전자 CORPUS를 이용한 정보통신 분야 영어학습(ESP)

한인석

도립청양대학

ishan@cheongyang.ac.kr

요 약

It is quite burden-some for non-native speakers of English to write and read more and more IT-related reports due to the rapid development of IT technology. Thus, this study aims at designing ESP materials by using huge volume of electronic ITU texts, corpora and concordancer SW. Various tests are designed to study the usage of articles, hyponym, agreement, synonym, and others. The results of this study will bring general and practical benefits to technical English writing and improving IT area students' lexical knowledge of actual English usage. The ESP materials produced by this study will also make an extensive contribution to other industries and academic areas in Korean society.

1.   서론

21세기 정보사회의 세계화 추세 속에서 정보통신(IT) 분야의 학생들 대부분은 물론 산.학.연 전문가들이 정보통신 분야의 영어보고서들을 독해하고 영작해야할 필요가 급증하고 있다. 이러한 현실에서 정보통신 분야 학생들이 우선 각자의 전공 학습을 하는 과정에서부터 영어구사의 큰 어려움을 겪고 있다. 산.학.연에서 영어 보고서와 기고서 등을 작성하고 독해해야할 필요가 더욱 급증한 것은 세계의 정보통신기술의 급진적 발전 속도에 속히 부응해야 하기 때문이다. 따라서 영문기고서와 기술논문의 발표가 개인은 물론 소속 기관 심지어 국가의 업적으로 평가되는 현실이다.

원어민은 물론 특히 외국인에게 정보통신 전문영작문에서 가장 어려운 문제가 되는 것은 전문 기술용어의 연어(Collocation), 즉 실제 상황맥락에서의 적합한 용어 구사 능력과 지식이다. 이러한 중요한 실제 사용 연어지식을 기존의 학습 교재, 사전, 여타 어느 학습지원 도구들도 충분히 제공해주지 못하고 있다. 이러한 문제의 타결을 위해서 세계적으로 공인되어 실제 사용하고 있는 영문 전자텍스트 또는 corpus를 제공해주

는 ITU의 영문보고서, CRATER corpus, BOE(Bank of English) corpus, BNC(British National Corpus)에서의 실제 사용 연어 예들을 기본자료로써 추출하고자 한다. 이 자료들을 활용하여 영어학습자들이 필수적인 전문 용어를 상황맥락에 적합하게 사용할 언어능력과 지식을 배양할 자기 발견학습 프로그램을 고안할 수 있을 것으로 예측한다.

2.　본론

2.1.　활용 corpus와 SW

Tagging이 되지 않은 전자텍스트는 ITU 사무국의 홈페이지에서 몇백만 단어 수준의 일부 텍스트라면 무상으로도 구할 수 있다. 그러나 기고서나 발간물 전체 텍스트는 구입 절차를 밟아야 가능하다. ITU-T, ITU-R 기고서 등은 모두 Full-text로 on-line 상에서 구입할 수 있으며, 분기마다 1회 CD-ROM 형태로 구입을 할 수도 있다. 기타 다양한 정보통신 관련 발간물들이 역시 전자텍스트로 구입가능하다. 본고에서는 주로 Lancaster University의 UCREL(University Center for Computer Research on English Language)에서 주관하여 개발한 CRATER corpus를 주로 다루고자 한다. 이 corpus는 ITU report text를 1백만 단어 수준에서 품사 tagging한 것이다. 아울러 UCREL에서 주도적으로 개발에 참여했던 BNC(British National Corpus) 중에서 언어학부문의 3백만 단어 수준의 corpus도 활용하고자 한다. BNC는 1990년대 초에 구축을 시작한 것이며, 1억 단어 수준의 최신 영어의 문어, 구어 텍스트를 포함하고 있다. 90% 이상이 문어 텍스트이며, 인터넷 홈페이지에서 구입이 가능한 상태이다. 정보통신용어는 주로 한국전자통신연구원(ETRI)와 한국정보통신 기술협회(TTA)의 전자통신용어사전과 정보통신 표준 용어집에서 참조하였다. 한편 Collins Cobuild사에서 제작한 2억 단어 수준의 텍스트를 포함하고 있는 BOE(Bank of English) corpus에서 만든 Cobuild English Collocation CD ROM 사전을 참고로 사용하였다. 2억 단어 수준의 텍스트에서 14만여 영어연어의 260만 실제 예문들을 추출할 수 있는 최근의 가장 방대한 전자사전을 이용하여, 쉽고도 신속하게 각종 통계자료를 구할 수 있었다.

이러한 corpus나 전자텍스트를 활용하여 학습자들이 스스로 발견학습을 하고, 특수 목적에 부합하는 영어 학습을 실제 자료 중심으로 진행할 수 있게 하기 위해서는 WordSmith같은 Concordancer SW가 반드시 필요하다. Johns(1991a, b)는 자신이 주장한 DDL(Data-Driven Learning)에서 Concordancer가 널리 공인된 언어 사용례의 증거자료들을 추출해 줄 수 있으며, 이 자료들을 학습자에게 맡겨서 각자가 그 안에서 문법을 스스로 발견케 하는 것이 학습과정의 중심이 되어야 한다고 역설한다. Leech(1987)와 Sinclair(1986) 같은 대표적 Corpus linguist들이 공통적으로 강조하는 것은 현대의 첨단정보기술의 발전이 인문과학적인 언어학은 물론 언어교육학의 방법론에 큰 변화를 초래했다는 것이다. 즉 학자들의 직관과 가정으로 미리 만든 문법규칙의 틀에 끼워 맞추는 것이 아니라, 광범한 실제 사용 언어자료들을 첨단 기술로 처리하

여 증거자료로 제시하고 이를 분석하여 문법규칙을 찾아내는 방법이 부각되고 있다.

2.2. 언어 자료의 생성

CRATER corpus에서 우선 정보통신용어 'interface'의 예를 추출해본 결과, 'interface'의 앞에서 쓰이는 관사들을 잘 분석할 실제 예문들을 구할 수 있었다. 특정화된 'interface'는 정관사 the를 필요로 하고 있다. 즉 the mentioned interface, the interface defined below, the v4 interface 등. 반면에 관사를 붙이지 않은 경우는 interface A, B, interface c2, interfaces 등에서 찾을 수 있었다.

1)is for each Administration to define the sources of permanent and impulsive noise to be applied to the **interface** of the sending and receiving parts in the light of

2)rec. E . 220 Ergonomic aspects of visual display terminals . Ccitt rec. E . 221 Human **interface** to visual display terminals .

3)Ccitt rec. Q . 920 Isdn user - network **interface** data link layer - General aspects. Ccitt rec. Q . 921 Isdn user - network inter

4)Ccitt rec. Q . 921 Isdn user - network **interface** data link layer specification . 5)network **interface** data link layer specification . Ccitt rec. Q . 930 Isdn user - network **interface** layer 3 - General aspects .

6)Z . 331 - Z . 333 Specification of man - machine **interface** .

7)The Recommendation describes the subscriber side **interface** characteristics .

8)The **interface** functions defined are those necessary for interworking with digital and analogue transmission

9)These cover the detailed transmission characteristics of the various types of **interface** that may be provided on a digital exchange .

10)The location of certain types of Cme relative to the exchange determines the choice of signalling **interface** .

11)The **interface** Isc and Dcme described below is intended to connect Deutsche Bundespost exchanges to the

12)The mentioned **interface** has three basic functions : dynamic load control Isc

13)For transmission of the signalling signals , the **interface** operates with 2 bits

14)Physical level of **interface**. For transmission of the signalling signals ,

15)To avoid a special **interface** at the Isc , Dcme / Isc signalling is transmitted in the same Pcm system to the Dcme as the

16)(In principle , any other physical **interface** with 2+2 bits is suitable for the forward and backward directions .)

17)**Interface** A is a digital **interface** to allow interconnection at the first level of

18)The frame structure at **interface** A should be identical to that of the first order multiplexes described in Recommendations G.

19)**Interface** B is a digital **interface** to allow interconnection at the second level of a digital transmission hierarchy towards

20)The electrical characteristics of **interface** B are described in Recommendation G . 703
. The frame structure at **interface B s**

21)Characteristics of analogue **interfaces** towards other exchanges

22)**interfaces** are specified in Recommendation Q . 551 . The transmission characteristics
of 2 - wire **interface** c2 are described in Recommendation

23)wire **interface** c2 are described in Recommendation Q . 552 and of 4 - wire **interface**
c1 in Recommendation Q . 553 .

24)associated with **interface** v1 include two B channels and one D channel as defined at
the user network **interface** in Recommendation I . 412 .

25)The characteristics of this **interface** may not be structured according to the Isdn prin-
ciples . In this case , only the elect

26)**Interface** v4 is a digital **interface** used to connect a digital access link which includes
a static multiplexer supporting several

27)The information on the status of the Crc processing shall be passed across the v4 in-
terface . 3 . 5 . 3

한편 유사어들도 corpus에서 추출한 실제 사용 예문을 통하여 비교, 분석할 수 있
다. 우선 두 유사어 large, great의 사용례들을 검색해본다. 그리고 그 중에서 특히 빈번
히 나타난 연어의 예로 large majority, great majority, large number, great number의
사용례를 추출해보았다. Large의 예들을 먼저 살펴 본다.

1)Selecting a rather **large** antenna may allow the use of both a low cost Lna

2)advantage in selecting a **large** antenna diameter (as explained above) may be

3)the usual choice for **large** and medium earth stations ; - simple , temperature -

4)Hpa combiner(s), are usually needed in **large** stations ; - Cascaded filter 5)combiners
may also be considered for very **large** stations.

6)In practice the **large** majority of international telephone traffic

7)transfer messages of different types having a **large** varietry of formats ; messages of
different types having a **large** variety of formats ;

8)for a **large** number of applications black and white reproduction is sufficient

9)owing to the **large** number of countries requiring two - digit codes ,

10)Hence in the case of **large** groups (for example , more than 40 circuits in

11)in circumstances necessitating the use of **large** groups , there are great

12)in some countries covering **large** geographic areas it may be appropriate to 13)extend
protocols should offer a **large** degree of compatibility with protocols

14)an increase of the economical transmission of **large** numbers of anisochronous

15)the transmission of **large** numbers of channels ; (c)that certain telex switches 16)ap-
plicable both to small and to **large** national networks ;

17)it should be possible to fix limit values for **large** countries and they should

18)difficult to lay down standards applicable both to small and to **large** national 19)small
and to **large** national networks. However , it has been possible to fix

20)need to exchange **large** amounts of data. It is considered that for these

21)send a large number of messages in a short amount of time , no flow control

22)Given the large potential number of destinations, measurements may have to be

23)based on a large sample of local lines in the British Telecom network have shown 24)the large economies in personnel that can result from the introduction of

25)In the case of signalling systems with large signalling capacity , it is 26)possible total system include a large diversity of functions and that further

27)If the rather large amount of effort needed is available and the importance of

28)case of a relatively large retransmit buffer, both the transmit buffer

29)is considered as of large size. (The value of n is for

30)the economic transmission of large numbers of anisochronous telegraph

이어서 great의 예를 다음과 같이 일부 추출해 보았다.

1)Agreeable circuit 3 Fair circuit. A moderate , but not too great effort is necessary. Not a very agreeable circuit 2 Poor circuit .

2)the Ccitt has reached the conclusion that four circuits is a representative figure to assume for the great majority of international connections.

3)Parameters related to operation of national networks may be affected by great increase in satellite propagation time , compared with the terrestrial value ,

4)Measurements under real conditions are complicated, time - consuming and not repro-ducible with great precision, especially when carbon microphones are involved .

5)mode (1) is concerned with the attenuation of signals subject to tropospheric

6)ropagation via near great circle paths .

7)The great promise of Isdn is the cost savings (equipment and more efficient use of the resources) and the

8)The light route service is of great importance for service to small communities .

9)this would no doubt prove to be a great advantage when this possibility is generally utilize, unanimously declares the following view

10)noise level as well as on other external conditions . There is , furthermore , a great variance individual listeners .

11)Since the satellite is located at a great distance from the Earth

12)The great majority of satellite telecommunication systems use Pama

13)- great flexibility for changes of services and traffic plans and for all changes in the earth segment.

14)been completely revised during the last few years due, in great part, to the introduction of satellite systems .

15)However, since in the case of satellite transmission, power efficiency is of great impor-tance , lower values are sometimes chosen

16)In order to insert successfully the satellite into the transfer orbit, the launcher must operate with great precision with regard to the magnitude and the orientation of the velocity vector .

17)- great sturdiness and long working life, - possibility of operating

18)The earth - station owner / user has **great** flexibility and freedom in deciding upon the best transmission method for his requirements

19)The Ccitt , considering that it will be a **great** advantage to use a standardized test chart to check the quality of facsimile transmissions and

20)Owing to the development of international document facsimile transmission services , a **great** variety of characters and symbols ,

21)that it should be possible to fix limit values for large countries and they should apply to the **great** majority of user locations taking part in the international service , unanimously declares

22)Message Transfer Part signalling performance Signalling performance is defined by a **great** number of different parameters .

23)A **great** many of such qualifications are defined in standard signalling system's

특히 빈번히 나타난 예로서, great majority와 large majority의 예문들을 아래와 같이 정리해 본다. great majority가 large majority보다 빈번한 사용을 보여준다.

1)for the **great majority** of international connections .

2)The **great majority** of satellite telecommunication systems use Pama, but in the case of sporadic traffic varying

3)in practice, the **large majority** of international telephone traffic is routed

4)the **large majority** of international telephone traffic ex " international 5)telephone they should apply to the great majority of user locations taking part

6)they could apply to the **great majority** of telex subscriber stations or gentex stations taking part in the

7)that the **great majority** of international telegraph circuits are routed on voice - frequency telegraph (Vft)

8)**overwhelming majority** of block codes in use are selected .

역시 빈번히 나타난 예로서, great number와 large number의 예문들을 아래와 같이 정리해 보니, great number보다 large number가 훨씬 더 빈번히 사용되고 있음을 알 수 있다.

1)supplementary services with a **large number** of variants without a proliferation of new messages ;

2)designed for a **large number** of channels the mean power handling capacity per telephone channel corresponds

3)System r2 enables transmission of information concerning a **large number** of different conditions of called subscriber' s line or reasons why call set - up

4)which have a **large number** of satellite circuits , e.g. when national satellites

5)simultaneous outage of a **large number** of terminations .

6)requirements for c22 interfaces of a **large number** of administrations :

7)section encompasses the requirements of a **large number** of administrations .

8)receiving sensitivity determined using a **large number** of real ears .

9)encompasses the requirements of a **large number** of administrations .

10)distance connections encompasses the requirements of a **large number** of administrations .

11)requirements of a **large number** of administrations :

12)For a carrier with a **large number** of channels , good linearity is also necessary

13)to produce a **large number** of dibit transitions to facilitate achieving

14)business communications to a very **large number** of remote locations .

15)a medium or **large number** of nodes (e.g. more than five) for possible use

16)a common resource to a **large number** of terrestrial users thus allowing

17)there is potential for a very **large number** of signals ;

18)the troposphere depends on a **large number** of parameters

19)China successfully conducted a **large number** of experiments on satellite .

20)a **large number** of errors at the decoder output will be produced .

21)due to the **large number** of possible vectors, some kind of step search algorithm

22)include a **large number** of earth stations (several dozen and several thousand)

23)supply a **large number** of transponders (e.g. , 50 transponders with a power

24)have established a **large number** of terrestrial / satellite paths as a means of

25)with methods by which a **large number** of earth stations can interconnect their

26)allow for a **large number** of interacting parameters : signal - to - noise ratio

27)For medium and **large numbers** of channels in each station , Fdm - Fm or Tdm -

28)Aloha systems in which a **large number** of users share the satellite transponder

29)for a **large number** of applications black and white reproduction is sufficient ;

30)owing to the **large number** of countries requiring two - digit codes ,

31)economical for the transmission of **large numbers** of channels ;

32)send a **large number** of messages in a short amount of time ,

33)common to a **large number** of the failures , then they are suspect as a

34)services with a **large number** of variants without a proliferation of new messages

35)a much **greater number** of earth stations can be connected in the Dama mode than

36)The greater **number** of code combinations available in ia5 means that not every

37)thus permits a **greater number** of simultaneous calls than would otherwise

38)where a great **number** of channels will be interconnected at random

39)the advantage of having a **greater number** of telegraph channels on the same

40)performance is defined by a **great number** of different parameters .

Johns (1991a)가 change의 앞에 유사어 big, great, huge, large가 쓰이는 빈도를 조사하여 비교해본 결과는 다음과 같다.

Times에서의 Frequency per 1M words

Big - 1.4

Great - 0.8

Huge -0.6

Large - 0.1

1) l of this will have certainly made a **big change** to Ma's thinking in the Bach

2) ence said the union was undergoing a **big change** of character and would, in

3) our way? The past decade has seen a **big change** in the way we teach mathematics,

4) Asprey was pragmatic. He knew that a **big** management **change** and a serious

5) 46;t listening and there is to be no **big change** . This could rebound on Major: MPs

6) , a fantasy world. Unless there is a **great change** of direction in the next few

7) ; she says, "when you undergo a **great change** and it was a colossal **change**

8) of the tables this year, which is no **great change** from years past. But what

9) strel Zhang Kou, was celebrating the **great** policy **change** from the Third Plenum

10) of the place. Ah, but then came the **great change** ; I don't know who was (and

11) Moving out of Corsham Court meant a **huge change** for the Peaches. George would

12) t after 30 years, and in the face of **huge** social **change** , the regulations are

13) ranello has emerged from a period of **huge change** . Boom when Ferrari sold a

14) omacy do not impinge greatly on that **huge change** in the world, which did more

15) egrity of the current system without **huge** organisational **change** Everyone

16) s that respond to deformation with a **large change** in conductivity. At

17) ing industry. The first **large change** to telecoms law in 60 years,

18) for under Mr Palk there has been **large-scale** organisational **change** including

2.3. 연어를 이용한 자기발견학습 과정

우리 나라 학생들이 가장 혼란스럽게 느끼는 관사 사용의 실례를 ITU corpus에서 구하여, 다음과 같이 문제를 만들어 학생들에게 연습을 시킬 수 있다. 기술 용어 중 하나인 'exchange'의 앞에서 관사 사용이 어떻게 되고 있는지를 학습자들이 직접 사용례에서 파악하게 한다.

[문제 1] 다음 글에서 _____에 적합한 the, a, an 또는 "관사필요 없음"을 기재하시오.

1)in an abstract way, _____ logical **exchange** of information and control the

2)need to be made to interconnect _____ **exchange** at the renter ' s premises with
on

*3)using one or more _____ private automatic **exchange** (s) as a transit switch

4)via _____ dedicated access lines and inter-**exchange** circuits and shared com-
mon

5)- on - idle , are used on _____ inter-**exchange** circuits , directly-homed s

*6)Signalling on _____ inter-**exchange** circuit is in accordance with

7)is controlled by _____ terminating **exchange** or Pabx in a conventional manne

*8)signalling on _____ inter-**exchange** trunks D . 4 . 1 Superv

*9)strument or its equivalent and _____ public **exchange** or Pabx in different countrie

*10)signalling unit (Isu) and _____ **exchange** signalling unit (Esu) .

*11)but in different systems which must _____ **exchange** information to achieve a common

12)in an abstract way, _____ logical **exchange** of information and control the

*13)information transfer can start, _____ **exchange** of a Sabme frame and an Unnumbe

14)in an abstract way, _____ logical **exchange** of information and control th (CRATER corpus에서 자료 추출함)

영어 전치사의 용법도 corpus에서 예문들을 추출하여 세밀하게 파악할 수 있다. 다음 문제를 통해서 전치사구의 내용을 기반으로 해서 밑줄 친 부분에 빠진 한 명사 단어를 추측해내는 연습을 학생들에게 시킬 수 있다.

[문제 2] 다음 글의 about 앞에 적합한 명사는 무엇인지 keys에서 골라 적으시오.

1) For example , _____ **about** network layer protocol in a message sent by a

2) This information identifies the protocol options used and gives other _____ **about** the message and its routing .

3) The subscriber - busy signal will also be sent in case of complete _____ **about** the place where the busy or congestion conditions are encountered and in the

4) calling-party's - category indicator Information sent in the forward _____ **about** the category of the calling party and

5) the category of the calling party and, in case of _____ **about** the service language to be spoken by the incoming , delay and assistance operators

6) Further _____ **about** the Hft performance may then be estimated by repeating the sending measurements

7) Note - Some _____ **about** reference equivalents can be found in the Ccitt Red Book

8) when the Administrations concerned agree that it is necessary to signal to _____ **about** a delay in transmission over the radio telex channel, delay signals shall be used having the

9) service message transfer unit(Smxu). Used to convey service _____ **about** messages .

keys: informations, data, ncertainty, direction, semi-automatic calls, information, telex subscribers, information

단어와 단어 사이의 의미관계를 나타내는 하의어(hyponym)과 상의어(Superordinate)의 대표적 예들은 such as에서 분명하게 나타나고 있다. 실제로 ITU corpus에서의 용례들을 어떠한지 알아보기 위해서, 다음과 같이 추출하여 문제로 만들어 보았다.

[문제 3] such as의 앞에서 가장 적합하게 쓰일 명사를 keys에서 찾으시오.

1)There should be efficient means of establishing, testing and changing call processing _____, such as translation and routing information .

2)which occur in telephone calls come from a number of _____, such as faulty construction of the switching equipment , defective earthing at exchanges

3)provided that there is a condition which must require interruption of the call setting - up process , such as time - out or congestion at any point of the national or international network and non -

4)Disc constructed of hard , stable and on - megnetic material, such as brass , having a diameter of 63 mm and 5 mm thick .

5)action to give appropriate trouble indications or alarms and perform other appropriate _____, such as removing circuits from service.

6)an attempt must be made to reduce these levels by various _____, such as the coordination of frequencies , which consists in separating the carriers so

7)they will take full advantage of their own _____, such as multiple access , broadcast capabilities , easy and rapid installation of new services

8)but other types of _____, such as analogue links with modem, can also be

9)The extent of the user part functions may differ significantly different categories of _____, such as : - users for which most user communication functions are defined within the

10)However, a burst error correcting code, such as a Reed - Solomon code , may be also used .

11)risk criteria have also been proposed for longer duration acoustic _____, such as tones .

12)representations for additional control functions , in particular for control functions with _____, such as presentation control functions

13)signalling information of the message includes the actual user _____, such as more telephone or data call control signals , management and maintenance information

14)- regional leased circuits may be provided with various forms of bandwidth economizing _____, such as speech interpolation systems and digital voice compression

keys: information, sources, action, means, features, data links, users, disturbances, parameters, information, systems

다음은 학습자가 복잡한 주어부에서 특히 복수명사가 끝에 나타나는 주어부에서 단수명사 주어를 식별해내는 능력을 키우는 문제들이다. 주어와 동사의 일치라는 규칙을 습득하게 하는 좋은 연습이 된다.

[문제 4] 다음 문장들에서 be동사 is의 주어가 되는 것을 찾아 밑줄을 치시오.

1) December could not agree. So an eight man <u>team</u> of scientists is to make a

2) ists of all-carbon rings. In a few rings one of the carbons is replaced by an

3) Britain's only major industry in digging up metals is based in Cornwall.

4) tive, not monopolistic. Each of the five serious dailies is separately owned and

5) due to reduce noise, The choice of species in tree plantings is also important.

6) The only party aggrieved by these arrangements is the Ministrv of

7) An intriguing project somewhere between the two options is the commissioning

8) that the Lransfer of genes between viruses and cells is a very common event.

9)ys that the attitude of doctors towards patients with colds is also significant.

10) The total number of people who will benefit from such drugs is unknown although

11) Base material which does not meet the normal specifications is very difficult to

12) cost of pushing a European patent through in nine countries is around ?0,000.

13) ome so specialised that communication across the disciplines is difficult to

14) 1 industrial pollution. A strategy for handling toxic wastes is beginning to

15) lf' chemical reagent, whose ability to produce free radicals is well known to

16) rials in Essen have proved that a trolley bus with batteries is too expensive to

17) rprise in Milan: that the amount of energy an animal creates is independent of

18) ing schedules), the only place we can make cuts or additions is at the end of

19) the market share for Europe's only manufacturer of recorders is still around 5

20) A most important consideration in laser techniques is the sensitivity of detecti

21) "I have always believed that telling stories is part of any social life.

22) ularly. A programme of Lraffic counts and deflection surveys is needed if

23) isotope selection effect observed by the San Piego chemists is not the only one

24) . The line with the steep slope that appears on both graphs is the best fit line

25) ofessional men. The report points out that quantifying risks is not without its

26)material which flows out under pressure from the capillaries is returned from the

27) rinsic character of much of the activity in these two fields is quite different.

28) job centre vacancies. He wants work. "You see, none of us is used to

29) the development of multi-million dollar industrial projects is neither feasible

30) nsiderably more successful. The biggest threat to the Tories is a confident,

[문제 5] 다음 문제의 예문들의 주어의 첫 부분이 임의 삭제되어 있다. 학생들 스스로 나머지 본문의 내용을 근거로 하여 삭제된 부분을 복원해 본다.

1) _isk_ of accidents causing large numbers of deaths or injuries is low.

2) _valuation_ of potential new varieties held in such gene banks is a long-term affair; for ex

3) _alysis_ of the human problems behind double blind drug trials is to be found not

4) _nitoring_ of airborne contamination and the health of workers is essential.

5) _tput_ per annum of scientific papers over the past four years is revealing.

6) _anism_ of both forms of integration in animal and plant cells is still poorly understood.

7) _ory_ of belief in the supernormal over the past hundred years is studded with the names of

8) _nt_ of tissue that must be destroyed to produce these effects is very small - destruction

[문제 6] 아래의 예문들에서 각기 주어부의 첫 부분이 임의 삭제되어 있다. 예문의 내용에 의거하여 삭제된 부분을 추측하여 복원시켜 본다. 첫 문제의 삭제된 부분은 "One of the most common ...."이다.

1) _he most common opportunistic infections associated with AIDS is a type of pneumonia, calle

2) ____between the expansion rates of metal and ceramic components is a big obstacle.

3) _purer programs to aid the appraisal of urban traffic demands is described by TPA

4) ____ssing samples drilled from beneath the sea for oil companies is a commercial winner, and a

5) ____uld receive treatment from currently recognised orphan drugs is in the region of

6) ____proving Britain's crumbling system of sewers and water mains is gargantuan.

7) ____merges from these recollections of one of his prize patients is evidently quite

8) ____ular "science-writing" should be stripped of technical terms is to misconstrue the nature(Tim Johns, 1991b:27-35)

3.   결론

정보통신 분야에서 가장 중요하게 우리가 다루어야할 방대한 양의 ITU 영문기고서를 최대한 효율적으로 분석, 처리하기 위해서는 corpus 또는 최소한 전자 텍스트 형태로 해당 자료를 입수하여야 함을 확인하였다. 현재 다양한 방법으로 ITU report text를 구할 수 있음을 알아보았다. 이러한 수천만 단어가 넘게 포함된 거대한 전자텍스트나 corpus를 효율적으로 검색하기 위해서 여러 SW가 있지만, 그 중 가장 용이하게 고속으로 검색을 해주는 WordSmith를 직접 이용해보고 활용가치가 높음을 알았다. 실제로 정보통신 용어 몇 개를 중심으로 collocation 예문들을 검색, 추출해보았다. ITU의 정보통신 보고서에서 현재 사용되고 있는 collocation의 사례들을 살펴보면서, 단어의 사전적 지식의 한계를 벗어나 실제 상황맥락에 관련한 collocation 지식을 구할 수 있었다.

이러한 확인 자료를 바탕으로 학생들이 가장 혼동을 많이 하고 어려움을 느끼는 관사, 용어 용례, 전치사, 유사어, 하의어 판별 능력 배양을 위한 문제를 만들어 제시하였다. 향후 보다 구체적으로 연어의 지식을 체계화하기 위하여, 주요 용어와 이의 수식어, 주요 동사와 부사(구) 등을 빈도순으로 조사, 분석할 필요가 있다. 본고의 이러한 시도는 이후에 다양한 여타 분야에서 특수목적 영어학습 프로그램 개발을 추진하는 과정에서도 바로 응용될 수 있다고 사료된다. 적절한 전자 Corpus들을 발굴하고 각종 컴퓨터 지원도구들을 활용하여 최적 효율의 언어학습지원 및 처리시스템을 고안할 수 있을 것으로 기대하며, 이러한 응용기법이 실용적으로 전반적인 학문과 산업분야에 확대될 경우 그 효과는 우리 사회 곳곳에서 광범위하게 나타날 수 있을 것으로 기대된다.

참고문헌

한인석. 1999. "A study on the electronic English collocation dictionary in IT sector and its applications to TEFL", 언어연구 15권 1호. 321-330.

한인석. 2000. "Corpus-based online collocation resources for EFL learners", 청양대학 논문집 제 1집. 47-60.

한국정보통신기술협회. 1997. 정보통신 표준 용어집.

한국전자통신연구원. 1990. 전기통신용어사전. 대전:동화출판.

Johns, Tim. 1991a. "Should you be persuaded; Two examples of data driven learning", In Johns, Tim, and Philip King Eds. Classroom concordancing, ELR Journal 4. Univ. of Birmingham. 1-16. Univ. of Birmingham. 1-16.

Johns, Tim. 1991b. "From printout to handout: Grammar and vocabulary teaching in the context of data-driven learning", In Johns, Tim, and Philip King Eds. Classroom concordancing, ELR Journal 4. Univ. of Birmingham. 1-16.

Leech, Geoffrey. 1987. "General introduction", In Garside, Roger and Leech, Geoffrey. Eds. The computational analysis of English. Harlow, Essex: Longman. 1-16.

Sinclair, john. 1986. "Basic computer processing of long texts", In Leech, Geoffrey and Christopher N. Candlin. Eds. Computers in English Language Teachng and Research. Harlow, Essex: Longman. 185-203.

Stevens, Vance. 1993. "Concordances as enhancements to language competence", TESIL Matters 2. 6-11.

Stevens, Vance. 1990. "Text manipulation: What's wrong with it, anyway?" CAELL Journal 1. 5-8.

Tribble, Chris. 1989. "The use of text structuring vocabulary in native and non-native speaker writing", MUESLI News (June). 11-13.