ABSTRACT
                Could corpus analysis be used to solve some of the problems
surrounding non-native speakers' understanding of seemingly synonymous words?
This research is needed because there is no single universal unified lexical
theory that can be applied to the study of semantics or of language as a
whole. Instead, there are a variety of views than can be condensed into two
competing positions: (1) language is a system, formal in nature, that does
not need to take into account the real world (Chomsky); or (2) the sharing
perspective of language, one that does take into consideration the link
between shared signs and shared concepts present in daily life. Another way
to put it: Do we want our students to be able to communicate with others
(sharing), or simply to manipulate words (system). Each word has a meaning
(system), but also a sense and way that it is used that is important for
communication (sharing). Furthermore, words are rarely if ever wholly
synonymous; the most that can be claimed is that they have varying degrees of
congruence. Six third-year Japanese high school students made up the study
group for this paper. Six pairs of target synonymic pairs (words) in English
were chosen for the experiment. The investigation was framed into five
distinct stages. In each stage the students investigated the meaning/sense of
the word using different methods, beginning with the dictionary and moving on
to more complicated exercises, including the use of concordances. It is
concluded that this series of exercises succeeds in allowing students to
discover and learn the different nuances of words needed for successful
second language learning. Five appendices with the exercises used in this
study are included. (Contains 28 references.) (KFT)

An investigation into how corpus analysis may be used in the second language classroom to solve some of the problems surrounding non-native speakers' understanding of seemingly synonymous words.

Sean Romano Maddalena
October 2001

1. Lexical Theory

1.1 Vocabulary: meaning and context

At present, there is no single unified lexical theory that can be systematically applied to a study of semantics or to the language as a whole. Rather, there exist a number of different perspectives that vary in their usefulness in relation to the purpose for which they are employed. These disparate views can be condensed into two competing propositions, namely that language may be seen as a *system*, formal in nature, that does not need to take into account the 'real' world, or an alternative *sharing* perspective of language; one that does take into account the link between shared signs and shared concepts present in our daily lives. In support of the system approach, Chomsky (1965), says, "Language learning requires use of language…grammar can be regarded as a theory of language," but Humboldt, (in Chomsky), argues, "We cannot really teach language, only present the conditions under which it will develop", which suggests a good deal more than language described solely in terms of its grammar. In the pedagogic domain, and in somewhat simpler terms, this issue may be defined in the following way: Do we want our students to be able to communicate with others, (sharing), or simply to manipulate words, (system)?

We may use many words to refer to the same thing, but feel very differently about each of them and their connotations for us as individuals. To put it another way, it is not the *meaning* of a word but its *sense* or the *way* that it is used in any given situation that is important for successful communication. For example, if we try to teach the word '*Christmas*' to our students, they may only ascertain that the word denotes the holiday on December 25th but not that to others in the relevant *speech community* the word connotes a good deal more: presents, turkey, mistletoe and chestnuts roasting on an open fire. In light of the two perspectives outlined above, there are two very different possible answers to the question of when can we be said to have *taught* a word or conversely, when our students can be said to have *learnt* a word. It is this notion of "meaning", whether connotative or denotative, that the structuralist, or systems, school does not follow. On the other hand, the social constructionist, or sharer, follows a much harder road, believing that there is no other way to explain to our learners the full richness of our language and to empower them to express to others their own meanings.

1

Context is a key component of the *language as sharing* perspective. As Benson says (1972), "Without context there is no communication". Regarding synonymic words, context plays an essential role in resolving ambiguity in meaning. Carroll (1964), says: "Students must be taught the meanings of unfamiliar words and idioms; they must be helped in recognizing unfamiliar ways in which familiar words may be used; and they must be aware of the ambiguity in meaning and the role of context in resolving it. Often the task that presents itself to the teacher is not merely to explain a new word in familiar terms, but to shape an entirely new concept in the mind of the student". By way of contrast, *co-text* is more related to a systems perspective.

In my description of the study that follows, the role of both context and co-text are seen as essential in molding students' sense of understanding new vocabulary.

2. Corpus, Concordance and Collocation

2.1 History and development of corpus linguistics

As stated above, studies of language can be divided into two main areas: studies of structure and studies of use. Corpus analysis (CA), focuses on the second of these, studying actual language used in naturally occurring texts. Ever since Firth (1957), stated that "You shall know a word by the company it keeps", it has been a practice in linguistics to classify words not only on the basis of their meanings, but on the basis of their co-occurrence with other words. However, in a purely practical sense it is only in recent times that machines have given us the ability to identify these relationships in a meaningful and significant way.

From the simple listing of words in the Middle Ages by hand, through to the first modern, electronically readable corpus, the Brown Corpus of Standard American English, and its close cousins the Lancaster-Oslo-Bergen corpus and the Kolhapur Corpus, the computer-aided analysis of vast amounts of authentic data has come a long way in a very short time. Almost half a century ago Firth (1957; 31), made the following prophetic statement: "The use of machines in linguistic analysis is now established". John Sinclair (1991: 1), describes the evolution through the last three decades in the following way: " Thirty years ago when this research started it was considered impossible to process texts of several million words in length. Twenty years ago it was considered marginally possible but lunatic. Ten years ago it was considered quite possible but still lunatic. Today it is very popular". This popularity has led to an increased understanding of the relationship of meaning to form as formal patterns, previously undetected, have come to light. Sinclair again, "At the very least, the quality of linguistic evidence is going to be improved out of all recognition...it is my belief that a new understanding of the nature and structure of language will shortly be available as a result of the examination by computer of large collections of texts", (1991b; 489). Stubbs (1996), concurs, "computer-assisted analysis of texts and corpora can provide new understanding of form-meaning relations".

2

2.2 Teachers and students as researchers

While much of this work has been undertaken by linguists in pursuit of lexical theory, there is reason to assume that language teachers and students alike may themselves make practical use of CA as a tool to improve their own understanding of language and so aid vocabulary acquisition. Technological advances in the power of modern computers, coupled with their now widespread availability, has meant that it has become increasingly easy for large corpora to be accessed by researcher, teacher and student alike.

In support of this notion, Barlow (1995; 1) says, "the use of text analysis tools and corpora allows everyone to become researchers: from the theoretical linguist to the student learning a second language". Johns (1991a; 2), echoes this sentiment: "The perception that 'research is too serious to be left to the researchers': that the language learner is also, essentially, a research worker whose learning needs to be driven by access to linguistic data". But Gavioli (1997), warns that "simply giving students direct access to the data produced by the computer is not enough to make them research workers", rather they need be guided in a series of methodological steps. In the case of the current investigation this was something that evolved as the investigation progressed.

2.2 Corpus analysis in the classroom

Much has been said about the use of 'real' rather than 'contrived' data in the classroom. It is generally accepted that it is extremely difficult to invent examples which simulate naturally occurring language Sinclair (1991; 6), warns, "we should never offer as an instance of language in use some combination of words which we cannot attest in usage". Once again we return to the *system* vs *shared* perspective. For example, if we expose to our students the following textbook citation, "It is a hot day today, isn't it?" we might say that an examination of the grammatical construction allows for a successful gleaning of meaning but, in truth, how often do any of us actually use this sentence in our normal, everyday lives? When faced with the authentic, "Jeez, how about this heat?" the systems approach falls woefully short; it is simply not capable of decoding the message carried in the words. Traditional theoretical grammars are simply not able to categorise the unexpected and diverse constructions that authentic material constantly exposes. Here then is a clear opening for the inclusion of genuine corpora in the teaching environment. Alderson (1996), says that an important contribution to the use of corpora in language learning will be "to make it possible for learners not only to have access to real rather than contrived language data, but also to explore language data on their own and to generate hypotheses and rules about the language". Corpus analysis is seen therefore as a starting point for further study rather than a means to an end.

3

3. A Classroom Question

3.1 Problem words

The focus for this paper was actually presented to me by one of my own students when, out of the blue, I was asked to explain to the difference between "ruin" and "destroy". At the time, my student had been struggling with her studies in preparation for an up-and-coming English proficiency test, and had come across the words in the course of her reading. Though confident in my own ability to differentiate between the two, I found it extremely difficult to elucidate the subtle differences of usage in a convincing way and left her, I am sure, almost none the wiser. That I was unable to provide a satisfactory answer to her inquiry was somewhat disconcerting to say the least and it set my mind on the track of addressing the wider issue of problematic synonyms in general.

3.2 The situation

As part of my teaching responsibilities I am involved in helping Japanese high school students in their attempt to pass the *Eiken* examination. These nationally administered tests, first administered in 1963 by the Society for Testing English Proficiency (STEP), are highly respected in social, educational and employment circles and taken by millions each year. The Ministry of Education endorses the *Eiken* tests, and recommends students take them. Consequently, its syllabus forms a significant part of most, if not all, Japanese junior and senior high schools' English programs. The tests are ranked from the highest, *I-kyu* (first level) to lowest, *go-kyu* (fifth level). At my school, I am involved in teaching a small group of six third-year high school students who are preparing to take the *ni-kyu* (second level), of this test.

In the Japanese textbook that we use, 'Eiken Ni-kyu Kyohon', the very first page lists a section of problem synonymic pairs, grouped under grammatical function. The inclusion of these words, rather than others, is based on the simple fact that they are the most commonly occurring problem pairs used in previous tests and are therefore considered worthy of special attention. However, the book merely presents each pair, along with the relevant Japanese translation, together with a short caveat that the pairs are not wholly congruous. No attempt is made to explain the sense in which one word may be different from another, or the situations in which one, rather than the other, is used. This, therefore, is my starting point: how can I use CA to help my students to better understand the usage of problem synonymic pairs and, in so doing, aid them in passing the *Eiken* examination?

3.3 What is a *synonym*?

The crux of the problem that faces non-native speakers, (NNS), is perhaps best illustrated when one examines the following two dictionary extracts: The Cobuild English Language Dictionary (1987), defines a synonym as follows: "A word or expression which means the same as another word or expression". The Longman Dictionary of Contemporary

4

English (1978), however, defines a synonym in a slightly different way: "A word with the same or nearly the same meaning as another word in the same language". With a cursory glance, the difference between the two may appear slight, but in actual fact there is a vast chasm lying between "the same" and "*nearly* the same"; a void that the language student needs help and guidance to traverse successfully. If synonymous words really did carry the exact same meaning then things would be very much easier for student and teacher alike. But, in truth, words are rarely wholly synonymous, and the most that can be claimed is for varying degrees of congruence.

In support of CA as a research tool in this area, Barlow (1995; 21) says, "A concordance search is an excellent tool for investigating a word or morpheme that is polysemous or has multiple functions. Once students have grasped the meaning of a word, they can then use a corpus to discover the other meanings or functions of the word". This study focuses on the first of Barlow's two-stage approach; "grasping" the meaning and identifying patterns of usage that may help to differentiate one synonymic word from another.

4. Methodology

4.1 The study

The research question for the study was as follows: Could corpus analysis be used to solve some of the problems surrounding non-native speakers' understanding of seemingly synonymous words?

4.2 The group

There were six third-year high school students making up the study group. At my school students are streamed in their English classes based on ability, from A to D, and the pupils involved in this inquiry were all representatives of the highest grade. With the help of a Japanese English teacher, the investigation was carried out over a period of four weeks and included both classroom and homework.

4.2 Target synonymic pairs

Six pairs were chosen for the project by the students from the 'problem' list as presented in the prescribed *Eiken* textbook, including the pair that had prompted the original inquiry.

1. Fertile : Rich (adjective)
2. Ripe : Mature (adjective)
3. Sick : Ill (adjective)
4. Fury : Anger (noun)
5. Contrary : Opposite (adjective)
6. Ruin : Destroy (verb)

5

4.3 The need for a relevant corpus of significant size

As Sinclair (1991;13) cautions, "the beginning of any corpus study is the creation of the corpus itself...the results are only as good as the corpus". However, despite extensive research I was unable to locate an *Eiken*-specific corpus for my investigation. Electronic media produced to accompany textbooks exists solely in the CD-audio format to be used in conjunction with the listening section of the test. Whilst it is possible to compile ones own corpora by hand, the need to produce a corpus of a significant size makes this approach impractical. Even corpora several million words in size may not produce significant evidence: "To study the meaning and use of words, we need a very large corpus", caution Biber, Conrad & Rippen (1998; 30). Clear (1993; 274 in Granger), agrees, "reliable evidence of patterning...can be obtained only from very substantial text corpora". Clear again, "By far the majority of lexical items have a relative frequency in current English of less than 20 per million".

Aside from sheer size, the role of register must also be taken into consideration as the inclusion of only moderately common or rare lexical words can depend a great deal on the different genres represented in the texts of the corpus. Consequently, a corpora formed across a wide range of registers may yield concordances that reflect such diversity.

Faced with the problems of size and register, I decided to make use of one of the well-establish corpora, and I finally settled on the Lancaster-Oslo-Bergen (LOB), corpus, a written text corpus of approximately 1,000,000 words of British English taken from 15 different registers. As this was my first attempt at corpus analysis I was concerned with two things: firstly that the corpus was relevant for my students in their current studies i.e. the concordances that the analysis would provide should contain the kind of English that my students were likely to encounter in their exam, and secondly that the corpus was manageable for me, in a purely practical sense, as an apprentice researcher in this field. More detailed studies usually use larger databases and others seek to compare across corpora. There are even 'mega' corpora available for analysis, such as the British National Corpus (BNC), which contains some 100 million words and the even larger Bank of English (BoE), which contains something in the order of 320 million words. However, with no domain-specific corpora to call upon, the general-purpose nature and manageable size of the LOB corpus appeared to satisfy my two criteria.

5. The Investigation

5.1 Five stages

The investigation was framed into five distinct stages. At first the students were introduced to the problem synonymic pairs by way of a simple dictionary exercise, (Appendix A). With the aid of a Japanese English teacher I then sought to clarify that the grammatical function of each word was clearly understood. For example, the students

6

were instructed to be aware that in the course of this investigation we were looking for the use of "opposite" as an adjective, and not as a preposition. It was also made clear at this stage that in order to keep the investigation manageable, we were not looking at lemmas or word families, but only at instances of the exact words as prescribed in the textbook.

Next the students were presented with a full set of concordances taken from the LOB corpus containing the target words. The students were left to work on these by themselves and to look, specifically, for patterns of usage. During the third stage of the investigation the students were asked to complete a 'cloze' exercise using the target words in a second set of concordances taken from the LOB corpus, (Appendix B). Next, the students were asked to produce sentences of their own based on what, if anything, they had learnt from the earlier work, (Appendix C). As a final part of the investigation there was a group discussion in which the students were asked to formulate, in their own way, some rules of usage as evidenced by the investigation in its entirety.

5.2 Results

As Biber et al note, "The goal of corpus-based activities is not simply to report quantitative findings, but to explore the importance of these findings for learning about the patterns of language use". In a qualitative study such as this, there was no preset goal to seek to establish rules of usage based around quantitative, measurable elements such as frequency. Rather, the idea was to simply use the concordances to expose the students to sections of real language and, in so doing, allow the learners to dig for clues of convention among the patterns which were generated.

With the aid of their dictionaries, the students were all able to complete the first exercise, matching synonymic words, successfully. The second task, however, caused significant problems as the learners were faced with the problem of also having to deal with unknown words in an unknown context. As reported earlier, with no domain-specific corpus to draw upon, it is unavoidable that generated concordances will vary greatly; their diversity reflects the very nature of the multi-registered LOB corpus. As Laufer and Shmueli (1997), remark, "vocabulary acquisition will not occur if unfamiliar words are not attended to: not noticed or not processed deeply". Some of these problems were offset with the help of lengthy discussions involving the students, my Japanese colleague and myself, but even this help was limited as concordances produce 'co-text' but often not the essential 'context'. As Ittzes (1991), in Singleton (1997; 218), and others have found, L2 learners of English are more successful at guessing the sense of words in context rather than words in isolation. Collocations may produce patterns left or right of the target word, *collocates*, but L2 learners may find it difficult to isolate such patterns in their struggle to come to terms with an unknown context.

It would have been possible to spend a great deal of time analyzing the concordances generated by just a single synonymic pair and indeed, this may form the basis for future lessons and further investigations. However, in order to remain focused on the task in hand, the role of the teacher here is to separate the wheat from the chaff. For example, a

7

search for the word 'contrary' produces a raw count of 35 instances but fully half of these are not the adjectival form under investigation, but the idiomatic phrase "on the contrary". When using an *untagged* corpus much pruning needs to be done as lexical words are not clearly distinct from their dual role as functional items.

The third task in the exercise proved inconclusive in the sense that in a fifty-fifty clozed exercise results are always going to be difficult to quantify. Discussions once again revealed the difficulties that the students were having in coming to terms with the context of the concordances. It had been my initial intention to end the investigation at this point, but the inconclusive nature of this exercise forced me to rethink. Taking some advice from Laufer and Shmueli (1997), "Associative context is most beneficial for retention when it is created by the learner himself", I embarked on the fourth part of the investigation, the students were asked to produce their own sentences containing the keywords, and was pleased to report that my students were better able to differentiate the uses of synonymic words than the cloze exercise had revealed, (Appendix D). As Cohen and Aphek (1981), found, associations created by the learner were more effective that those provided by the teacher, or, in this case, the corpus.

The fifth and final part of the study, an open group discussion conducted mostly in L1, and later translated for me, produced the following summary:

Fertile and rich
Adjectives, meaning to have an abundant supply of something. Fertile is the preferred choice when talking about land.

Ripe and mature
Adjectives, meaning to come to full development. Ripe is the preferred choice for fruit and vegetables. Mature is the preferred choice when the subject or object is human.

Sick and ill
Adjectives, meaning to feel unwell or unhealthy. Both have a tendency to collocate with adverbs: suddenly, seriously, extremely, violently…etc.

Fury and anger
Nouns, meaning a strong feeling of hostility. Fury is the preferred choice for personification.

Contrary and opposite
Adjective, meaning marked by a natural or innate and irreconcilable opposition. Contrary is the preferred choice when stressing extreme divergence, opposite marks two things which are altogether different.

Ruin and destroy
Verbs, meaning to cause destruction. To ruin implies irretrievable harm but not necessarily total destruction. To destroy suggests complete destruction.

8

6. Conclusion

6.1 A return to the shared and systems perspectives

As before, we return once again to the *system* vs *shared* perspectives. The *Eiken* test, like many English proficiency tests, may be seen as systems oriented, consisting of standardized multiple-choice exercises based around a focus on grammar and vocabulary. As Dieterich et al comment, (1979) "such tests often have validity weaknesses...they may confuse knowledge *about* a language with ability to *use* a language". In recent years the Ministry of Education in Japan has come under a lot of fire in regard to its insistence on focusing on grammatical correctness rather than on communicative competence. That Japanese students continually prop up the league table of English ability in Asia would seem to highlight this failing. Be this as it may, it is not the purpose of this paper to present a critique of current testing procedures as McGregor (1997), and many others have already done. As teachers we are paid to teach in accordance with syllabi as dictated by our employers. A focus on grammar and word manipulation may not lead to the skills necessary for successful communication, but it may help students to pass such systems-oriented examinations.

There is also a cultural angle that must be dealt with: In the Japanese English language classroom students are simply not used to being asked, "Please take a look at this and try and work out what it means". Rather they are told, "This is the rule – learn it and apply it". This 'system*atic*' approach produces something that can be tested, it produces an academic world in which something is either right or it is wrong with little regard for anything in-between, and it is precisely this "in-between" that CA forces the student to address.

My initial belief that CA could be successfully adopted into the classroom with relative ease underwent its own evolution as the study progressed. The revisions made to my own methodological approach reflect this change. Concordances may say a great deal to a teacher or researcher, but a lot less to a student; supervision and guidance are key components in the learning process. Though I feel I am now better equipped to answer the initial question (What is the difference between ruin and destroy?), with greater clarity than before, it took my students a great deal of time and effort to come to a similar understanding.

My own small investigation would seem to indicate that with students of this age, and at this level of English instruction, concordances produced from a relevant corpus should not be used on their own to help learners to come to terms with the problems of meaning and usage associated with synonymic pairs. If we are to successfully utilize CA in the classroom, then there is a need here for a combination of approaches introduced in carefully organized steps and guided at all stages by the teacher. The "context theory of meaning" proposed by Carroll (1967; 574, 575), "words vary in meaning according to their context", cannot be adopted in isolation, rather a systems approach of "focus over

context" as proposed by Laufer and Shmueli, (1997), must also be employed. If, as Firth suggests, "we know a word by the company it keeps" then the reverse is also true: *we must also know the company in order to understand the word.*

SYNONYMS TEST

Choose the answer which best corresponds to the word above.

1. Fertile

a) rich b) broken c) angry d) special

2. Ripe

a) attractive b) mature c) sarcastic d) tidy

3. Sick

a) ill b) quiet c) great d) fun

4. Fury

a) approval b) chance c) anger d) charm

5. Contrary

a) essential b) equal c) basic d) opposite

6. Ruin

a) destroy b) lift c) demand d) hope

11

12

towards Apalache . They were approaching a *fertile* country , with numerous settlements and plantations
It has been suggested that the *fertile* corn-growing regions in the hinterland also
in the course of an hour's *fertile* imagination . I said as much to

to land returned with armfuls of *rich* grass for the exhausted horses and
called Paracoxi in a land of *rich* maize-fields . \ De Soto forthwith despatched a
they reminded them of Maine . Small , *rich* fields interspersed with fingerlings of forest

countries of Africa that the time is *ripe* to have more frequent consultations between
is the pear that isn't quite *ripe* . And warn your children not to
Alsloot or Sallaert with burghers . The *ripe* hue of the red and dun

example , he has chosen a group *mature* film actors * - men like
that children in co-educational schools often *mature* earlier than those who are segregated
The touchstone , for a man of *mature* years considering what to take up

he returned to England with a *sick* wife in the spring of 1878
marriage . I felt by now extremely *sick* again and practically suffocated with excitement
they both leaned against the well , *sick* and dizzy , hardly able to see

Europe , Clarence Paget had become seriously *ill* with a supposed abscess on the
the India Office on account of *ill* health and other duties in the
Men . In 1941 I was dangerously *ill* with pneumonia in Leeds Castle Hospital

now it's over . The waves , their *fury* spent , are plashing lazily on the
swept through the house with a *fury* that disarranged everything and left a
he . He tried to control his *fury* and his hammering heart by taking

man , " Miss Charlotte said coldly , and *anger* glinted in her pale eyes . " Your
pretended to be completely staggered . Then *anger* came into her face . " What a
his thumb , his thoughts boiling with *anger* and disgust and humiliation . He reached

to all appearances unmoved , which is *contrary* to what might have been expected
such work is commenced or completed *contrary* to the provisions of this section
he was not drinking . This was *contrary* indeed to my expectations , for he

quite mistaken and are having the *opposite* result . There is already ample statistical
below . Spectators are gathered on the *opposite* cliff, cut off from me by
away for their walk in the *opposite* direction , without another glance at her

is important . A fussy referee can *ruin* a bout . There were southpaws galore
things , wobbling at every step ? They'll *ruin* your feet * - and the carpets , too
unsteady .) I thought that Dackson would *ruin* my son . I did not think

had entered the city , ready to *destroy* it . At that moment , the song
It remained for the Nazis to *destroy* its inhabitants . On the walls of
bring Russia down before she could *destroy* America's nuclear bases . It is a

12

13

<u>Appendix C</u>

<u>FERTILE & RICH</u>
1. Picasso was a creative man with a fertile imagination
2. That farmland is so fertile that almost anything will grow there.
3. The Amazon jungle is dense and fertile, full of life.
4. Osaka castle is old and rich in history.
5. Even in the deepest oceans life is rich and diverse.
6. This company has rich experience in business.

<u>RIPE & MATURE</u>
1. Fruit is most delicious when it is fully ripe.
2. The time was ripe for a change of government.
3. I love to eat ripe melons.
4. Parmesan cheese is mature and very tasty.
5. Even though he is young, he is mature for his age.
6. When the trees become mature, the garden will look even more beautiful.

<u>SICK & ILL</u>
1. Owing to ill health, he decided to retire early.
2. The poisonous mushrooms made him seriously ill.
3. After drinking the dirty water she was taken ill.
4. Roller –coasters always make me feel sick.
5. I am worried sick about my new job.
6. I was sick last weekend.

<u>FURY & ANGER</u>
1. The victim was unable to control her anger.
2. In all his life he had never once raised his voice in anger.
3. He could clearly see the anger in her eyes.
4. The waves crashed onto the rocks with a terrible fury.
5. It was impossible to hear anything amidst the fury of the storm.
6. The crazy man flew into a fury.

<u>CONTRARY & OPPOSITE</u>
1. We approached the junction from opposite directions.
2. Two pictures were hung on opposite walls in the living room.
3. That will have the opposite effect.
4. The government acted contrary to its earlier promises.
5. On the contrary, I have always liked gin.
6. His behaviour was contrary to what was normally expected.

<u>RUIN & DESTROY</u>
1. The bomb was able to completely destroy the building.
2. If his wife knew of his secret, it would destroy her.
3. The orchard was destroyed by the typhoon.
4. My father came home late work and his dinner was ruined.
5. Because of her mistakes, she had ruined her chance of winning that day.
6. If you add too much sugar, you will ruin it.

References

Alderson, C. (1996) Corpora are becoming mainstream.
In Thomas, J. and Short, M. (Eds.): *Using Corpora for Language Research.*
Longman: London.

Barlow, M. (1995) Corpora for Theory and Practice.
International Journal of Corpus Linguistics 1/1.

Bateson, G. (1972) *Steps to an Ecology of Mind.*
Ballantine: New York.

Biber, D., Conrad, S., & Reppen, R. (1998) *CORPUS LINGUISTICS: Investigating Language Structure and Use.*
Cambridge: Cambridge University Press.

Carroll, J. (1964) Words, Meanings and Concepts.
In Jakobivits & Miron (1967) *Readings in the Psychology of Language.*

Dieterich, T., Freeman, C., & Crandall, J (1979) A linguistic analysis of some English proficiency tests.
TESOL Quarterly 13: 535-550.
In Brown. H. D. (1993) *Principles of Language Learning and Teaching.*
Prentice-Hall: New Jersey.

Chomsky, N. (1965) *Aspects of the Theory of Syntax.*
MIT Press: Massachusetts.

Clear, J. (1993) From Firth Principles. Computational Tools for the Study of Collocation.
In Baker, M., Francis, G., & Tognini-Bonelli, E. (Eds.): *Text and Technology.*
John Benjamins.

Cohen, A.D., & Aphek, E. (1981). Easifying second-language learning.
Studies in Second Language Acquisition, 3, 221-235.

Collins Cobuild English Language Dictionary (1987).
Collins: London.

Eiken Ni-kyu Kyohon (1972)
Obunsha: Tokyo.

Firth, J. R. (1957) A synopsis of linguistic theory.
Studies in linguistic analysis.
Oxford: Oxford University Press.

14

Gavioli, L. (1997) Exploring Texts through the concordancer: Guiding the Learner.
In Wichmann, A., Fligelstone, S., McEnery, T., & Knowles, G. (Eds.): *Teaching and Language Corpora.*
London: Longman.

Ittzes, K. (1991) Lexical guessing in isolation and context.
Journal of Reading, 34, 360-366.

Johns, T. (1991a) Should you be persuaded – Two Examples of Data-Driven Learning Materials.
English Language Research Journal (4) 1-16.
University of Birmingham.

Kenkyusha's English-Japanese Dictionary For The General Reader (1984).
Matsuda, T. (Ed.).
Kenkyusha: Tokyo

Lancaster-Oslo/Bergen corpus (1978)

Longman Dictionary of Contemporary English (1978).
Longman: London.

Laufer, B and Shmueli, K. (1997) Memorizing New Words: Does Teaching Have Anything to do with it?
RELC Journal, Vol28, No.1, June 1997

McGregor, L. (1997) The *Eiken* Test: An Investigation.
JALT Journal, Vol.19, No.1, May 1997

Mindt, D. (1996) English corpus linguistics and the foreign language teaching syllabus.
In Thomas, J. and Short, M. (Eds.): *Using Corpora for Language Research.*
Longman: London.

Nattinger J. R. and DeCarrico J. S. (1992) *Lexical Phrases and Language Teaching.*
Oxford: Oxford University Press.

Quirk, R., Greenbaum, G., Leech, G., & Svartvik, J. (1985) *A Comprehensive Grammar of the English Language.*
Longman: London.

Renouf, A. (1997) Teaching Corpus Linguistics to Teachers of English.
In Wichmann, A., Fligelstone, S., McEnery, T., and Knowles, G. (Eds.): *Teaching and Language Corpora.*
Longman: London.

15

Sinclair, J. (1991) *Corpus, Concordance, Collocation.*
Oxford: Oxford University Press.

Singleton, D. (1997) Learning and processing L2 vocabulary.
Lang.Teach. 30, 213-225.
Cambridge: Cambridge University Press.

Stubbs, M. (1996) *Text and Corpus Analysis.*
Blackwell: London.

Svartvik, J. (1996) Corpora are becoming mainstream.
In Thomas, J. and Short, M. (Eds.): *Using Corpora for Language Research.*
Longman: London.

16

17

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
Educational Resources Information Center (ERIC)

ERIC

REPRODUCTION RELEASE
(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: AN INVESTIGATION INTO HOW CORPUS ANALYSIS MAY BE USED IN THE SECOND LANGUAGE CLASSROOM TO SOLVE SOME OF THE PROBLEMS SURROUNDING NON-NATIVE SPEAKERS' UNDERSTANDING OF SEEMINGLY SYNONYMOUS WORDS

Author(s): SEAN ROMANO MADDALENA

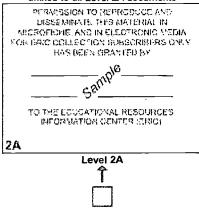Corporate Source: | Publication Date:

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce the identified document, please CHECK ONE of the following three options and sign at the bottom of the page.

The sample sticker shown below will be affixed to all Level 1 documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

_____Sample_____

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

Level 1

☑

Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g., electronic) *and* paper copy.

The sample sticker shown below will be affixed to all Level 2A documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY HAS BEEN GRANTED BY

_____Sample_____

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

2A

Level 2A

☐

Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only.

The sample sticker shown below will be affixed to all Level 2B documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY

_____Sample_____

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

2B

Level 2B

☐

Check here for Level 2B release, permitting reproduction and dissemination in microfiche only.

Documents will be processed as indicated provided reproduction quality permits. If permission to reproduce is granted, but neither box is checked, documents will be processed at Level 1.

I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce this document as indicated above. Reproduction from the ERIC microfiche or electronic/optical media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.

Sign here please

Signature: *S. Maddalena*

Printed Name/Position/Title: MR. SEAN R. MADDALENA

Organization/Address: UNIVERSITY OF ASHIYA, ROKUROKUSO-CHO; ASHIYA 〒659-001 JAPAN

Telephone: 81-798-72-9282 | FAX: 81-797-31-7773
E-Mail Address: Srm@gol.com | Date: 26/10/2001

(HOME ADDRESS:
SHUKUGAWA ANNEX III #301,
6-12 MATSUOI-CHO;
NISHINOMIYA,
〒662-0076.
JAPAN

ERIC
Full Text Provided by ERIC

III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, or, if you wish ERIC to cite the availability of the document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents that cannot be made available through EDRS).

Publisher/Distributor:
Address:
Price Per Copy:
Quantity Price:

IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant a reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

Name:
Address:

V.WHERE TO SEND THIS FORM:

You can send this form and your document to the ERIC Clearinghouse on Languages and Linguistics, which will forward your materials to the appropriate ERIC Clearinghouse.

Acquisitions Coordinator
ERIC Clearinghouse on Languages and Linguisitics
4646 40th Street NW
Washington, DC 20016-1859

(800) 276-9834/ (202) 362-0700
e-mail: eric@cal.org