2 Essentially, design recovery tools read existing code and translate it into the language in which CASE is conversant — definitions and structured diagrams.

We discuss sentence ordering and language generation approaches in Section 4. The examples above clearly demonstrate the need for such approaches in order to build realistic, human-like, summarization systems. Yet the majority of current systems rely on sentence extraction for selecting content and do not use any of the text-to-text generation techniques, leaving the opportunity for significant improvements with further progress in language generation.

1.3 Evaluation Issues

The tension between generic and query-focused summarization, sentence-extraction and more sophisticated methods was also apparent in the context of the DUC (Document Understanding Conference) Evaluation Workshops [163]. Despite its name, DUC was initially formed in 2001 to evaluate work on summarization and was open to any group interested in participating. Its independent advisory board was charged with identifying tasks for evaluation. Generic summarization was the initial focus, but in later years it branched out to cover various task-based efforts, including a variation on query-focused summarization, topic-based summarization. The first of the topic-based tasks was to provide a summary of information about a person (similar to a biography), given a set of input documents on that person, while later on, a system was provided with a paragraph-length topic and a set of documents and was to use information within the documents to create a summary that addressed the topic.

Generic single document summarization of news was discontinued as a task at DUC after the first two years of the evaluations because no automatic summarizer could outperform the simple baseline consisting of the beginning of the news article, when using manual evaluation metrics. Similarly, for the task of headline generation — creating a 10-word summary of a single news article — no automatic approach outperformed the baseline of using the original headline of the article. For both tasks, human performance was significantly higher than that of the baselines, showing that while not yet attainable, better performance for automatic systems is possible [148].

DUC, which was superseded by the Text Analysis Conference (TAC) in 2007, provided much needed data to the research community, allowing the development of empirical approaches. Given the difficulty of evaluation, DUC fostered much research on evaluation. However, because the metrics emphasized content selection, research on the linguistic quality of a summary was not necessary. Furthermore, given the short time-frame within which tasks were introduced, summarization researchers who participated in DUC were forced to come up with a solution that was quick to implement. Exacerbating this more, given that people like to win, researchers were more likely to try incremental, safe approaches that were likely to come out on top. Thus, DUC in part encouraged the continuation of the "safe" approach, sentence extraction, even while it encouraged research on summarization and evaluation.

1.4 Where Does Summarization Help?

While evaluation forums such as DUC and TAC enable experimental setups through comparison to a gold standard, the ultimate goal in development of a summarization system is to help the end user perform a task better. Numerous task-based evaluations have been performed to establish that summarization systems are indeed effective in a variety of tasks. In the TIPSTER Text Summarization Evaluation (SUMMAC), single-document summarization systems were evaluated in a task-based scenario developed around the tasks of real intelligence analysts [113]. This large-scale study compared the performance of a human in judging if a particular document is relevant to a topic of interest, by reading either the full document or a summary thereof. It established that automatic text summarization is very effective in relevance assessment tasks on news articles. Summaries as short as 17% of the full text length sped up decision-making by almost a factor of two, with no statistically significant degradation in accuracy. Query-focused summaries are also very helpful in making relevance judgments about retrieved documents. They enable users to find more relevant documents more accurately, with less need to consult the full text of the document [203].

Multi-document summarization is key for organizing and presenting search results in order to reduce search time, especially when the goal of the user is to find as much information as possible about a given query [112, 131, 181]. In McKeown et al. [131], users were given a task of writing reports on specified topics, with an interface containing news articles, some relevant to the topic and some not. When articles were clustered and summaries for the related articles were provided, people tended to write better reports, but moreover, they reported higher satisfaction when using the information access interface augmented with summaries; they felt they had more time to complete the task. Similarly, in the work of Mana-López et al. [112], users had to find as many aspects as possible about a given topic. Clustering similar articles returned from a search engine together proved to be more advantageous than traditional ranked list presentation, and considerably improved user accuracy in finding relevant information. Providing a summary of the articles in each cluster that conveys the similarities between them, and single-document summaries highlighting the information specific to each document, also helped users in finding information, but in addition considerably reduced time as users read fewer full documents.

In summarization of scientific articles, the user goal is not only to find articles relevant to their interest, but also to understand in what respect a scientific paper relates to the previous work it describes and cites. In a study to test the utility of scientific paper summarization for determining which of the approaches mentioned in the paper are criticized and which approaches are supported and extended, automatic summaries were found to be almost as helpful as human-written ones, and significantly more useful than the original article abstract [199].

Voicemail summaries are helpful for recognizing the priority of the message, the call-back number, or the caller [95]; summaries of threads in help forums are useful in deciding if the thread is relevant [151], and summaries of meetings are a necessary part of interfaces for meeting browsing and search [205].

Numerous studies have also been performed to investigate and confirm the usefulness of single document summaries for improvement of other automated tasks. For example, Sakai and Sparck Jones [182] present the most recent and extensive study (others include [22] and several studies conducted in Japan and published in Japanese) on the usefulness of generic summaries for indexing in information retrieval. They show that, indeed, indexing for retrieval based on automatic summaries rather than full document text helps in certain scenarios for precision-oriented search. Similarly, query expansion in information retrieval is much more effective when potential expansion terms are selected from a summary of relevant documents instead of the full document [100].

Another unexpectedly successful application of summarization for improvement of an automatic task has been reported by [23]. They examined the impact of summarization on the automatic topic classification module that is part of a system for automatic scoring of student GMAT essays. Their results show that summarization of the student essay significantly improves the performance of the topical analysis component. The conjectured reason for the improvement is that the students write these essays under time constraints and do not have sufficient time for revision and thus their writing contains some digressions and repetitions, which are removed by the summarization module, allowing for better assessment of the overall topic of the essay.

The potential uses and applications of summarization are incredibly diverse as we have seen in this section. But how do these systems work and what are the open problems not currently handled by systems? We turn to this discussion next.

1.5 Article Overview

We begin our summarization overview with a presentation of research on sentence extraction in Section 2. In that section, we first present earlier research on summarization that experimented with methods for determining sentence importance that are based on variants of frequency. Machine learning soon became the method of choice for determining pertinent features for selecting sentences. From there, we move to graph-based approaches that select sentences based on their relations to other sentences. Finally, we close Section 2 by looking at the use of sentence extraction for query-focused summarization. One case of query-focused summarization is the generation of biographical summaries and we see that when the task is restricted (here to one class of queries), researchers begin to develop approaches that differ substantially from the typical generic extraction based approach.

In Section 3, we continue to survey extractive approaches, but move to methods that do more sophisticated analysis to determine importance. We begin with approaches that construct lexical chains which represent sentence relatedness through word and synonym overlap across sentences. The hypothesis is that each chain represents a topic and that topics that are pursued for greater lengths are likely to be more salient. We then turn to approaches that represent or compute concepts and select sentences that refer to salient concepts. Finally, we turn to methods that make use of discourse information, either in the form of rhetorical relations between sentences, or to augment graph-based approaches.

In Section 4, we examine the different sub-fields that have grown up around various forms of sentence revision. We look at methods that compress sentences by removing unnecessary detail. We then turn to methods that combine sentences by fusing together repeated and salient information from different sentences in the input. Next, we turn to work that edits summary sentences, taking into account the new context of the summary. We close with research on ordering of summary sentences.

In the final section on approaches, Section 5, we survey research that has been carried out for specific genres and domains. We find that often documents within a specific genre have an expected structure and that structure can be exploited during summary generation. This is the case, for example, with journal article summarization. At other times, we find that while the form of the genre creates problems (e.g., speech has disfluencies and errors resulting from recognition that cause difficulties), information beyond the words themselves may be available to help improve summarization results. In speech summarization, acoustic and prosodic clues can be used to identify important information, while in very recent work on web summarization, the structure of the web can be used to determine importance. In some domains, we find that domain dependant semantic resources are available and the nature of the text is more regular so that semantic interpretation followed by generation can be used to produce the summary; this is the case in the medical domain.

Before concluding, we provide an overview in Section 6 on research in summarization evaluation. Much of this work was initiated with DUC as the conference made evaluation data available to the community for the first time. Methodology for evaluation is a research issue in itself. When done incorrectly, evaluation does not accurately reveal which system performs better. In Section 6, we review *intrinsic methods* for evaluation. Intrinsic refers to methods that evaluate the quality of the summary produced, usually through comparison to a gold standard. This is in contrast to *extrinsic evaluation* where the evaluation measures the impact of the summary on task performance such as the task-based evaluations that we just discussed. We review metrics used for comparison against a gold standard as well as both manual and automatic methods for comparison. We discuss the difference between evaluation of summary content and evaluation of the linguistic quality of the summary.

References

[1] J. Allan, C. Wade, and A. Bolivar, "Retrieval and novelty detection at the sentence level," in *Proceedings of the Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 314–321, 2003.

[2] E. Amitay and C. Paris, "Automatically summarizing web sites - Is there a way around it?," in *Proceedings of the ACM Conference on Information and Knowledge Management*, pp. 173–179, 2000.

[3] R. H. Baayen, *Word Frequency Distributions*. Kluwer Academic Publishers, 2001.

[4] B. Baldwin and T. Morton, "Dynamic coreference-based summarization," in *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pp. 1–6, 1998.

[5] M. Banko, V. O. Mittal, and M. J. Witbrock, "Headline generation based on statistical translation," in *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, pp. 318–325, 2000.

[6] R. Barzilay and M. Lapata, "Modeling local coherence: An entity-based approach," *Computational Linguistics*, vol. 34, no. 1, pp. 1–34, 2008.

[7] R. Barzilay and M. Elhadad, "Text summarizations with lexical chains," in *Advances in Automatic Text Summarization*, (I. Mani and M. Maybury, eds.), pp. 111–121, MIT Press, 1999.

[8] R. Barzilay and N. Elhadad, "Sentence alignment for monolingual comparable corpora," in *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pp. 25–32, 2003.

[9] R. Barzilay, N. Elhadad, and K. McKeown, "Inferring strategies for sentence ordering in multidocument news summarization," *Journal of Artificial Intelligence Research*, vol. 17, pp. 35–55, 2002.

[10] R. Barzilay and L. Lee, "Catching the drift: Probabilistic content models, with applications to generation and summarization," in *Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics*, pp. 113–120, 2004.

[11] R. Barzilay and K. R. McKeown, "Sentence fusion for multidocument news summarization," *Computational Linguistics*, vol. 31, no. 3, pp. 297–328, 2005.

[12] M. Becher, B. Endres-Niggemeyer, and G. Fichtner, "Scenario forms for web information seeking and summarizing in bone marrow transplantation," in *Proceedings of the International Conference on Computational Linguistic*, pp. 1–8, 2002.

[13] A. Berger and V. Mittal, "OCELOT: A system for summarizing web pages.," in *Proceedings of the Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 144–151, 2000.

[14] G. Berland, M. Eilliot, L. Morales, J. Algazy, R. Kravitz, M. Broder, D. Kanouse, J. M. noz, J.-A. Puyol, M. Lara, K. Watkins, H. Yang, and E. McGlynn, "Health information on the Internet: Accessibility, quality and readability in English and Spanish," *American Medical Association*, vol. 285, no. 20, pp. 2612–2621, 2001.

[15] F. Biadsy, J. Hirschberg, and E. Filatova, "An Unsupervised Approach to Biography Production Using Wikipedia," in *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, pp. 807–815, 2008.

[16] S. Blair-Goldensohn, K. R. McKeown, and A. H. Schlaikjer, "DefScriber: A hybrid system for definitional QA," in *Proceedings of the Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 462–462, 2003.

[17] B. K. Boguraev and M. S. Neff, "Discourse segmentation in aid of document summarization," in *Proceedings of the Hawaii International Conference on System Sciences-Volume 3*, p. 3004, 2000.

[18] B. Boguraev and C. Kennedy, "Salience-based content characterization of text documents," in *Advances in Automatic Text Summarization*, pp. 2–9, The MIT Press, 1997.

[19] D. Bollegala, N. Okazaki, and M. Ishizuka, "A machine learning approach to sentence ordering for multidocument summarization and its evaluation," in *Proceedings of the International Joint Conference on Natural Language Processing*, pp. 624–635, 2005.

[20] D. Bollegala, N. Okazaki, and M. Ishizuka, "A bottom-up approach to sentence ordering for multi-document summarization," in *Proceedings of the International Conference on Computational Linguistics and the Annual Meeting of the Association for Computational Linguistics*, pp. 385–392, 2006.

[21] S. Branavan, H. Chen, J. Eisenstein, and R. Barzilay, "Learning document-level semantic properties from free-text annotations," in *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, pp. 263–271, 2008.

[22] R. Brandow, K. Mitze, and L. F. Rau, "Automatic condensation of electronic publications by sentence selection," *Information Processing and Management*, vol. 31, no. 5, pp. 675–685, 1995.

[23] J. Burstein and D. Marcu, "Toward using text summarization for essay-based feedback," in *Proceedings of Le Conference Annuelle sur Le Traitement Automatique des Langues Naturelles*, pp. 51–59, 2000.

[24] O. Buyukkokten, O. Kaljuvee, H. Garcia-Molina, A. Paepcke, and T. Winograd, "Efficient web browsing on handheld devices Using page and form summarization," in *ACM Transactions on Information Systems*, vol. 20, no. 1, pp. 82–115, January 2002.

[25] J. Carbonell and J. Goldstein, "The use of MMR, diversity-based rerunning for reordering documents and producing summaries," in *Proceedings of the Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 335–336, 1998.

[26] G. Carenini and J. C. K. Cheung, "Extractive vs. NLG-based abstractive summarization of evaluative text: The effect of corpus controversiality," in *Proceedings of the International Natural Language Generation Conference*, pp. 33–41, 2008.

[27] G. Carenini, R. T. Ng, and X. Zhou, "Summarizing email conversations with clue words," in *Proceedings of the International Conference on World Wide Web*, pp. 91–100, 2007.

[28] Y. Chali, S. Hasan, and S. Joty, "Do automatic annotation techniques have any impact on supervised complex question answering?," in *Proceedings of the Joint Conference of the Annual Meeting of the ACL and the International Joint Conference on Natural Language Processing of the AFNLP*, pp. 329–332, 2009.

[29] Y. Chali and S. Joty, "Improving the performance of the random walk model for answering complex questions," in *Proceedings of the Annual Meeting of the Association for Computational Linguistics, Short Papers*, pp. 9–12, 2008.

[30] Y. Choi, M. Fontoura, E. Gabrilovich, V. Josifovski, M. Mediano, and B. Pang, "Using landing pages for sponsored search ad selection," in *Proceedings of the International Conference on World Wide Web*, 2010.

[31] H. Christensen, B. Kolluru, Y. Gotoh, and S. Renals, "From text summarization to style-specific summarization for broadcast news," in *Proceedings of the European Conference on IR Research*, 2004.

[32] J. Clarke and M. Lapata, "Models for sentence compression: A comparison across domains, training requirements and evaluation measures," in *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, pp. 377–384, 2006.

[33] J. Clarke and M. Lapata, "Modelling compression with discourse constraints," in *Proceedings of the Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*, pp. 1–11, 2007.

[34] A. Cohen, "Unsupervised gene/protein entity normalization using automatically extracted dictionaries," in *Proceedings of the ACL-ISMB Workshop on Linking Biological Literature, Ontologies and Databases: Mining Biological Semantics*, pp. 17–24, Association for Computational Linguistics, 2005.

[35] W. W. Cohen, R. E. Schapire, and Y. Singer, "Learning to order things," *Journal of Artificial Intelligence Research*, vol. 10, pp. 243–270, 1998.

[36] T. Cohn and M. Lapata, "Sentence compression beyond word deletion," in *Proceedings of the International Conference on Computational Linguistic*, pp. 137–144, 2008.

[37] J. Conroy, J. Schlessinger, D. O'Leary, and J. Goldstein, "Back to basics: CLASSY 2006," in *Proceedings of the Document Understanding Conference*, 2006.

[38] J. Conroy, J. Schlesinger, J. Goldstein, and D. O'Leary, "Left-brain/right-brain multi-document summarization," in *Proceedings of the Document Understanding Conference*, 2004.

[39] J. M. Conroy and D. P. O'Leary, "Text summarization via hidden Markov models," in *Proceedings of the Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 406–407, 2001.

[40] J. M. Conroy, J. D. Schlesinger, and D. P. O'Leary, "Topic-focused multi-document summarization using an approximate oracle score," in *Proceedings of the International Conference on Computational Linguistics and the annual meeting of the Association for Computational Linguistics*, pp. 152–159, 2006.

[41] T. Copeck and S. Szpakowicz, "Leveraging pyramids," in *Proceedings of the Document Understanding Conference*, 2005.

[42] S. Corston-Oliver, E. Ringger, M. Gamon, and R. Campbell, "Task-focused summarization of email," in *Proceedings of the ACL Text Summarization Branches Out Workshop*, pp. 43–50, 2004.

[43] H. Daume III and D. Marcu, "Generic Sentence Fusion is an Ill-Defined Summarization Task," in *Proceedings of the ACL Text Summarization Branches Out Workshop*, pp. 96–103, 2004.

[44] H. Daumé III and D. Marcu, "A phrase-based HMM approach to document/abstract alignment," in *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pp. 119–126, 2004.

[45] H. Daumé III and D. Marcu, "Bayesian query-focused summarization," in *Proceedings of the International Conference on Computational Linguistics and the annual meeting of the Association for Computational Linguistics*, pp. 305–312, 2006.

[46] S. Deerwester, S. Dumais, G. W. Furnas, T. K. Landauer, and R. Harshman, "Indexing by Latent Semantic Analysis," *Journal of the American Society for Information Science*, pp. 391–407, 1990.

[47] J.-Y. Delort, B. Bouchon-Meunier, and M. Rifqi, "Enhanced web document summarization using hyperlinks," in *Proceedings of the ACM Conference on Hypertext and Hypermedia*, pp. 208–215, 2003.

[48] R. L. Donaway, K. W. Drummey, and L. A. Mather, "A comparison of rankings produced by summarization evaluation measures," in *Proceedings of the NAACL-ANLP Workshop on Automatic summarization*, pp. 69–78, 2000.

[49] B. Dorr, D. Zajic, and R. Schwartz, "Hedge Trimmer: A parse-and-trim approach to headline generation," in *Proceedings of the HLT-NAACL Workshop on Text Summarization*, pp. 1–8, 2003.

[50] P. A. Duboue, K. R. McKeown, and V. Hatzivassiloglou, "ProGenIE: Biographical descriptions for intelligence analysis," in *Proceedings of NSF/NIJ Symposium on Intelligence and Security Informatics*, vol. 2665, pp. 343–345, Springer-Verlag, 2003.

[51] T. Dunning, "Accurate methods for the statistics of surprise and coincidence," *Computational Linguistics*, vol. 19, no. 1, pp. 61–74, 1994.

[52] H. P. Edmundson, "New methods in automatic extracting," *Journal of the ACM*, vol. 16, no. 2, pp. 264–285, 1969.

[53] N. Elhadad, M.-Y. Kan, J. Klavans, and K. McKeown, "Customization in a unified framework for summarizing medical literature," *Journal of Artificial Intelligence in Medicine*, vol. 33, pp. 179–198, 2005.

[54] N. Elhadad, K. McKeown, D. Kaufman, and D. Jordan, "Facilitating physicians' access to information via tailored text summarization," in *Proceedings of the AMIA Annual Symposium*, pp. 226–300, 2005.

[55] G. Erkan and D. Radev, "The University of Michigan at DUC 2004," in *Proceedings of the Document Understanding Conference*, 2004.

[56] G. Erkan and D. R. Radev, "LexRank: Graph-based centrality as salience in text summarization," *Journal of Artificial Intelligence Research*, 2004.

[57] D. Feng and E. Hovy, "Handling biographical questions with implicature," in *Proceedings of the Conference on Human Language Technology and Empirical Methods in Natural Language Processing*, pp. 596–603, 2005.

[58] E. Filatova and V. Hatzivassiloglou, "A formal model for information selection in multi-sentence text extraction," in *Proceedings of the International Conference on Computational Linguistic*, pp. 397–403, 2004.

[59] K. Filippova and M. Strube, "Sentence fusion via dependency graph compression," in *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pp. 177–185, 2008.

[60] M. Fuentes, E. Alfonseca, and H. Rodríguez, "Support Vector Machines for query-focused summarization trained and evaluated on Pyramid data," in *Proceedings of the Annual Meeting of the Association for Computational Linguistics, Companion Volume: Proceedings of the Demo and Poster Sessions*, pp. 57–60, 2007.

[61] P. Fung and G. Ngai, "One story, one flow: Hidden Markov Story Models for multilingual multidocument summarization," *ACM Transactions on Speech and Language Processing*, vol. 3, no. 2, pp. 1–16, 2006.

[62] S. Furui, M. Hirohata, Y. Shinnaka, and K. Iwano, "Sentence extraction-based automatic speech summarization and evaluation techniques," in *Proceedings of the Symposium on Large-scale Knowledge Resources*, pp. 33–38, 2005.

[63] M. Galley, "A skip-chain conditional random field for ranking meeting utterances by importance," in *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pp. 364–372, 2006.

[64] M. Galley and K. McKeown, "Improving word sense disambiguation in lexical chaining," in *Proceedings of the International Joint Conference on Artificial Intelligence*, pp. 1486–1488, 2003.

[65] M. Galley and K. McKeown, "Lexicalized Markov grammars for sentence compression," in *Human Language Technologies: The Conference of the*

North American Chapter of the Association for Computational Linguistics, pp. 180–187, 2007.

[66] M. Galley, K. McKeown, J. Hirschberg, and E. Shriberg, "Identifying agreement and disagreement in conversational speech: Use of Bayesian networks to model pragmatic dependencies," in *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, pp. 669–676, 2004.

[67] D. Gillick, K. Riedhammer, B. Favre, and D. Hakkani-Tur, "A global optimization framework for meeting summarization," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 4769–4772, 2009.

[68] J. Glass, T. Hazen, S. Cyphers, I. Malioutov, D. Huynh, and R. Barzilay, "Recent progress in the MIT spoken lecture processing project," in *Proceedings of the Annual Conference of the International Speech Communication Association*, pp. 2553–2556, 2007.

[69] Y. Gong and X. Liu, "Generic text summarization using relevance measure and latent semantic analysis," in *Proceedings of the Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 19–25, 2001.

[70] S. Gupta, A. Nenkova, and D. Jurafsky, "Measuring importance and query relevance in topic-focused multi-document summarization," in *Proceedings of the Annual Meeting of the Association for Computational Linguistics, Demo and Poster Sessions*, pp. 193–196, 2007.

[71] I. Gurevych and T. Nahnsen, "Adapting lexical chaining to summarize conversational dialogues," in *Proceedings of the Recent Advances in Natural Language Processing Conference*, pp. 287–300, 2005.

[72] I. Gurevych and M. Strube, "Semantic similarity applied to spoken dialogue Summarization," in *Proceedings of the International Conference on Computational Linguistic*, pp. 764–770, 2004.

[73] B. Hachey, G. Murray, and D. Reitter, "Dimensionality reduction aids term co-occurrence based multi-document summarization," in *SumQA '06: Proceedings of the Workshop on Task-Focused Summarization and Question Answering*, pp. 1–7, 2006.

[74] D. Hakkani-Tur and G. Tur, "Statistical sentence extraction for information distillation," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, vol. 4, pp. IV–1 –IV–4, 2007.

[75] D. Harman and P. Over, "The effects of human variation in DUC summarization evaluation," in *Proceedings of ACL Text Summarization Branches Out Workshop*, pp. 10–17, 2004.

[76] V. Hatzivassiloglou, J. L. Klavans, M. L. Holcombe, R. Barzilay, M. yen Kan, and K. R. McKeown, "SIMFINDER: A flexible clustering tool for summarization," in *Proceedings of the NAACL Workshop on Automatic Summarization*, pp. 41–49, 2001.

[77] M. Hirohata, Y. Shinnaka, K. Iwano, and S. Furui, "Sentence-extractive automatic speech summarization and evaluation techniques," *Speech Communication*, vol. 48, no. 9, pp. 1151–1161, 2006.

[78] C. Hori and S. Furui, "Speech summarization: An approach through word extraction and a method for evaluation," *IEICE Transactions on Information and Systems*, vol. 87, pp. 15–25, 2004.

[79] C. Hori, S. Furui, R. Malkin, H. Yu, and A. Waibel, "Automatic speech summarization applied to English broadcast news speech," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 1–9, 2002.

[80] C. Hori, T. Hori, and S. Furui, "Evaluation methods for automatic speech summarization," in *Proceedings of the European Conference on Speech Communication and Technology*, pp. 2825–2828, 2003.

[81] E. Hovy and C.-Y. Lin, "Automated Text Summarization in SUMMARIST," in *Advances in Automatic Text Summarization*, pp. 82–94, 1999.

[82] G. Hripcsak, J. Cimino, and S. Sengupta, "WebCIS: Large scale deployment of a web-based clinical information system," in *Proceedings of the AMIA Annual Symposium*, pp. 804–808, 1999.

[83] M. Hu, A. Sun, and E.-P. Lim, "Comments-oriented blog summarization by sentence extraction," in *Proceedings of the ACM Conference on Information and Knowledge Management*, pp. 901–904, 2007.

[84] D. Ji and Y. Nie, "Sentence ordering based on cluster adjacency in multi-document summarization," in *Proceedings of the International Joint Conference on Natural Language Processing*, pp. 745–750, 2008.

[85] H. Jing, R. Barzilay, K. McKeown, and M. Elhadad, "Summarization evaluation methods: Experiments and analysis," in *AAAI Symposium on Intelligent Summarization*, pp. 51–59, 1998.

[86] H. Jing and K. McKeown, "The Decomposition of Human-Written Summary Sentences," in *Proceedings of the Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 129–136, 1999.

[87] H. Jing, "Sentence reduction for automatic text summarization," in *Proceedings of the Conference on Applied Natural Language Processing*, pp. 310–315, 2000.

[88] H. Jing, "Using Hidden Markov modeling to decompose human-written summaries," *Computational linguistics*, vol. 28, no. 4, pp. 527–543, 2002.

[89] H. Jing and K. R. McKeown, "Cut and paste based text summarization," in *Proceedings of the North American chapter of the Association for Computational Linguistics Conference*, pp. 178–185, 2000.

[90] H. Jing, *Cut-and-paste text summarization.* PhD thesis, Columbia University, 2001.

[91] M. Johnson and E. Charniak, "A TAG-based noisy-channel model of speech repairs," in *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, pp. 33–39, 2004.

[92] M.-Y. Kan, "Combining visual layout and lexical cohesion features for text segmentation," Tech. Rep. CUCS-002-01, Columbia University, 2001.

[93] T. Kikuchi, S. Furui, and C. Hori, "Automatic speech summarization based on sentence extraction and compaction," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 384–387, 2003.

[94] K. Knight and D. Marcu, "Summarization beyond sentence extraction: A probabilistic approach to sentence compression," *Artiitial Intelligence*, vol. 139, no. 1, pp. 91–107, 2002.

[95] K. Koumpis and S. Renals, "Automatic summarization of voicemail messages using lexical and prosodic features," *ACM Transactions on Speech and Language Processing*, vol. 2, no. 1, pp. 1–24, 2005.

[96] E. Krahmer, E. Marsi, and P. van Pelt, "Query-based sentence fusion is better defined and leads to more preferred results than generic sentence fusion," in *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, pp. 193–196, 2008.

[97] J. Kupiec, J. Pedersen, and F. Chen, "A trainable document summarizer," in *Proceedings of the Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 68–73, 1995.

[98] F. Lacatusu, A. Hickl, S. Harabagiu, and L. Nezda, "Lite_GISTexter at DUC2004," in *Proceedings of the Document Understanding Conference*, 2004.

[99] D. Lam, S. L. Rohall, C. Schmandt, and M. K. Stern, "Exploiting e-mail structure to improve summarization," in *Proceedings of the ACM Conference on Computer Supported Cooperative Work*, 2002.

[100] A. M. Lam-Adesina and G. J. F. Jones, "Applying summarization techniques for term selection in relevance feedback," in *Proceedings of the Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 1–9, 2001.

[101] M. Lapata and R. Barzilay, "Automatic evaluation of text coherence: models and representations," in *Proceedings of the International Joint Conference on Artificial Intelligence*, pp. 1085–1090, 2005.

[102] K. Lerman, S. Blair-Goldensohn, and R. McDonald, "Sentiment summarization: evaluating and learning user preferences," in *Proceedings of the Conference of the European Chapter of the Association for Computational Linguistics*, pp. 514–522, 2009.

[103] J. Leskovec, N. Milic-frayling, and M. Grobelnik, "Impact of linguistic analysis on the semantic graph coverage and learning of document extracts," in *Proceedings of the National Conference on Artificial Intelligence*, pp. 1069–1074, 2005.

[104] C. Lin and E. Hovy, "From single to multi-document summarization: A prototype system and its evaluation," in *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, pp. 457–464, 2002.

[105] C.-Y. Lin, "ROUGE: A package for automatic evaluation of summaries," in *Proceedings of ACL Text Summarization Branches Out Workshop*, pp. 74–81, 2004.

[106] C.-Y. Lin and E. Hovy, "The automated acquisition of topic signatures for text summarization," in *Proceedings of the International Conference on Computational Linguistic*, pp. 495–501, 2000.

[107] C.-Y. Lin and E. Hovy, "Automatic evaluation of summaries using n-gram cooccurance statistics," in *Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics*, 2003.

[108] Y. Liu, E. Shriberg, A. Stolcke, D. Hillard, M. Ostendorf, and M. Harper, "Enriching speech recognition with automatic detection of sentence boundaries and disfluencies," *IEEE Transactions on Audio, Speech and Language Processing*, vol. 14, pp. 1526–1540, 2006.

[109] Y. Liu and S. Xie, "Impact of automatic sentence segmentation on meeting summarization," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 5009–5012, 2008.

[110] A. Louis, A. Joshi, and A. Nenkova, "Discourse indicators for content selection in summarization," in *Proceedings of the Annual Meeting of the Special Interest Group on Discourse and Dialogue*, pp. 147–156, 2010.

[111] H. P. Luhn, "The automatic creation of literature abstracts," *IBM Journal of Research and Development*, vol. 2, no. 2, pp. 159–165, 1958.

[112] M. Mana-López, M. D. Buenaga, and J. M. Gómez-Hidalgo, "Multidocument summarization: An added value to clustering in interactive retrieval," *ACM Transactions on Informations Systems*, vol. 22, no. 2, pp. 215–241, 2004.

[113] I. Mani, G. Klein, D. House, L. Hirschman, T. Firmin, and B. Sundheim, "SUMMAC: A text summarization evaluation," *Natural Language Engineering*, vol. 8, no. 1, pp. 43–68, 2002.

[114] I. Mani, *Automatic Summarization*. John Benjamins, 2001.

[115] I. Mani and E. Bloedorn, "Summarizing similarities and differences among related documents," *Information Retrieval*, vol. 1, no. 1-2, pp. 35–67, April 1999.

[116] I. Mani, B. Gates, and E. Bloedorn, "Improving summaries by revising them," in *Proceedings of the annual meeting of the Association for Computational Linguistics*, pp. 558–565, 1999.

[117] W. Mann and S. Thompson, "Rhetorical Structure Theory: Towards a functional theory of text organization," *Text*, vol. 8, pp. 243–281, 1988.

[118] C. D. Manning and H. Schutze, *Foundations of Natural Language Processing*. MIT Press, 1999.

[119] D. Marcu, "From discourse structure to text summaries," in *Proceedings of ACL/EACL 1997 Summarization Workshop*, pp. 82–88, 1997.

[120] D. Marcu, "To build text summaries of high quality, nuclearity is not sufficient," in *Working Notes of the AAAI-98 Spring Symposium on Intelligent Text Summarization*, pp. 1–8, 1998.

[121] D. Marcu, "The automatic construction of large-scale corpora for summarization research," in *Proceedings of the Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 137–144, 1999.

[122] D. Marcu, *The Theory and Practice of Discourse and Summarization*. The MIT Press, 2000.

[123] E. Marsi and E. Krahmer, "Explorations in sentence fusion," in *Proceedings of the European Workshop on Natural Language Generation 2005*, pp. 109–117, 2005.

[124] S. Maskey and J. Hirschberg, "Summarizing speech without text using Hidden Markov Models," in *Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics, Companion Volume: Short Papers*, pp. 89–92, 2006.

[125] S. Maskey, A. Rosenberg, and J. Hirschberg, "Intonational phrases for speech summarization," in *Proceedings of the Annual Conference of the International Speech Communication Association*, pp. 2430–2433, 2008.
[126] R. McDonald, "Discriminative sentence compression with soft syntactic evidence," in *Proceedings of the Conference of the European Chapter of the Association for Computational Linguistics*, pp. 297–304, 2006.
[127] R. McDonald, "A study of global inference algorithms in multi-document summarization," in *Proceedings of the European Conference on IR Research*, pp. 557–564, 2007.
[128] K. McKeown, R. Barzilay, D. Evans, V. Hatzivassiloglou, B. Schiffman, and S. Teufel, "Columbia multi-document summarization: Approach and evaluation," in *Proceedings of the Document Understanding Conference*, 2001.
[129] K. McKeown, S.-F. Chang, J. Cimino, S. Feiner, C. Friedman, L. Gravano, V. Hatzivassiloglou, S. Johnson, D. Jordan, J. Klavans, A. Kushniruk, V. Patel, and S. Teufel, "PERSIVAL, a system for personalized search and summarization over multimedia healthcare information," in *Proceedings of the ACM/IEEE-CS Joint Conference on Digital Libraries*, pp. 331–340, 2001.
[130] K. McKeown, R. Barzilay, D. Evans, V. Hatzivassiloglou, J. Klavans, A. Nenkova, C. Sable, B. Schiffman, and S. Sigelman, "Tracking and summarizing news on a daily basis with Columbia's Newsblaster," in *Proceedings of the International Conference on Human Language Technology Research*, 2002.
[131] K. McKeown, R. J. Passonneau, D. K. Elson, A. Nenkova, and J. Hirschberg, "Do summaries help? A task-based evaluation of multi-document summarization," in *Proceedings of the Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 210–217, 2005.
[132] K. Mckeown, L. Shrestha, and O. Rambow, "Using question-answer pairs in extractive summarization of email conversations," in *Proceedings of the International Conference on Computational Linguistics and Intelligent Text Processing*, pp. 542–550, 2007.
[133] K. R. McKeown, J. L. Klavans, V. Hatzivassiloglou, R. Barzilay, and E. Eskin, "Towards multidocument summarization by reformulation: progress and prospects," in *Proceedings of the National Conference on Artificial Intelligence*, pp. 453–460, 1999.
[134] Q. Mei and C. Zhai, "Generating impact-based summaries for scientific literature," in *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, pp. 816–824, 2008.
[135] R. Mihalcea and P. Tarau, "Textrank: Bringing order into texts," in *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pp. 404–411, 2004.
[136] R. Mihalcea and P. Tarau, "An algorithm for language independent single and multiple document summarization," in *Proceedings of the International Joint Conference on Natural Language Processing*, pp. 19–24, 2005.
[137] G. Miller, R. Beckwith, C. Fellbaum, D. Gross, and K. J. Miller, "Introduction to WordNet: An on-line lexical database," *International Journal of Lexicography (special issue)*, vol. 3, no. 4, pp. 235–312, 1990.

[138] T. Miller and W. Schuler, "A syntactic time-series model for parsing fluent and disfluent speech," in *Proceedings of the International Conference on Computational Linguistic*, pp. 569–576, 2008.

[139] J. Mrozinski, E. W. D. Whittaker, P. Chatain, and S. Furui, "Automatic sentence segmentation of speech for automatic summarization," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 981–984, 2006.

[140] H. Murakoshi, A. Shimazu, and K. Ochimizu, "Construction of deliberation structure in email conversation," in *Proceedings of the Conference of the Pacific Association for Computational Linguistics*, pp. 570–577, 2004.

[141] S. Muresan, E. Tzoukermann, and J. L. Klavans, "Combining linguistic and machine learning techniques for email summarization," in *Proceedings of the Workshop on Computational Natural Language Learning*, pp. 1–8, 2001.

[142] G. Murray, S. Renals, and J. Carletta, "Extractive summarization of meeting recordings," in *Proceedings of 9th European Conference on Speech Communication and Technology*, pp. 593–596, 2005.

[143] G. Murray, S. Renals, J. Carletta, and J. Moore, "Evaluating automatic summaries of meeting recordings," in *Proceedings of the ACL Workshop on Evaluation Measures for MT/Summarization*, 2005.

[144] G. Murray and S. Renals, "Term-weighting for summarization of multi-party spoken dialogues," in *Proceedings of the International Workshop on Machine Learning for Multimodal Interaction*, vol. 4892, pp. 155–166, 2007.

[145] H. Nanba and M. Okumura, "Towards multi-paper summarization using reference information," in *Proceedings of the International Joint Conference on Artificial Intelligence*, pp. 926–931, 1999.

[146] H. Nanba and M. Okumura, "Producing more readable extracts by revising them," in *Proceedings of the International Conference on Computational Linguistic*, pp. 1071–1075, 2000.

[147] A. Nenkova and K. McKeown, "References to named entities: A corpus study," in *Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics*, pp. 70–72, 2003.

[148] A. Nenkova, "Automatic text summarization of newswire: lessons learned from the document understanding conference," in *Proceedings of the National Conference on Artificial Intelligence*, pp. 1436–1441, 2005.

[149] A. Nenkova, "Entity-driven rewrite for multi-document summarization," in *Proceedings of the International Joint Conference on Natural Language Processing*, 2008.

[150] A. Nenkova, *Understanding the process of multi-document summarization: content selection, rewrite and evaluation*. PhD thesis, Columbia University, January 2006.

[151] A. Nenkova and A. Bagga, "Facilitating email thread access by extractive summary generation," in *Proceedings of the Recent Advances in Natural Language Processing Conference*, 2003.

[152] A. Nenkova and R. Passonneau, "Evaluating content selection in summarization: The pyramid method," in *Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics*, pp. 145–152, 2004.

[153] A. Nenkova, R. Passonneau, and K. McKeown, "The Pyramid method: Incorporating human content selection variation in summarization evaluation," *ACM Transactions on Speech and Language Processing*, vol. 4, no. 2, 2007.

[154] A. Nenkova, A. Siddharthan, and K. McKeown, "Automatically learning cognitive status for multi-document summarization of newswire," in *Proceedings of the Conference on Human Language Technology and Empirical Methods in Natural Language Processing*, pp. 241–248, 2005.

[155] A. Nenkova, L. Vanderwende, and K. McKeown, "A compositional context sensitive multi-document summarizer: Exploring the factors that influence summarization," in *Proceedings of the Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 573–580, 2006.

[156] P. S. Newman and J. C. Blitzer, "Summarizing archived discussions: A beginning," in *Proceedings of the International Conference on Intelligent user Interfaces*, pp. 273–276, 2003.

[157] M. L. Nguyen, A. Shmazu, S. Horiguchi, T. B. Ho, and M. Fukushi, "Probabilistic sentence reduction using support vector machines," in *Proceedings of the International Conference on Computational Linguistic*, pp. 743–49, 2004.

[158] K. Ohtake, K. Yamamoto, Y. Toma, S. Sado, S. Masuyama, and S. Nakagawa, "Newscast Speech Summarization via Sentence Shortening Based on Prosodic Features," in *Proceedings of ISCA and IEEE Workshop on Spontaneous Speech Processing and Recognition*, pp. 247–254, 2003.

[159] K. Ono, K. Sumita, and S. Miike, "Abstract generation based on rhetorical structure extraction," in *Proceedings of the International Conference on Computational Linguistic*, pp. 344–348, 1994.

[160] M. Osborne, "Using maximum entropy for sentence extraction," in *Proceedings of the ACL Workshop on Automatic Summarization*, pp. 1–8, 2002.

[161] J. Otterbacher, G. Erkan, and D. Radev, "Using random walks for question-focused sentence retrieval," in *Proceedings of the Conference on Human Language Technology and Empirical Methods in Natural Language Processing*, pp. 915–922, 2005.

[162] J. Otterbacher, D. Radev, and A. Luo, "Revisions that improve cohesion in multi-document summaries: A preliminary study," in *Proceedings of the Workshop on Automatic Summarization*, pp. 2–7, 2002.

[163] P. Over, H. Dang, and D. Harman, "DUC in context," *Information Processing and Managemant*, vol. 43, no. 6, pp. 1506–1520, 2007.

[164] C. D. Paice, "The automatic generation of literature abstracts: An approach based on the identification of self-indicating phrases," in *Proceedings of the Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 172–191, 1981.

[165] C. D. Paice, "Constructing literature abstracts by computer: Techniques and prospects," *Information Processing and Management*, vol. 26, no. 1, pp. 171–186, 1990.

[166] K. Papineni, S. Roukos, T. Ward, and W.-J. Zhu, "Bleu: A method for automatic evaluation of machine translation," in *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, pp. 311–318, 2002.

[167] F. Peng, R. Weischedel, A. Licuanan, and J. Xu, "Combining deep linguistics analysis and surface pattern learning: A hybrid approach to Chinese definitional question answering," in *Proceedings of the Conference on Human Language Technology and Empirical Methods in Natural Language Processing*, pp. 307–314, 2005.

[168] G. Penn and X. Zhu, "A critical reassessment of evaluation baselines for speech summarization," in *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, pp. 470–478, 2008.

[169] V. Qazvinian and D. R. Radev, "Scientific paper summarization using citation summary networks," in *Proceedings of the International Conference on Computational Linguistic*, pp. 689–696, 2008.

[170] D. Radev and K. McKeown, "Generating natural language summaries from multiple on-line sources," *Computational Linguistics*, vol. 24, no. 3, pp. 469–500, 1998.

[171] D. Radev, J. Otterbacher, A. Winkel, and S. Blair-Goldensohn, "NewsInEssence: Summarizing online news topics," *Communications of the ACM*, vol. 48, no. 10, pp. 95–98, 2005.

[172] D. Radev and D. Tam, "Single-document and multi-document summary evaluation via relative utility," in *Proceedings of the ACM Conference on Information and Knowledge Management*, pp. 508–511, 2003.

[173] D. R. Radev, S. Blair-Goldensohn, Z. Zhang, and R. S. Raghavan, "NewsInEssence: A system for domain-independent, real-time news clustering and multi-document summarization," in *Proceedings of the International Conference on Human Language Technology Research*, pp. 1–4, 2001.

[174] D. R. Radev, H. Jing, and M. Budzikowska, "Centroid-based summarization of multiple documents: Sentence extraction, utility-based evaluation, and user studies," in *Proceedings of the NAACL-ANLP Workshop on Automatic summarization*, pp. 21–30, 2000.

[175] D. R. Radev and K. R. McKeown, "Building a generation knowledge source using internet-accessible newswire," in *Proceedings of the Conference on Applied Natural Language Processing*, pp. 221–228, 1997.

[176] D. R. Radev, S. Teufel, H. Saggion, W. Lam, J. Blitzer, H. Qi, A. Çelebi, D. Liu, and E. Drabek, "Evaluation challenges in large-scale document summarization: The MEAD project," in *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, pp. 375–382, 2003.

[177] O. Rambow, L. Shrestha, J. Chen, and C. Lauridsen, "Summarizing email threads," in *Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics*, 2004.

[178] G. J. Rath, A. Resnick, and R. Savage, "The formation of abstracts by the selection of sentences: Part 1: Sentence selection by man and machines," *American Documentation*, vol. 2, no. 12, pp. 139–208, 1961.

[179] K. Riedhammer, D. Gillick, B. Favre, and D. Hakkani-Tur, "Packing the meeting summarization knapsack," in *Proceedings of the Annual Conference of the International Speech Communication Association*, pp. 2434–2437, 2008.

[180] S. Riezler, T. H. King, R. Crouch, and A. Zaenen, "Statistical sentence condensation using ambiguity packing and stochastic disambiguation methods for

lexical-functional grammar," in *Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics*, pp. 118–125, 2003.

[181] D. G. Roussinov and H. Chen, "Information navigation on the web by clustering and summarizing query results," *Information Processing and Management*, vol. 37, no. 6, pp. 789–816, 2001.

[182] T. Sakai and K. Sparck Jones, "Generic summaries for indexing in information retrieval," in *Proceedings of the Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 190–198, 2001.

[183] G. Salton, A. Singhal, M. Mitra, and C. Buckley, "Automatic text structuring and summarization," *Information Processing and Management*, vol. 33, no. 2, pp. 193–208, 1997.

[184] G. Salton and C. Buckley, "Term-weighting approaches in automatic text retrieval," *Information Processing and Management*, vol. 24, pp. 513–523, 1988.

[185] B. Schiffman, I. Mani, and K. Concepcion, "Producing biographical summaries: Combining linguistic knowledge with corpus statistics," in *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, pp. 458–465, 2001.

[186] B. Schiffman, A. Nenkova, and K. McKeown, "Experiments in multidocument summarization," in *Proceedings of the International Conference on Human Language Technology Research*, pp. 52–58, 2002.

[187] L. Shrestha and K. Mckeown, "Detection of question-answer pairs in email conversations," in *Proceedings of the International Conference on Computational Linguistic*, pp. 889–895, 2004.

[188] E. Shriberg, A. Stolcke, D. Hakkani-Tür, and G. Tür, "Prosody-based automatic segmentation of speech into sentences and topics," *Speech Communication*, vol. 32, no. 1-2, pp. 127–154, 2000.

[189] A. Siddharthan and K. McKeown, "Improving multilingual summarization: using redundancy in the input to correct MT errors," in *Proceedings of the Conference on Human Language Technology and Empirical Methods in Natural Language Processing*, pp. 33–40, 2005.

[190] A. Siddharthan, A. Nenkova, and K. Mckeown, "Syntactic simplification for improving content selection in multi-document summarization," in *Proceedings of the International Conference on Computational Linguistic*, pp. 896–902, 2004.

[191] A. Siddharthan and S. Teufel, "Whose idea was this, and why does it matter? Attributing scientific work to citations," in *Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics*, pp. 316–323, 2007.

[192] H. G. Silber and K. F. McCoy, "Efficiently computed lexical chains as an intermediate representation for automatic text summarization," *Computational Linguistics*, vol. 28, no. 4, pp. 487–496, 2002.

[193] K. Sparck Jones, "A statistical interpretation of term specificity and its application in retrieval," *Journal of Documentation*, vol. 28, pp. 11–21, 1972.

[194] K. Sparck Jones, "Automatic summarizing: factors and directions," in *Advances in Automatic Text Summarization*, pp. 1–12, MIT Press, 1998.

[195] J. Steinberger, M. Poesio, M. A. Kabadjov, and K. Jeek, "Two uses of anaphora resolution in summarization," *Information Processing and Management*, vol. 43, no. 6, pp. 1663–1680, 2007.

[196] A. Stolcke and E. Shriberg, "Statistical language modeling for speech disfluencies," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 405–408, 1996.

[197] J.-T. Sun, D. Shen, H.-J. Zeng, Q. Yang, Y. Lu, and Z. Chen, "Web-page summarization using clickthrough data," in *Proceedings of the Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 194–201, 2005.

[198] W. tau Yih, J. Goodman, L. Vanderwende, and H. Suzuki, "Multi-document summarization by maximizing informative content-words," in *Proceedings of the International Joint Conference on Artificial Intelligence*, pp. 1776–1782, 2007.

[199] S. Teufel, "Task-based evaluation of summary quality: describing relationships between scientific papers," in *Proceedings of the NAACL Workshop on Automatic Summarization*, pp. 12–21, 2001.

[200] S. Teufel and M. Moens, "Summarizing scientific articles: experiments with relevance and rhetorical status," *Computational Linguisics.*, vol. 28, no. 4, pp. 409–445, 2002.

[201] S. Teufel and H. van Halteren, "Evaluating information content by factoid analysis: Human annotation and stability," in *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, 2004.

[202] D. R. Timothy, T. Allison, S. Blair-goldensohn, J. Blitzer, A. elebi, S. Dimitrov, E. Drabek, A. Hakim, W. Lam, D. Liu, J. Otterbacher, H. Qi, H. Saggion, S. Teufel, A. Winkel, and Z. Zhang, "MEAD — a platform for multidocument multilingual text summarization," in *Proceedings of the International Conference on Language Resources and Evaluation*, 2004.

[203] A. Tombros and M. Sanderson, "Advantages of query biased summaries in information retrieval," in *Proceedings of the Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 2–10, 1998.

[204] S. Tucker, N. Kyprianou, and S. Whittaker, "Time-compressing speech: ASR transcripts are an effective way to support gist extraction," in *Proceedings of the International Workshop on Machine Learning for Multimodal Interaction*, pp. 226–235, 2008.

[205] S. Tucker and S. Whittaker, "Temporal compression of speech: An evaluation," *IEEE Transactions on Audio, Speech and Language Processing*, vol. 16, pp. 790–796, 2008.

[206] J. Turner and E. Charniak, "Supervised and unsupervised learning for sentence compression," in *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, (Ann Arbor, Mi.), pp. 290–297, June 2005.

[207] A. Turpin, Y. Tsegay, D. Hawking, and H. E. Williams, "Fast generation of result snippets in web search," in *Proceedings of the Annual International*

ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 127–134, 2007.

[208] E. Tzoukermann, S. Muresan, and J. L. Klavans, "GIST-IT: Summarizing email using linguistic knowledge and machine learning," in *Proceedings of the Workshop on Human Language Technology and Knowledge Management*, pp. 1–8, 2001.

[209] J. Ulrich, G. Murray, and G. Carenini, "A publicly available annotated corpus for supervised email summarization," in *Proceedings of the AAAI EMAIL Workshop*, pp. 77–87, 2008.

[210] UMLS, "UMLS Knowledge Sources," National Library of Medicine, Bethesda, Maryland, 9th edition, 1998.

[211] H. van Halteren and S. Teufel, "Examining the consensus between human summaries: Initial experiments with factoid analysis," in *Proceedings of the HLT-NAACL Workshop on Automatic Summarization*, 2003.

[212] L. Vanderwende, H. Suzuki, C. Brockett, and A. Nenkova, "Beyond SumBasic: Task-focused summarization with sentence simplification and lexical expansion," *Information Processing and Managment*, vol. 43, pp. 1606–1618, 2007.

[213] R. Varadarajan and V. Hristidis, "A system for query-specific document summarization," in *Proceedings of the ACM Conference on Information and Knowledge Management*, 2006.

[214] A. Waibel, M. Bett, and M. Finke, "Meeting browser: Tracking and summarizing meetings," in *Proceedings of the DARPA Broadcast News Workshop*, pp. 281–286, 1998.

[215] S. Wan and K. McKeown, "Generating state-of-affairs summaries of ongoing email thread discussions," in *Proceedings of the International Conference on Computational Linguistic*, 2004.

[216] X. Wan and J. Yang, "Improved affinity graph based multi-document summarization," in *Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics, Companion Volume: Short Papers*, pp. 181–184, 2006.

[217] R. Weischedel, J. Xu, and A. Licuanan, "A hybrid approach to answering biographical questions," in *New Directions In Question Answering*, (M. Maybury, ed.), pp. 59–70, 2004.

[218] M. J. Witbrock and V. O. Mittal, "Ultra-summarization (poster abstract): A statistical approach to generating highly condensed non-extractive summaries," in *Proceedings of the Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 315–316, 1999.

[219] F. Wolf and E. Gibson, "Paragraph-, word-, and coherence-based approaches to sentence ranking: A comparison of algorithm and human performance," in *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, pp. 383–390, 2004.

[220] K.-F. Wong and W. Wu, Mingli sand Li, "Extractive summarization using Supervised and semi-supervised learning," in *Proceedings of the International Conference on Computational Linguistic*, pp. 985–992, 2008.

[221] S. Xie and Y. Liu, "Using corpus and knowledge-based similarity measure in Maximum Marginal Relevance for meeting summarization," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 4985–4988, 2008.

[222] E. Yamangil and S. M. Shieber, "Bayesian synchronous tree-substitution grammar induction and its application to sentence compression," in *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, pp. 937–947, 2010.

[223] J. Yang, A. Cohen, and W. Hersh, "Automatic summarization of mouse gene information by clustering and sentence extraction from MEDLINE abstracts," in *Proceedings of the AMIA Annual Symposium*, pp. 831–835, 2007.

[224] S. Ye, T.-S. Chua, M.-Y. Kan, and L. Qiu, "Document concept lattice for text understanding and summarization," *Information Processing and Management*, vol. 43, no. 6, pp. 1643–1662, 2007.

[225] W. Yih, J. Goodman, L. Vanderwende, and H. Suzuki, "Multi-document summarization by maximizing informative content-words," in *Proceedings of the International Joint Conference on Artificial Intelligence*, pp. 1776–1782, 2007.

[226] D. Zajic, B. J. Dorr, J. Lin, and R. Schwartz, "Multi-candidate reduction: Sentence compression as a tool for document summarization tasks," *Information Processing and Management*, vol. 43, no. 6, pp. 1549–1570, 2007.

[227] K. Zechner, "Summarization of spoken language - challenges, methods, and prospects," *Speech Technology Expert eZine*, January 2002.

[228] K. Zechner and A. Lavie, "Increasing the coherence of spoken dialogue summaries by cross-speaker information linking," in *Proceedings of the NAACL Workshop on Automatic Summarization*, pp. 22–31, 2001.

[229] K. Zechner and A. Waibel, "Minimizing word error rate in textual summaries of spoken language," in *Proceedings of the North American chapter of the Association for Computational Linguistics Conference*, pp. 186–193, 2000.

[230] J. Zhang, H. Y. Chan, and P. Fung, "Improving lecture speech summarization using rhetorical information," in *Proceedings of the IEEE Automatic Speech Recognition and Understanding Workshop*, pp. 195–200, 2007.

[231] J. Zhang and P. Fung, "Speech summarization without lexical features for Mandarin broadcast news," in *Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics; Companion Volume, Short Papers*, pp. 213–216, 2007.

[232] L. Zhou and E. Hovy, "A web-trained extraction summarization system," in *Proceedings of the Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology*, pp. 205–211, 2003.

[233] L. Zhou and E. Hovy, "On the summarization of dynamically introduced information: Online discussions and blogs," in *Proceedings of AAAI Spring Symposium 2006*, 2006.

[234] L. Zhou, M. Ticrea, and E. Hovy, "Multi-document biography summarization," in *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pp. 434–441, 2004.

[235] X. Zhu and G. Penn, "Evaluation of sentence selection for speech summarization," in *Proceedings of the RANLP Workshop on Crossing Barriers in Text Summarization Research*, pp. 39–45, 2005.

[236] X. Zhu and G. Penn, "Summarization of spontaneous conversations," in *Proceedings of the Annual Conference of the International Speech Communication Association*, pp. 1531–1534, 2006.