e-Learning for English based on Multimedia Database and Internet

Ying-Hong Wang

Department of Computer Science & Information Engineering
Tamkang University
Tamshui, Taiwan, R.O.C.
E-mail: inhon@mail.tku.edu.tw

Abstract

In this time of Internet delivery, learning through Internet will be popular and enhance the efficiency of teaching. This paper presents an Internet-based distance learning system for English learning through multimedia database and Internet technologies, it is called "multimedia English corpus". It includes two major learning functions. One of them provides Articles, Dialogs, and Videos databases in English. An English learner can study English writing, reading, and listing by Web browser to connect the Corpus server. In the system, "semantic query" and "Link grammar annotation" are applied. It can promote the query level from keyword-base and content-based query to semantic level. These skills of "semantic query" and "link grammar" have been used to construct the English multimedia corpus system. The main function of this system is to query the English sentence pattern by keywords from the English multimedia corpus. And the other function is to detect the grammar error in the sentence, which is written by student. It does not only provide learners to find their mistakes of English grammar, but also the teachers can understand the most frequent mistakes made by learners through the records of this corpus.

Key Words **:** Distance Learning, Semantic-Level Query, Corpus, Multimedia, Link Grammer

1. Introduction

Traditional relation databases always store the character and numeral data. One of the main functions deals with basic character and numeral data, for examples are “students’ data of schools”, “employee and financial affair data of the companies.” Following the development of computer technologies, multiple types of data, such as image, audio, video, and hypermedia document, are applied in the representation of computer information. Thus the technologies of database need to support multimedia data, called multimedia database. There are many issues in the multimedia database, such like the content-based retrieval [6, 19], shape detection, and object recognition [7, 12], …etc. At the same time, there are many application domains of multimedia database are promoted by these technique issues. Many examples of distance learning, digital television, distance medical and others are implemented based on those techniques. This paper will present an Internet-based English learning environment applied the technologies of multimedia database.

Learning English is always a big problem for Chinese students. Some experts suggest that the best way to improve the efficient of English learning is to get a good English study environment and practice English in several different ways to be a satisfying learning experience. In this viewpoint, it is appropriate that

teachers should consider how to provide students a better environment for English learning. Distance learning based on Internet is one of the best solutions to build a English learning environment.

Currently, the highly distributed computing environment has been supported and applied. It is particularly important that networking services act as tutoring tools. Different kinds of innovative approaches have been investigated these years. Many advocates of computer-mediated education emphasize its positive aspects and undertake the kind of work that it requires for students. There are some tutoring systems of English learning based on computing environment have been developed by a number of academic research group [20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32].

By the way, as students learning a foreign language with computing devices, they might suffer from a lot of difficulties and have many problems. Some articles discuss whether students' frustrations inhibit their educational opportunity [21, 22, 24]. However, students' frustrations were found in three interrelated sources: lack of prompt feedback, ambiguous instructions on the web, and technical problems.

As mentioned above, it encourages us a great deal of inspirations to develop an easy-used English learning system. This means that several preparations, which are certainly relevant to English learning, are necessary for the proposed system. While dealing with natural language processing, the system is constructed based on the Link Grammar. In order to keep some differences from the conventional approaches, some modification will be made to improve its ability.

Thus, there is a new kind of Corpus system is presented for the students. One of the functions of the proposed English-learning corpus, it is built to store the mistakes that students usually make. All the errors that students will probably make should be constructed in advance. There are more information for each type of mistakes, such as correcting suggestions and error descriptions. The motivation that the Corpus constructed can not only act as suggestion provider but also record the mistakes that Chinese students might happen in English. The corpus can provide students, teachers and researchers a good tool in the Internet. Another function of this multimedia corpus, user can find some example sentences from movies scripts, live dialogs and textbook reading and etc. Each sentence from living dialog and the caption of the movies is parsed and analyzed by link grammar, and some special tags are included, and then stored together in this proposed database system. The system provides users to query English sentence pattern on-line through Internet. Consequently, the system provides to the query methodology base on the semantic retrieval.

In this paper, English sentences will be analyzed firstly. We can divide the sentence pattern into nine classes. Secondly, each sentence pattern is parsed by link grammar, and the phenomenon is that every pattern has their special tag set. Consequently, the system can construct the query engine using these important tag sets.

In generally, the learning corpus stores the single sentences or vocabularies and provides user to query by keywords [20, 21, 22, 23, 25, 26, 28, 29, 30, 31, 32]. This paper will discuss how to use the appended information, the special tag sets, generated by Link grammar to promote the query function from keyword-based to semantic-based.

This paper has surveyed several web sites about learning corpus [20, 21, 22, 23], and study several papers beside link grammar [8, 10, 15, 16]. Furthermore, Schulenburg proposed the processing with realistic feedback [10], Brill's machine learning and automatic linguistic analysis [14], Chang presented Automatic linguistic resolution: framework and applications [4], and the other important manuscripts have been referenced [1, 8, 9, 16, 18].

In section 2, this paper discusses the related works, which are about link grammar and XML. The system analysis for the proposed learning functions will be introduced in the section 3, The section 4 presents the architecture of the proposed system. And the conclusion and the discussions of future research in the section 5.

2. The Related Work

2.1 Link Grammar Terminology

Link grammar is an English grammar parser system proposed by school of computer science of CMU. Link grammar is a context-free formalism for the description of natural language [11].

A link grammar consists of a set of ***words***, called the terminal symbols of the grammar, each of which has a ***linking requirement***. The linking requirements of each word are contained in a ***dictionary***. To illustrate the linking requirements, Figure 1 shows a simple dictionary for the words *a/the*, *cat/mouse*, *John*, *ran*, *chased*. The diagram shows the link requirement is represented on the above of each word.

Each of the intricately shaped boxes has a label called ***connector***. A pair of compatible

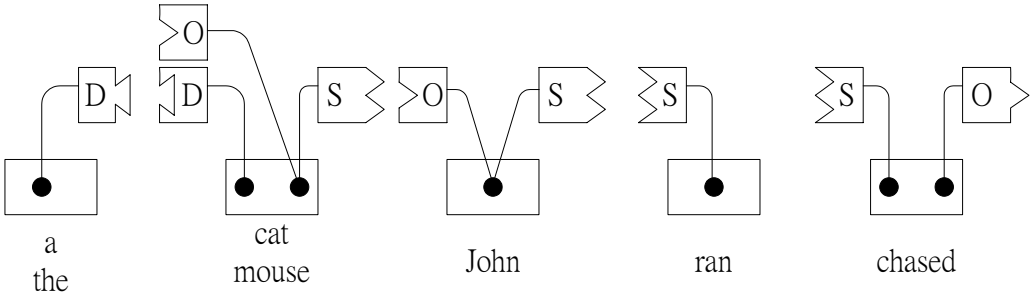

Figure 1. Words and connectors in the dictionary

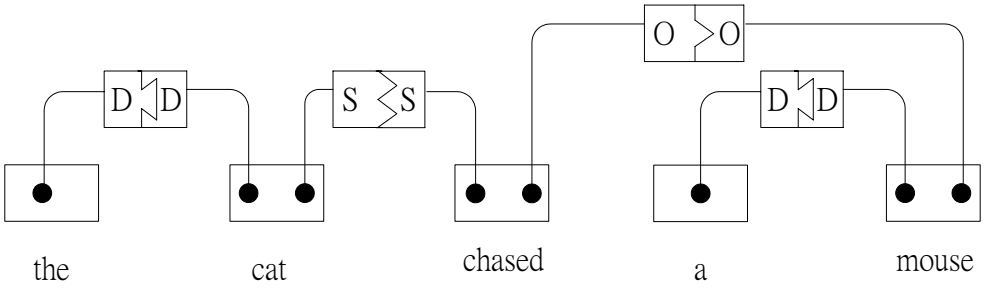

Figure 2. All linking requirements are satisfied to form a linkage

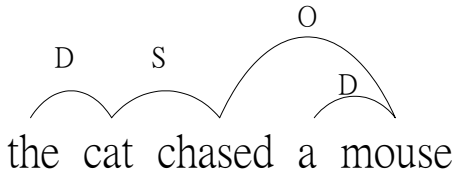

Figure 3. The simplified form of Fig. 2

Table 1. The words and linking requirements in a dictionary

words	formula
a / the	D+
cat / mouse	D- & (O- or S+)
John	O- or S+
ran	S-
chased	S- & O+

connectors will match to each other, and only one of the connectors attached to a given black dot must be satisfied. In the figure 2, it shows how the linking requirements are satisfied in the sentence, "*The cat chased a mouse*."

The sentence consists of five words. We may think the linkage as a graph. And the words can be treated as vertices, which are connected by arcswith labels on them, so that the graph is connected and planar. The labeled arcs connecting words to other words on their lefts or rights are called ***links.*** A valid parse is called a ***linkage.*** Here is the simplified form of the diagram showing that *the cat chased a mouse* is part of this language.

The following abridged dictionary, presented as Table 1, encodes some linking requirements of the example above.

The linking requirement for each word is expressed as a formula involving the operators *&*, and *or*, *parentheses*, and *connector names*. The + or – suffix on a connector name indicates the direction in which the matching connector must lie. The farther left a connector is in the expression, the nearer the word to which it connects must be.

A sequence of words is a sentence of the language defined by the grammar if there exists a way to draw links among the words so as to satisfy the formula of each word, and the following ***meta-rules:***

Planarity: The links are drawn above the sentence and do not cross.

Connectivity: The links suffice to connect all the words of the sequence together.

Ordering: When the connectors of a formula are traversed from left to right, the words to which they connect proceed from near to far. In other words, consider a word, and consider two links connecting that word to word to its left.

The link connecting the near word (the shorter link) must satisfy a connector appearing to the left (in the formula) of that of the other word. Similarly, a link to the right must satisfy a connector to the left (in the formula) of a longer link to the right.

Exclusion: No two links may connect the same pair of words.

The use of formula to specify a link grammar dictionary is convenient for creating natural language grammars, but it is cumbersome for mathematical analysis of link grammar, and in describing algorithms for parsing link grammar. Here is a different way of expressing a link grammar called ***disjunctive form***. In disjunctive form, each word of the grammar has a set of disjuncts associated with it. A disjunct will be denoted:

$$\mathbf{((L_1, L_2, \ldots, L_m)(R_n, R_{n-1}, \ldots, R_1))}$$

Where L_1, L_2, …, L_m are the connectors that must connect to the left, and R_n, R_{n-1}, …, R_1 are connectors that must connect to the right.

It is easy to see how to translate a link grammar in disjunctive form to one in standard form. It can be done simply by rewriting each disjunct as

$$\mathbf{(L_1\&L_2\&\ldots\&L_m\&R_1\&\ R_2\&\ldots\&R_n)}$$

and combining all the disjuncts together with the *or* operator to make an appropriate formula.

It is also easy to translate a formula into a set of disjuncts. This is done by enumerating all ways that the formula can be satisfied. For example, the formula

(A- or ( )) & D- & (B+ or ( )) & (O- or S+)

corresponds to the following eight disjuncts:

((A,D) (S,B))
((A,D,O) (B))
((A,D) (S))
((A,D,O) ( ))
((D) (S,B))
((D,O) (B))
((D) (S))
((D,O) ( ))

Any of the eight disjuncts may be used in some linkages.

2.2 XML Terminology

The extraordinary growth of the World Wide Web has been fueled by the ability it gives authors to easily and cheaply distribute electronic documents to an international audience. As Web documents have become larger and more complex, however, Web content providers have begun to experience the limitations of a medium that does not provide the extensibility, structure, and data checking needed for large-scale commercial publishing. To address the requirements of commercial Web publishing and enable the further expansion of Web technology into new domains of distributed document processing, the World Wide Web Consortium (W3C) has developed an eXtensible Markup Language (XML) for applications that require functionality beyond the current Hypertext Markup Language (HTML).

XML is not a single, predefined markup language: it's a metalanguage – a language for describing other languages – which lets you design your own markup. It can do this because it's written in SGML, the international standard meta-language for markup.

SGML had been used for many years in sophisticated and highly complex publishing applications. Despite the fact that HTML, an application of SGML, is the standard for Web publishing, SGML itself never got any traction with the Web development community. This is primary due to the complexity of SGML and the resulting overhead required using it.

XML is an abbreviate version of SGML, to make it easier for authors to define their own document types, and to make it easier for programmers to write programs to handle them. It omits the more complex and less-used parts of SGML in return for the benefits of being easier to write applications for, easier to understand, and more suited to delivery and interoperability over the Web. But it is still SGML, and XML files may still be parsed and validated the same as any other SGML file.

XML is a low-level syntax for representing structured data. You can use this simple syntax to support a wide variety of applications. This idea is put across in a simplistic way in the diagram below, which shows how XML now underpins a number of Web mark-up languages and applications.

XML-based data does not contain information about how data should be displayed. To display the data, either the Web server or the

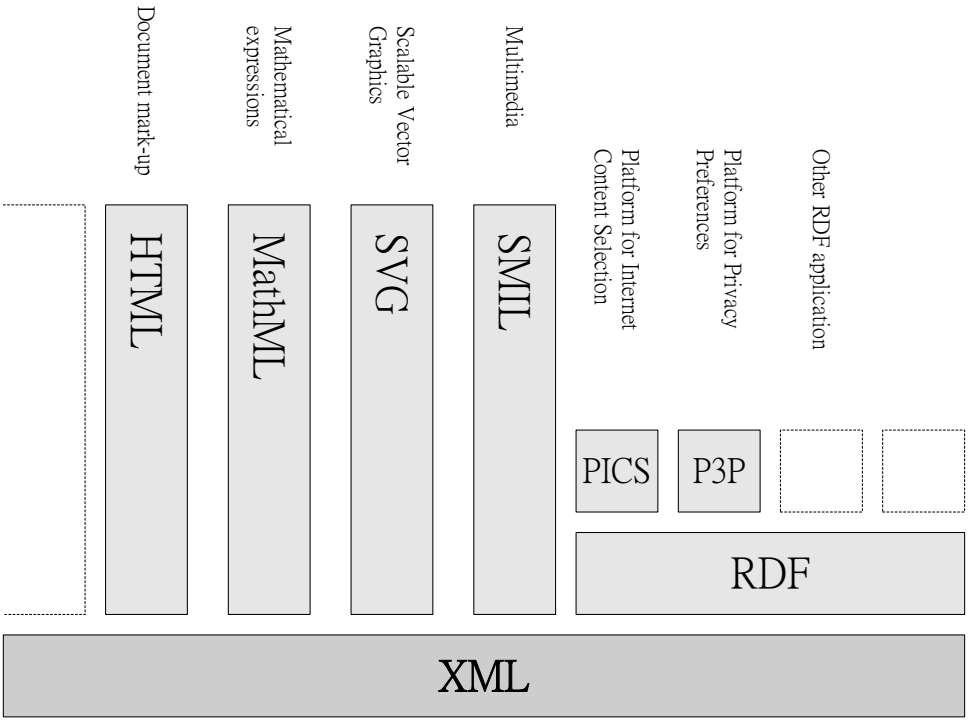

Figure 4. Applications with XML

Web browser will need to move XML-based data into HTML for display, or alternatively transform XML data into HTML. In the future, developers will be able to use XSL to generate HTML without complex scripting. XSL is a transformation language that defines rules for mapping structured XML data to HTML, or other display formats, using an XSL Processor.

This system will be based on those above two kinds of technology, to build a new English learning environment.

3. System Analysis

In this paper, the proposed system provides two kinds of learning functions. We can consider that there are two kinds of sub-systems in the system. The names of the two sub-systems are English pattern query system and English sentence error-detect system, respectively. The detail will be discussed in the following.

3.1 The English Pattern Query System

To develop this system, English sentence patterns are analyzed firstly. In general analysis, there are five English sentence patterns shown as follows:

(1) Simple sentence pattern
(2) Negate sentence pattern
(3) Interrogative sentence pattern
(4) WH question sentence pattern
(5) Imperative sentence pattern

And then they can be divided more detail by tense and they and their examples are presented follows:

Simple pattern

The simple present tense: He writes a letter everyday.

The simple past tense: He wrote a letter yesterday.

The simple future tense: He will write a letter tomorrow.

Perfective pattern

The present perfective tense: He has written a letter.

The past perfective tense: He had written a letter when I came.

The future perfective tense: He will have written a letter before I come.

Continuous pattern

The present continuous tense: He is writing a letter now.

The past continuous tense: He was writing a letter when I came.

The future continuous tense: He will be writing a letter when I come.

Perfective proceed pattern

The present perfective continuous tense: He is writing a letter now.

The past perfective continuous tense: He was writing a letter when I came.

After parsed by link grammar, that can find the label relation of those sentence patterns. That shown by table 2~5 as follow.

Table 2. Simple sentence pattern and link grammar tags

		Active voice	Passive voice
Simple pattern	present	Ss or Sp	Ss or Sp + Pv
	past	Ss or Sp	Ss or Sp + Pv
	future	Ss or Sp + I	Ss or Sp + Ix + Pv
Continuous pattern	present	Ss or Sp + PP	Ss or Sp + ppf + Pv
	past	Ss or Sp + PP	Ss or Sp + ppf + Pv
	future	Ss + If + PP	Ss or Sp + If + ppf + Pv
Proceed pattern	present	Ss or Sp + Pg	Ss or Sp + Pg + Pv
	past	Ss or Sp + Pg	Ss or Sp + Pg + Pv
	future	Ss or Sp + Ix + Pg	none
Perfective continuous pattern	present	Ss or Sp + ppf + Pg	
	past	Ss or Sp + ppf + Pg	
Illustrate: Ss and Sp: connects subject nouns to finite verbs I: connects infinitive verb forms to certain words such as modal verbs and "to" PP: connects forms of "have" with past participles. PV: connects forms of the verb "be" to past participles. Pg: connects forms of the verb "be" to present participles. ppf: connects forms of the verb "be" to past participles “been”.			

Table 3. Negate sentence pattern and link grammar tags

		Original＋not	“Not” condensation
Simple pattern	present	Ss or Sp + N	Ss or Sp + I*d
	past	Ss or Sp + N	Ss or Sp + I*d
	future	Ss or Sp + N + I	Ss or Sp + I
Continuous pattern	present	Ss or Sp + N + PP	Ss or Sp + PP
	past	Ss or Sp + N + PP	Ss or Sp + PP
	future	Ss or Sp + N + If + PP	Ss or Sp + If + PP
Proceed pattern	present	Ss or Sp + N + Pg	Ss or Sp + Pg
	past	Ss or Sp + N + Pg	Ss or Sp + Pg
	future	Ss or Sp + N + Ix + Pg	Ss or Sp + Ix + Pg
Perfective continuous pattern	present	Ss or Sp + N + ppf + Pg	Ss or Sp + ppf + Pg
	past	Ss or Sp + N + ppg +Pg	Ss or Sp + ppf + Pg
Illustrate: There must be label “N” in the negate sentence pattern. N: connects the word "not" to preceding auxiliaries.			

Table 4. Yes/No question sentence and link grammar tags

		Yes/No question sentence
Simple pattern	present past future	Qd + SIs + I*d Qd + SIs + I*d Qd + SIs + I
Continuous pattern	present past future	Qd + SIs + PP Qd + SIs + PP Qd + SIs + If + PP
Proceed pattern	present past future	Qd + SIs + Pg Qd + SIs + Pg Qd + SIs + Ix + Pg
Perfective continuous pattern	present past	Qd + SIs + ppf + Pg Qd + SIs + ppf + Pg
Illustrate: Yes/no question sentence pattern must be star by label “Qd”. Qd: is used in questions.		

Table 5. WH question sentence and link grammar tags

		what	Where and how
Simple pattern	present past future	Wq + Sid + I*d + Bsw Wq + Sid + I*d + Bsw Wq + SIs + I + Bsw	Wq + Q + SIs + I*d Wq + Q + SIs + I*d Wq + Q + SIs + I*d
Continuous pattern	present past future	Wq + SIs + I + Bsw Wq + SIs + I + Bsw	Wq + Q + SIs + PP Wq + Q + SIs + PP
Proceed pattern	present past future	Wq + SIs + Pg + Bsw Wq + SIs + Pg + Bsw	Wq + Q + SIs + PP Wq + Q + SIs + PP
Perfective continuous pattern	present past	none	none
Illustrate: WH question sentence must be star by label “Wq”. Wq: connects the subjects of main clauses to the most questions (except yes-no questions). most questions (except yes-no questions). Sid: connects subject nouns to finite verbs in cases of subject-verb inversion Bsw: connects auxiliary verb will、have/has/had to past participle.			

After parsed these English sentence patterns by link grammar, each pattern has their special kinds of label constituent type. It is easily using the characteristic to parse each sentence in the system. And then the result label set is stored in the database. For example, (Ss, ppf, pg) is a label set after parsed by link grammar. And stored in the database. Then user can use query language to query the English sentence pattern what they want to read or learn. In the beginning, analysis only one kind of pattern is ensuring pattern is pure, and then the system can be upgraded to deal with multi-pattern in the future. The system can also deal with multi query in several kinds of pattern. But there are some difficulties, we will discuss them in latter section.

3.2 English Sentence Error Detect System

In this section, we illustrate the Enhanced

Link Grammar based on the dictionary with modified disjuncts. Firstly, this section introduces how the words link together. And the second part goes for the workings of the Enhanced Link Grammar.

3.2.1 Modified Dictionary

All of the recognized words of link grammar are stored in its dictionary. The following diagram, Figure 5, shows what the dictionary looks like.

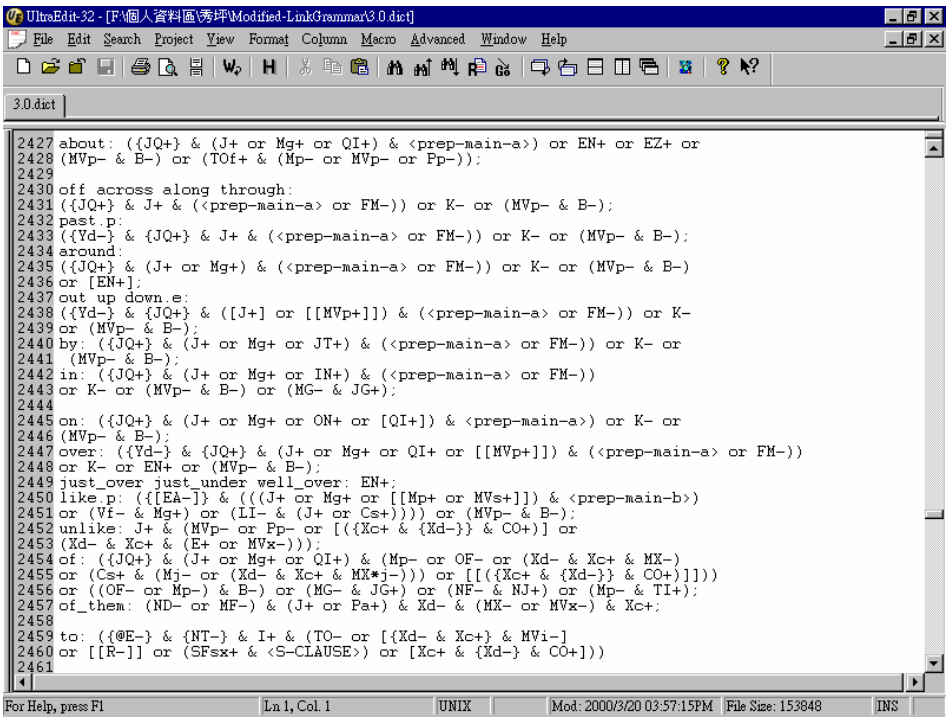

Figure 5. The original dictionary in the link grammar

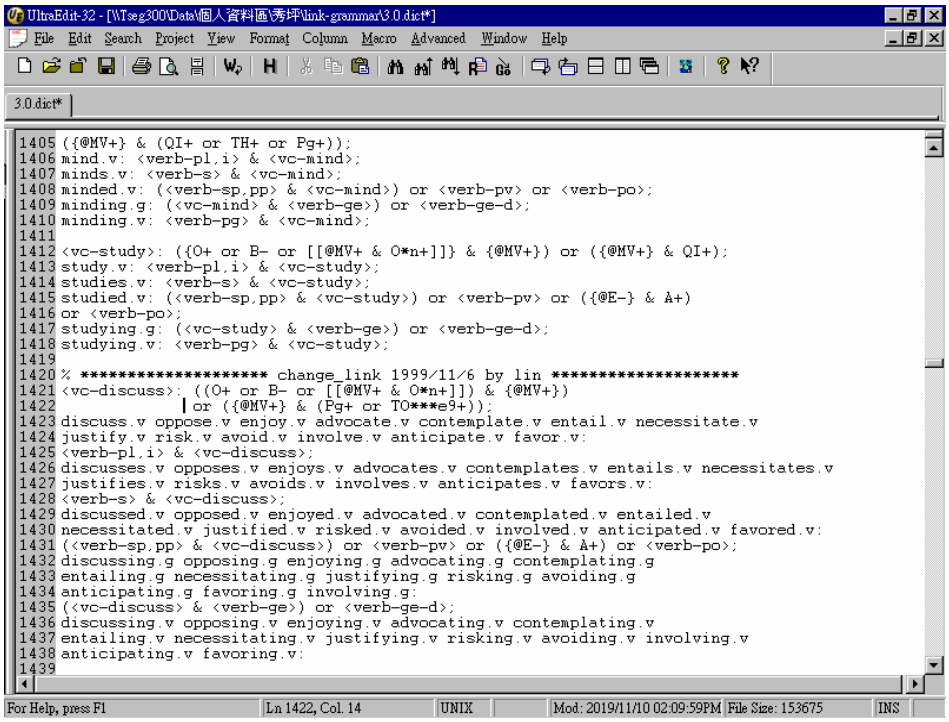

Figure 6. The modified dictionary

In general, the formula of each word is complicated. The first step for us is to analyze the word that we want to change its connectors. The characteristics of different formulas impose different constrains on the design of our work. Some changes maybe work for our modification, but a lot of them maybe not. It is worth choosing some simple words to test before going into the details.

For example, some verbs must be followed by infinitive *–to*, some must be followed by gerund, and some must be followed by participle. There are some verbs that can accept several different types. People say "We enjoy walking in the rain." but not "We enjoy to walk in the rain." By the way, we can change the link of the word "enjoy" to make it acceptable for infinitive *–to*.

Comparing with Figure 5, in Figure 6, we change part of the formula from "*(Pg)*" to "*(Pg or TO***e9+)*". By this method, the word "enjoy" can link to not only a gerund but also an infinitive *–to* with the additional connector "*TO***e9+*". The new connector contains a subscript "*****e9* " which means there would be an error of type 9.

And then, we type another two sentences "We enjoy swimming." and "We enjoy to swim." to the Link grammar with original dictionary and modified dictionary, respectively.

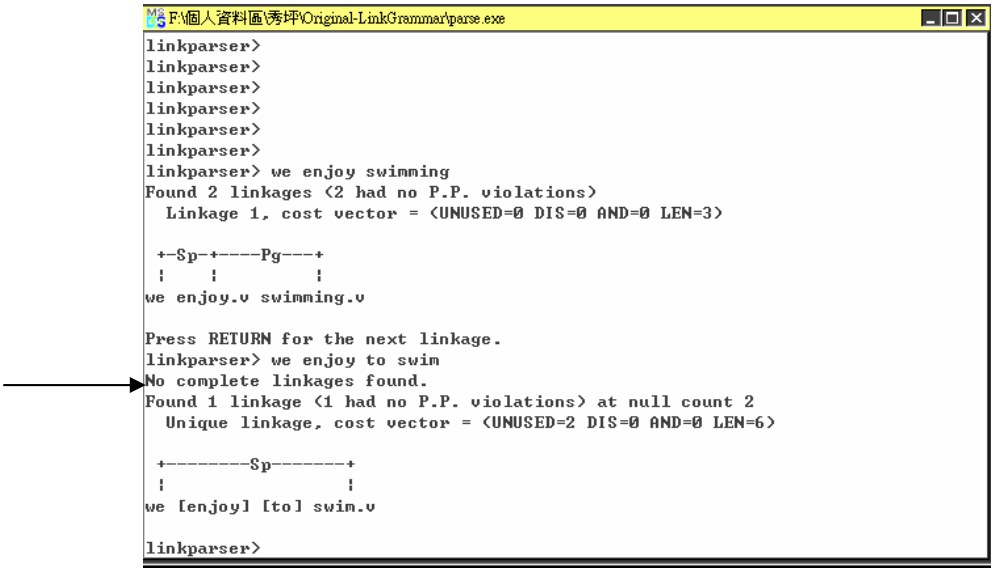

Figure 7. The original link grammar parser

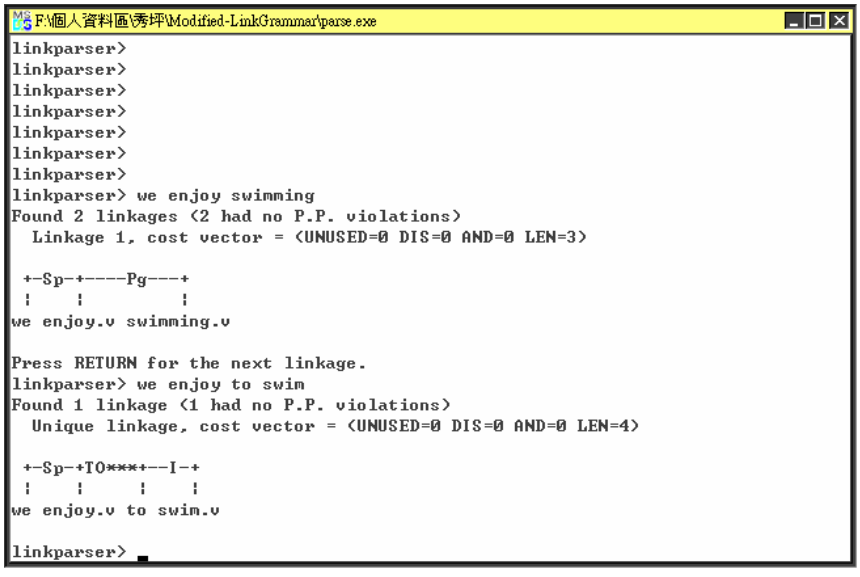

Figure 8. The modified link grammar parser

As Figure 7 presented, the link grammar with original dictionary can not accept the wrong sentence. An error message “No complete linkage found.” will appear.

When the same two sentences are parsed by the link grammar with modified dictionary, the parsing result is shown as below:

As shown in Figure 8, the word “enjoy” can link to either a gerund or an infinitive *–to*. Of course, there is more error tags that are designed for error detection. To sum up, we catalog ten types of errors. In the next section, the error tag itself will be discussed in detail, including the links that can and cannot be modified.

4. The System Architecture and Examples

4.1 The System Architecture

Corpus is a database that it is not only to provide dictionary but also to record the mistakes of learners, comments of teachers, and system feedback. The proposed English-learning corpus has positive study data, just like English essay, dialog, sentence, word, terms, attributes …etc. It can promote language teaching and research. In this paper, we will construct a multimedia corpus. Except the original data, it also stores many kinds of multimedia data, for example voice, teaching films, dialog audio, movies video, …etc. in the database. Therefore, it can provide a vivid learning environment to students and attract more learners to use it. This system, presented in figure 9 and 10, is a multimedia corpus that provides four main functions of the basic way for language learning; they are listen, speech, read and write. User can read essay, listen standard English pronounce, writing compositions, and practice dialog and recitation a text. The system can also provide research issues to linguistic researchers through collecting the results of learners' situation. Philologists can analysis some mistakes what the non-native student usually make. It is a very important information to linguistic researchers.

User can use this system over the Internet to access the multimedia data from the multimedia learning corpus. Figure 9 shows a user can reach this system anywhere.

There are two sub systems in this proposed system, they are English pattern query system and English sentence error detect system, respectively. Both of them share the same database. And a learner corpus is built by XML.

The architecture of the “English pattern query system” is shown in the figure 11.

The data must be parse by the link grammar with original dictionary and produced the link labels. And then system will analysis and filters the link labels. Then the link labels will be stored into the database. When a user want to query some sentence pattern. They can use the interface of the

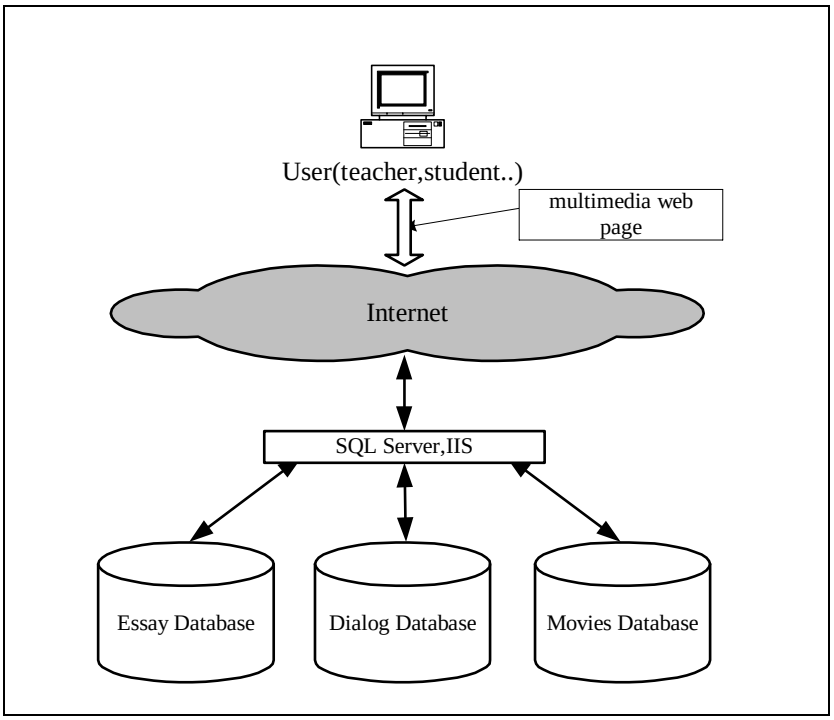

Figure 9. System architecture overview

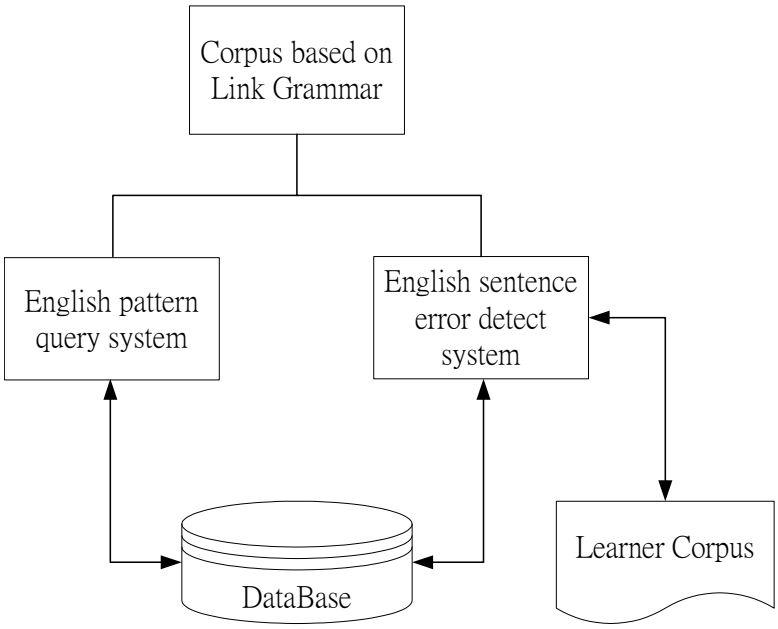

Figure 10. System architecture

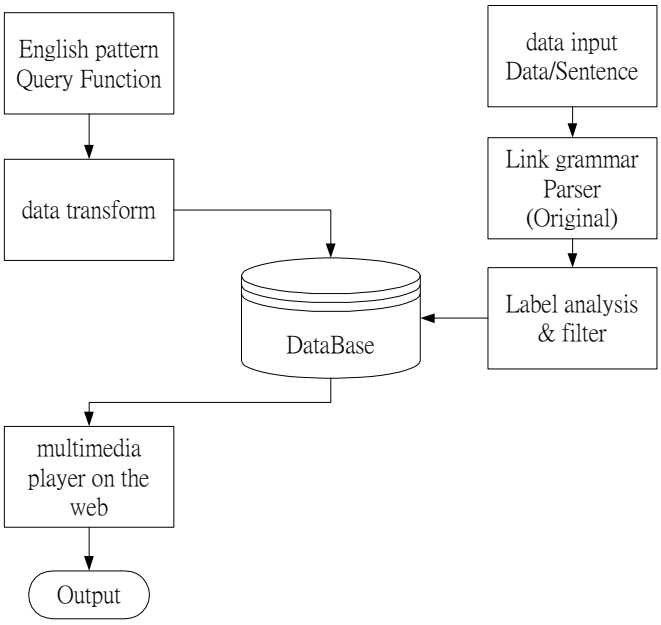

Figure 11. English pattern query system architecture

system to select the pattern what they want. And then the system will transform the query into the data-type that the database can readable. And find the same link label in the database. If the system can find some record in the database. The system will retrieve the data and send them to the multimedia player on the web. The advantage is that user can listen several times until they really clear understand the correct pronunciation of the sentence, and can make the user understand that they can using the sentence in what situation. That can make the user know the relation of the sentence and the circumstance in really world.

And the other sub-system is “English sentence error detect system”. The architecture of the English sentence error detect system is shown in the figure 12.

The user can input the sentence what they write into the system. And the link grammar with modified dictionary will parse the sentence, after

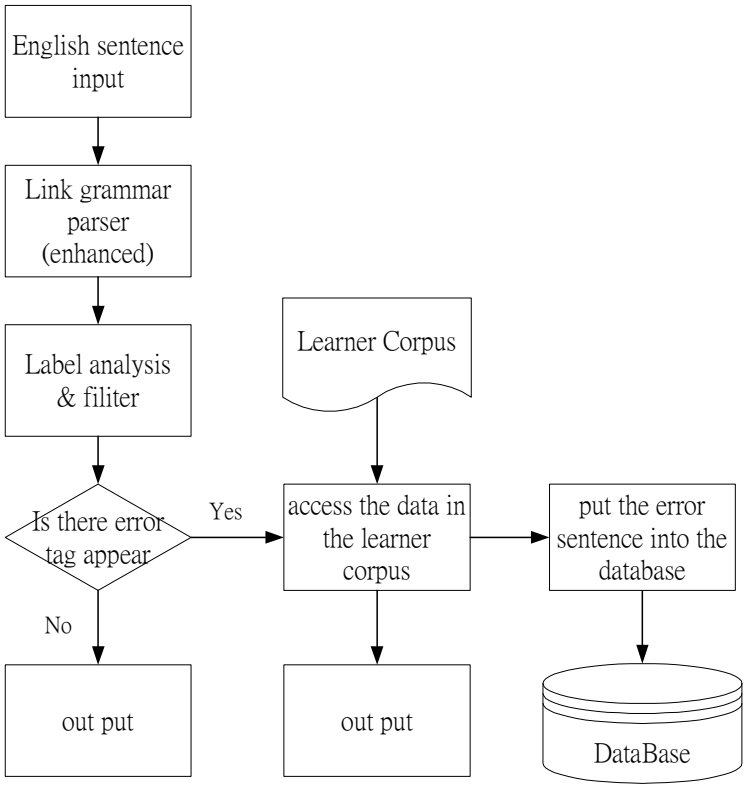

Figure 12. English sentence error detect system architecture

analysis and filter the link labels if error tags are detect by the system. The system will find the description and error examples of the learner corpus and then the system will show that information on the web page. If the system didn't understand what kind of the error is. Maybe the system did not defined yet or the system can not deal with. Those sentences will be stored in the database. Philologists can analysis these kinds of the sentences. And then they can find the new rules to enhance this system.

This system construction in SQL Server 7.0 and that have three main database, essay, dialog and movies, as previous figure 9 shown. In this paper, only essay database is discussed. There is data schema of essay database：

Table 6. Essay_meta

Field	datatype	Length
Essay_no	Int	4
Title	varchar	100
From_book	varchar	50
Created_date	datetime	8
Keywords	varchar	20
Class_level	int	4
Type	int	4
Authors	varchar	20
Annotation	int	4
Essay	varchar	8000
Voice	carchar	50
No_para	int	4

Keyword: the keywords of the essay.
class_level: the hard degree of essay.
type: the type of the essay issue.
annotation: the annotation by manual.
essay: the whole sentence of the essay.
voice: the audio files of the essay.
no_para: the number of paragraph.

And the schema of the essay sentence Essay sentence

Table 7. Essay_sentence

Field	datatype	Length
Sentence_no	int	4
Essay_no	int	4
Class_level	int	4
Sentence	varchar	200
Chinese	varchar	200
Voice	varchar	50
Annotation	int	4
Paragraphs	int	4
Link_g	varchar	4

Chinese: the translation to Chinese.
annotation: the annotation by manual.
paragraphic: the nth paragraphic.
Voice: the audio files of the sentence.
Link_g: the label set.

Figure 13. English pattern query system interface

4.2 System Examples

In this section, this paper shows the operation examples of the system. It is a web-based user interface. The figure 13 shows the interface of the English pattern query system.

Users have an easy use interface; they just need to drop down the menu and chose the pattern that they want to read. And then the system will return the query result with the multimedia data on the web browser. It is shown in figure 14.

User can see the movies and listen the audio through the hyper-media web browser. And the user can practice the conversation or declaim the essay several times on the Internet. This system also support keyword-based query. The user can also use this kind of query to retrieval the sentence.

Next, the figure 15 will show the interface of the English sentence error detect system.

In this system, users can input their sentence. And the system will check this input sentence is error or not based on the link grammar with modified dictionary. Figure 16 shows the example: when a user input a correct English sentence.

As showed in figure 17, here an error sentence is generated and the response of this system.

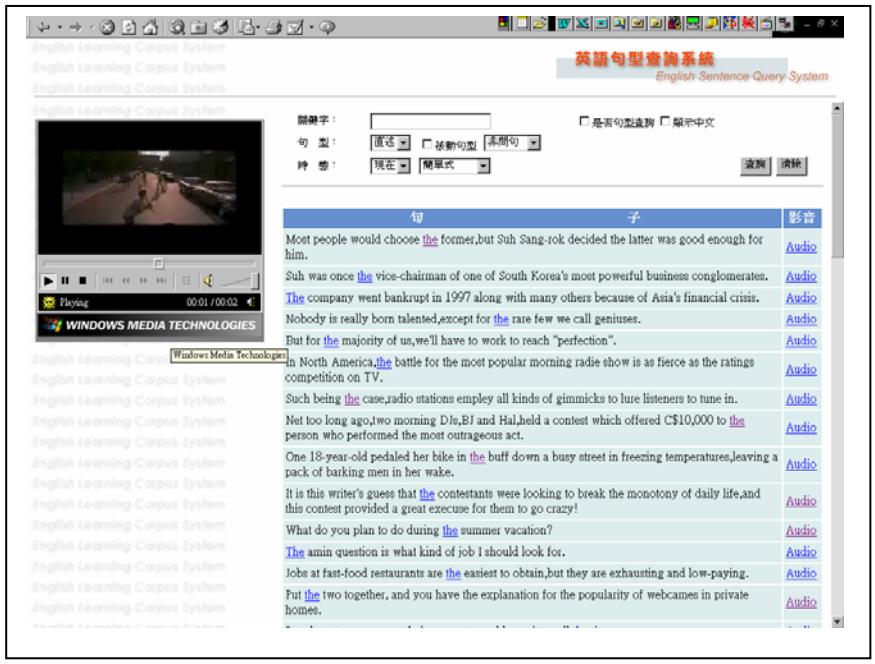

Figure 14. Query result

Figure 15. English sentence error detect system interface

Figure 16. Correct sentence result

Figure 17. Error sentence result

When the linkage of the user's input appears one error tag, the system will search the Learner Corpus and the correct and incorrect examples will be given to the user. Of course, the description also is included. Figure 18 shows the result.

5. Conclusion and Future Work

5.1 The Contribution and Remarks

As mentioned before, learning English is always a big problem for foreign students. After analysis and experiment, the systems can easily use the label set which store in the corpus parsed by link grammar. It will help user to find the English sentence pattern, when they are learning these patterns in junior or senior high school. Furthermore, this system promotes a high interaction multimedia system. And some linguistic experts suggest that the best way to improve English ability is to get a good English-learning environment and practice English in several different ways.

Figure 18. Error suggestion

The Link Grammar proposed by Daniel D.K. Sleator and Davy Temperley is a word based parsing mechanism. It is believed that the system should be a good grammar checker. However, the system proposed by the paper focuses on the fault tolerance ability, especially for non-native students. The contributions of the system come from that it can parse the sentences that the original Link Grammar cannot do.

By the way, we have made steady progress in the quality of our modified dictionary. It appears that the idea proposed by the paper can apply to other natural language processing and its corresponding applications.

The system provides not only query the English sentence pattern, but also can parse input sentences on-line. It is importance to philologist to analysis the sentences generated by learners, and they can easy to find the common or special mistakes. And then they can change the focus of teaching course in the future for the complete learn environment.

This system can provide a better environment for teacher and student. That can give a more interactive relationship to teachers and students. And it combines the humanism education with technology. The system can advance these two sciences for a better achievement.

5.2 Future Works

Generally speaking, words are a basic unit of communication common to all natural language texts and speech-processing activities. For English teaching, teachers always want to know what kinds of mistakes students may make. The system can also be extended to cover more scalable ranges according to different functional necessities. The descendants can develop other applications by using the idea of the paper.

In the other way, XML has the self-descriptive characteristic. Such a feature makes it more and more widely used for the data interchange in the world. The applications and research areas may become a popular subject in the future.

Reference

[1] Agius, IHarry W., Angelides, Marios C., “Developing Knowledge-Based Intelligent Multimedia Tutoring Systems Using Semantic Content- Based Modeling”, Artificial Intelligence Review 13: pp. 55 - 83 (1999).

[2] Al-Khatib, Wasfi, Francis Day, Y., Ghafoor Arif, Bruce Berra, P., “Semantic Modeling and Knowledge Representation in Multimedia Databases”, IEEE Transactions on Knowledge and Data Engineering, Vol. 11 NO. 1, January/February (1999).

[3] Beeferman, D., Berger, A. and Lafferty, J.,

"Text segmentation using exponential models", Proceedings of the Second Conference On Empirical Methods in NLP, Providence, RI (1997).

[4] Brill, E., "Machine learning and automatic linguistic analysis: the next step", Acoustics, Speech and Signal Processing, 1998. Proceedings of the 1998 IEEE International Conference on Volume: 2, pp. 1033 – 1036 (1998).

[5] Chan, S. W. K., "Automatic linguistic resolution: framework and applications", Systems, Man and Cybernetics, 1996, IEEE Internation Conference on Volume: 1, pp. 625 – 630 (1996).

[6] Chang, Shih-Fu, Chen, W., Meng, H. J., Sundaram, H., Di Zhong Circuits and Systems for Video Technology, "A Fully Automated Content-Based Video Search Engine Supporting Spatiotemporal Queries", IEEE Transactions on Volume: 8, pp. 602 – 615, 5, Sept. (1998).

[7] Chang, S. -K.; Hsu, A.,"Image Information systems: where do we go from here?" Knowledge and Data Engineering, IEEE Transactions on Volume: 4, pp. 431 – 442, 5, Oct. (1992).

[8] Chung, C. M., Wang, Yuan-Kai, Wang, Ying-Hong and Lin, Chih-Hao, "The multimeaia corpus based on link grammmar", National Computer Symposium 99, Vol. 1, pp. A-300 ~ 304.

[9] Chung, C. M., Wang, Yuan-Kai, Wang, Ying-Hong and Lin, Chih-Hao, "The Design of Multimedia Database for English Learning Application" proceeding of 2000 International Conference on Chinese Language Computing (ICCLC2000), pp. 77 ~ 81.

[10] Grinberg, D., Lafferty, J. and Sleator, D., "A robust parsing algorithm for link grammars", Proceedings of the Fourth International Workshop on Parsing Technologies, Prague and Karlovy Vary, pp. 111 – 125, September (1995).

[11] Lafferty, John., Sleator Daniel., and Temperley, Davy., **Grammatical Trigrams: A Probabilistic Model of Link Grammar** Presented to the 1992 AAAI Fall Symposium on Probabilistic Approaches to Natural Language.

[12] Myron, F., Harpreet, S., Wayne, N., Jonathan, A., Qian, H., Byron, D., Monika, G., Jim, H., Denis, L. and Dragutin, P., "Query by Image and Video Content: The QBIC System", Computer, IEEE Los Alamitos, CA. USA, Vol. 28, No. 9, pp. 23 – 32, Sept. (1995).

[13] Nakamura, Yuichi., Kanade, Takeo., "Semantic Analysis for Video Contents Extraction – Spotting by Association in New Video", ACM Multimedia Conference (1997).

[14] Schulenburg, D.,"Sentence processing with realistic feedback", Neural Networks, 1992. IJCNN., International Joint Conference on Volume: 4, pp. 661 – 666 (1992).

[15] Sleator, D., and Temperley, D., "Parsing English with a link grammar", technical report CMU-CS-91-196, Department of Computer Science, Carnegie Mellon University (1991).

[16] Wang, Yuan-Kai, Wang, Ying-Hong, Lin, Hsiu-Ping and Chao, Chih-Feng, "An English-Learning Application based on extended link grammar formalism", National Computer Symposium 99, Vol. 2, pp. B-306 ~ 310.

[17] Wu, Jian-Kang, "Content-Based Indexing of Multimedia Databases", IEEE Transactions on Knowledge and Data Engineering, Vol. 9, NO. 6, November/December (1997).

[18] Yamada, Chiro., Sumiyoshi, Hideki., Kim, Yeun-Bea and Shibata, Masahiro., "TV program skimming method using similarity of scene descriptions", SPIE Conference on multimedia Storage and Archiving Systems III, November (1998).

[19] Yoshitaka, A.; Ichikawa, T., "A survey on content-based retrieval for multimedia databases", Knowledge and Data Engineering, IEEE Transactions on Volume: 11, pp. 81 – 93, 1, Jan.-Feb. (1999).

[20] EFLWEB Home Page, http://www.eflweb.com/

[21] English grammar clinic, http://forums.educhat.com:8080/~grammar

[22] Foreign Languages for Travelers, http://www.travlang.com/languages/

[23] G.R. Sampson: SUSANNE Scheme, http://www.cogs.susx.ac.uk/users/geoffs/RSue.html

[24] **Interactive Javascript Quizzes for ESL Students:** http://www.aitech.ac.jp/~iteslj/quizzes/js/

[25] Language Dictionaries and Translators: http://rivendel.com/~ric/resources/dictionary.html

[26] Linguistic Data Consortium LDC, http://www.ldc.upenn.edu/

[27] Link Grammar, http://bobo.link.cs.cmu.edu/link/

[28] Longman Dictionaries Home, http://www.awl-elt.com/dictionaries/

[29] Merriam-Webster Online,
http://www.m-w.com/home.htm
[30] On line English Grammar,
http://www.edunet.com/english/grammar/
[31] Project Gutenberg Official Home Site,
http://promo.net/pg/
[32] YourDictionary.COM,
http://www.yourdictionary.com/

Manuscript Received: Feb. 26, 2001
Revision Received: Mar. 8, 2001
and Accepted: Mar. 14, 2001