Intralinguistic (Genre[1]) transferability and prototypicality of meaning of polysemous words: the case of BREAK.[2]

Toru Kinoshita and Tomohiro Sakai

1. Introduction

The problem of the transferability of the various meanings of a polysemous word from one language to another has been studied by many researchers (e.g., Cunningham and Graham, 2000; Kellerman,1978; 1986; Viberg, 1999). The underlying theory of 'core' and 'prototype' has been used in many of psychology and psycholinguistics, such as perception (e.g. Rosch, 1973), cognitive psychology (e.g., Miller, 1978), language development (e.g., Nelson, 1985), semantics (e.g., Jackendoff, 1984, and Pause, Botz, and Egg, 1995), historical linguistics (e.g. Michaelis, 1996 on the diachronic analysis of the adverbial STILL, and Uhlenbeck, 1996, on the distinction between homonymy and polysemy), and by Bell (1998) on the "cancellative discourse markers").

The concepts of 'core' and 'prototype' have been used in the field of second language acquisition as well: Shirai (1990), for example, studied the prototype and metaphorical extensions of the verb PUT. The concepts of core and prototype are not identical: the notion of 'core' is regarded as a context-free, abstract essence of one set (or group), while 'prototype' is seen as the most typical member of a given set defined by a concept of core. Since 'core' is a concept which is free of context, the concrete meaning of each member belonging to a set defined by the core is determined by the context in which the member appears. In other words, the embodiment of core is determined by contextual information. Limiting himself to cases involving polysemous words, Miller (1978) points out two necessary conditions for actualization of the core: "(a) it is possible to identify central or core senses of polysemous words, and (b) it is possible to formulate construable rules governing the ways a core sense can be extended or provide other senses (p.102)."

---

[1] Here the word 'genre' is used without strict definition to refer to field grouping in the Brown Corpus. For example, later codes such as A01, B02, etc., will be given as genre identifiers.   "A" means "Press:

In a slightly different approach to the concepts of core and prototype, Coleman and Kay (1981), for example, define 'semantic prototype' as a 'prelinguistic cognitive schema or image' with which speakers judge the degree to which an object matches as a member of the category represented by the schema or image. Using this semantic concept of prototype, Kellerman (1978) carried out an experiment on transferability judgments between a Dutch polysemous word BREKEN and its English equivalent BREAK. Kellerman's subjects, who were all native speakers of Dutch learning English as second language, judged that the closer to the 'core' a specific meaning was, the more likely the meaning would be transferable from their first language to the target language. That is, it was judged "safe" to do a direct translation from one language to another.

Building on Kellerman (1978), Tanaka and Abe (1985) studied cross-linguistic transferability between Japanese and English. They came to a conclusion to that of analogous to Kellerman. They found that the less idiomatic an expression was, i.e., the closer to the "core" the word was, the more likely it would be judged transferable from one language to another.

In a related study, Caramazza and Grober (1976) claimed that the grammatical category of a word sense did not affect similarity judgments; however, Kellerman (1978) reported that the grammatical category of a lexeme in a polysemous word did indeed affect the transferability judgment. At the same time, however, Kellerman claimed that the degree to which a polysemous word was judged as similar to a corresponding word in another language does not predict the transferability judgment reliably. According to Kellerman, in comparison with a similarity judgment, a prototypicality judgment shows a much stronger correlation with the judgment of transferability.

Kellerman (1986) attempted to improve the predictability of transferability judgments through the use of prototypicality judgments. In order to do this, he introduced the notion of subjective frequency -- how frequently a learner thinks a sense of a word is used regardless of its objective frequency. He used the concept of subjective frequency in addition to a previously-used notion, i.e., prototypicality, which was defined as "perceived similarity to the primary sense of a word (XXX)." Using these two factors simultaneously, he claimed the predictability of transferability judgments could be significantly improved when expressed as a product of these two factors.

Similar to Kellerman's (1986) experiment using the frame work of prototype

Intralinguistic (Genre) transferability and prototypicality of meaning of polysemous words: the case of BREAK.

with regards to the prototypical meanings of PLAY. Ueda (1998b) also reported a similar result in which both Japanese learners of English and English native speakers used the prototypical meanings of PLAY to understand more peripheral ones of the same verb.

The above-mentioned researchers, however, focused mainly on the transferability of individual senses of a polysemous word in an exclusively crosslinguistic / interlinguistic situation. It seems possible, however, to consider an analogous problem of transferability in an *intra*-linguistic situation as well. Intralinguistic transferability here refers to the possibility that a possible meaning of a polysemous word used in one particular genre can also be used in another genre. In the present study, "genre variability" of senses in polysemous words will be considered equivalent to and used interchangeably with the term "intralinguistic transferability".

There seem to be a very small number of studies conducted concerning the 'intralinguistic transferability' of any given sense of a polysemous word. Evidence for this paucity comes from the relatively small number of intralinguistic transferability studies in large-scale data base searches. For instance, the *Language and Linguistic Behavior Abstract* regularly scans domestic and international journals and books in language-related research and has accumulated 234,000 records (as of 2001). This database holds 182 records for the key words 'polysemous word(s)' and 1856 records for the key word 'genre.' However, this same database has only twoo records containing both of these two keywords found together.

Based on the theoretical framework of contextual analysis (Celce-Murcia, 1980; 1990), the present study, therefore, attempts to investigate the relationships between the intralinguistic transferability (or genre variability) and the prototypicality of each lexeme of the polysemous verb BREAK[3] using (a) native and near-native speakers' typicality judgments on lexemes in given sentences, and (b) subjects' most typically generated sentences as measurements of prototypicality for the various senses of BREAK.

Research questions and hypothesis

The study attempts to answer the following research questions:
(1) What is a working representation of various meanings of BREAK?
(2) What is the most prototypical sense of BREAK?

(3) What is the relationship between prototypicality and intralinguistic transferability?

The hypothesis

The following hypothesis is tested in this study:

The more 'core-like' or the more 'prototypical' a meaning is, the more widely it is used across various genres, i.e., the higher the degree of intralinguistic transferability.

2. Method

The present study was conducted following four major steps. First, a classification of the meanings of BREAK was broadly constructed based on several dictionaries' categorization of them. Second, the actual use of each sense of the verb BREAK in a large-scale database was observed using the Brown Corpus, which contains more than 1,000,000 written words.

It is true that there are much larger corpora of English than the Brown Corpus, such as the Bank of English and the British National Corpus (see   2001 for the status quo of larger corpora). It is also true, however, that the Brown Corpus is still the major reference corpus for many researchers given its design and principles of construction its word frequency information, and its provision for targeted words in an immediate context. Thus, Hennoste, Koit, Roosmaa, and Saluveer (1998) created the first written Estonian corpus based on LOB and Brown Corpora design principles. Sinclair (1995) also adopted the original principles of the Brown Corpus with minor modifications to describe and classify various large scale linguistic corpora. As for the use of the frequency information, Grabowski and Mindt (1995), for instance, relied upon the Brown Corpus and the LOB corpus to compile a learning list of irregular verbs in English, and they confirmed the list with Longman/Lancaster English Language Corpus, as well. Furthermore, Burgess and Livesay (1998) compared the word frequency estimates from the Brown Corpus (about 1 million words) and those of the HAL corpus (about 130 million words) in terms of their predicting abilities for word recognition time and found that the two corpora have virtually no difference as far as their high-frequency words (such as BREAK) are

justifiable to use the Brown Corpus for the current study in spite of the relatively early date of its construction.

Third, each meaning of the verb BREAK was evaluated in terms of prototypicality based on (a) the 'receptive typicality' judgment test by linguistically sophisticated native and near native speakers of English, and (b) the 'productive typicality' test given to the same subjects. Finally using descriptive statistics, the present study examined whether meanings judged 'prototypical' were indeed used more widely in various genres than other senses judged less prototypical.

2.1. Classification of the meanings of the verb BREAK

According to Tanaka (1987) and Shirai (1990), There are three major approaches to the classification of the various meaning of polysemous words, i.e., (a) the dictionary approach, (b) the feature approach, and (c) the core-prototype approach. The dictionary approach starts with collecting as many samples as possible which contain the target word. Those samples are then divided into multiple groups. After the grouping has been verified, each sense group is named according to its most representative meaning.

The feature approach to the categorization of multiple meanings of a word dissects those meanings into basic features so that a combination of the existence and the non-existence of each feature will define the specific meaning of one sense. Shirai (1990) cites Jackendoff (1983), who illustrates this approach using PUT, as in "Amy put the flower in the vase" in terms of [+CAUSE] [+ GO] and [+ positional/+circumstantial].

The core and the prototype approach set the core meaning and the prototypical example of a target word at the center and arrange various meanings in concentric circles until the examples at the outer-most borders can no longer be regarded as examples of the word.

The three approaches discussed above have both positive and negative aspects. Tanaka (1987) criticizes the dictionary approach by saying that its enumerating nature makes intergroup relations opaque. The feature approach is of interest due to its binary nature which makes it possible to classify a virtually infinite number of separate entities based on features. As the number of meanings of a word increases, it is necessary to increase the number of features in order for each meaning to possess a unique

Toru Kinoshita and Tomohiro Sakai

This approach has another problem involving feature expansion, since each word may need a completely different set of features from others to distinguish its own meanings. Thus, although each word may need just a few features, the total number of features necessary for classifying all words could become quite large.

The core and prototype approach also has problems stemming from the operationalization of prototypicality and coreness. Givón (1984), for example, discusses the limitations of the core approach by arguing that there is almost always something that the context-free core of a word alone cannot cover. In other words, Givón sees the meaning of words as "a mixed compromise system," and admits that there are some meanings which are context-free and others which are context-dependent.

In terms of feasibility and theoretical superiority, the present study uses a combination of the dictionary and the core and prototype approaches, since (a) sample classifications are readily available from a large number of dictionaries and (b) the core / prototypical approach has shown some success in previous studies. The following are some examples of the core and prototype approach and the dictionary approach for various meanings of the verb BREAK.

The core and prototypical approach
   (A). Tanaka and Kawade (1987)
       1. (with movement) interruption
       2. (without movement) damage

The dictionary approach
   (B). West's (1953)' *A General Service List of English Word s*
       1.fracture
           e.g.,. break the window/ break the skin/ The rope broke
       2.figurative
           e.g., break her heart/ break the attack
       3.not be retained by
           e.g., break the prison/ break a contract
       4.interrupt
           e.g., break one's journey/ break the silence
       5.phrasal expressions
           (Phrasal expressions will not be included in this study.)

   (C) . *Oxford Advanced Learner's Dictionary of Current English* (Hornby, 1974)
       1. (of a whole thing) (cause to) go or come into two or more separate parts as the result of force, a

e.g., He broke a branch from the tree.
3. make something useless by injuring an essential part (of a machine, apparatus, etc.)
e.g.,break a clock
4. (with adjective). break even, break loose etc.
(Phrasal expressions will not be in this study.)
5. (with various subjects):
e.g., The abscess blister bubble broke.
6. (with various objects):
e.g., break somebody's back.
7. train or discipline
e.g., break a horse to a harness
8. subdue, keep under, end by force:
e.g., break the enemy's resistance
9. act in opposition to; infringe:
e.g., break the law
10. interrupt or destroy the continuity of; end the operation or duration of
e.g., break the silence by speaing
11. (with adverbial particles and prepositions) phrasal expressions
(Phrasal expression will not be used in this study.)

(D) *Longman Dictionary of Contemporary English* (Procter, 1978)
1. to (cause to) separate into parts suddenly or violently, but not by cutting or tearing:
e.g., to break a window / a leg. The rope broke when they were climbing.
2. to (cause to) become separated from the main part suddenly or violently, but not cutting or tearing:
e.g., to break a branch off a tree.
3. to (cause to) become unusable by damaging to one or more parts:
e.g., He broke his wristwatch by dropping it.
4. to (cause to) become, suddenly or violently:
e.g., The prisoner broke free/loose.
5. to open the surface of:
e.g., to break the skin/the soil.
6. to disobey; not keep; not act in accordance with:
e.g., to break the law/a promise
7. to force a way (into,out of, or through):
e.g., He broke into the shop.
8. to bring under control:
e.g., to break a horse/ a child's spirit
9. to do better than:
e.g., to break a record in sports

        e.g., We broke the enemy at the battle of Harlow Fields.
12. to make known (especially something bad.)
        e.g., Break the bad news gently to him, please.
13. to interrupt (an activity):
        e.g., We broke our journey to Rome at Venice.
        e.g., The bushes will break his fall.
14. to (cause to) come to an end:
        e.g., to break the silence; The cold weather at last broke at the end of March.
15. to come especially suddenly into being or notice:
        e.g., As day breaks; The storm broke.
16.to fail as a result of pressure from inside or outside:
        e.g., His health broke. He may break under continuous   questioning.
17.to (cause to) change suddenly in direction, level, loudness:
        e.g., His voice broke with strong feeling
18.to discover the secret of :
        e.g.,.She broke their code(=secret writing).

Among these various categorizations, the present study used the *Longman Dictionary's* system for its basic classification of the meanings.[4] The choice of the *Longman Dictionary* is partly due to the fact that this dictionary is made for second language learners so that it is easier to identify the meaning of a certain use of BREAK in this system than in more complicated ones. Secondly, what is necessary to differentiate among the various meanings of a word in order to master the word in second language learning may in fact reflect better a native speaker's mental categorization of the meanings of the word. Finally, Longman's list of definitions seems to be appropriate as indicated by Kellerman's (1986) use of it to define the prototypical meanings of BREAK. *The Oxford Advanced Learner's Dictionary* (OALD) (Hornby, 1974) is also a learner dictionary, but the *Longman Dictionary of Contemporary English* is more recently compiled and thus reflects the present-day use of words better than the OALD.

In order to make the Longman system applicable to the data used in this study, however, a modification to the Longman system was devised. This was partly because the number of categories seems to exceed the observed uses in the Brown corpus and partly because some differentiations seemed to be too minute to be practical for investigation in the present study. The following is an adapted categorization based on the *Longman Dictionary of Contemporary English* (Procter 1978).

Intralinguistic (Genre) transferability and prototypicality of meaning of polysemous words: the case

of BREAK.

(E) **Modified Longman system**

G1[5] Destruction[6]

1[7]. to (cause to) separate into parts suddenly or violently, but not by cutting or tearing: e.g., to break a window / a leg. The rope broke when they were climbing.

2. to (cause to) become separated from the main part suddenly or violently, but not cutting or tearing: e.g., to break a branch off a tree.

10.to ruin: e.g., if that young man tries to marry my daughter, I'll break him!

16. to fail as a result of pressure from inside or outside: e.g., His health broke. He may break under continuous questioning.

G2 Damage

3.to (cause to) become unusable by damaging one or more parts:

e.g., He broke his radio.

G3 Exposure/ uncovering

5. to open the surface of: e.g., to break the skin/the soil.

12. to make known (especially something bad.): e.g., He broke the bad news to her.

18.to discover the secret of : e.g.,.She broke their code(=secret writing).

G4 Go against expectation/ convention

6.to disobey; not keep; not act in accordance with: e.g., to break the law/a promise

G5 Overpower / control

8. to bring under control: to break a horse/ a child's spirit

11. to destroy as an effective force: e.g., We broke the enemy at the battle of Harlow Fields.

9. to do better than: e.g., to break a record in sports

G6 Interruption

13. to interrupt (an activity): e.g., We broke our journey to Rome at Venice. The bushes will break his fall.

G7 (+/- change in state or existence)

14. to (cause to) come to an end: e.g., to break the silence; The cold weather at last broke at the end of March.

15. to come especially suddenly into being or notice: e.g., As day breaks; The storm broke.

17. to (cause to) change suddenly in direction, level, loudness: e.g., His voice broke with strong feeling

G8 None of the above

One note which seems necessary here is that Sense Group 1 does not differentiate between physical and figurative ("imagery") objects. This is based on previous research by Kellerman (1978, 1986). Kellerman reports that the degree of prototypical sense or

---

attribute shared by various meanings is minimally influenced by a consideration of *concreteness* or *imagery*. Kellerman (1978) elaborates as follows:

Senses of break as represented by *She broke his heart* or *The accident left him a broken man* are seen as more related to the 'prototypical sense' than such concrete or high imagery senses as contained in *The waves broke on the shore* or *The tree broke his fall*.(p.38).

Also, with regard to idiomatic or phrasal expressions, all expressions listed as idioms and phrasal verbs in Longman and Cobuild or judged as such by a sophisticated native speaking researcher were excluded from further analysis. This follows Kellerman's (1978: 38) methodology.

2.2. Data

2.2.1. Corpus

In this research, the Brown Corpus[8] was used as the source of data. Computer software called the Oxford Citation Program was used to select all sentences containing any form of BREAK in the form of a concordance with tags indicating the genre of the source.

2.2.2. Limiting the data and the classification of the sentences from the Brown Corpus.

This study used only the "simple" forms of BREAK based on Kellerman (1978). This means that the present study excluded the cases using BREAK as (1) nouns, (2) phrasal verbs, (3) prepositional expressions, or (4) a combination of both phrasal and prepositional expressions. In addition, cases of gerundive use to form a compound noun were also excluded because of their noun-like nature.

In order to decide whether each example of BREAK should be included in the data for analysis in the current study, the following steps were taken. First, using the tag information in the Brown Corpus regarding the part of speech of the target word, all examples of BREAK classified as nouns were eliminated, even if they co-occurred with tokens of BREAK used as verbs in a given segment of output from the Brown Corpus.

Second, the remaining sentences containing BREAK were examined individually.

Intralinguistic (Genre) transferability and prototypicality of meaning of polysemous words: the case of BREAK.

sentences were classified according to the meaning of BREAK in each sentence, using the modified Longman system of classification. Whenever the researcher was not certain of how to classify a token of BREAK, it was discussed with linguistically sophisticated native English speakers so that an appropriate classification could be made.

Toru Kinoshita and Tomohiro Sakai

2.3. The receptive prototypicality judgment test

2.3.1. Subjects

Thirty-seven linguistically sophisticated native speakers and near-native speakers of English who were either graduate students or faculty members of the Applied Linguistics Department at the University of California, Los Angeles (UCLA) participated in this test in June 1991.

2.3.2. Instrument

A five-point Likert scale-type questionnaire adapted from Kellerman (1978) was used. See Appendix A.

2.3.3. Procedure

The subjects were asked to judge the 'typicality' of each sense of BREAK. The questionnaire sheets were collected after one week.

2.3.4. Analysis

The left-most column of Appendix A shows the sense group to which each sentence belongs. When more than one sentence was classified as a member of a sense group, the scores of these sentences were averaged. The results of this test were summarized as the receptive prototypicality scores in Table 4.

2.4 The productive prototypicality test

The productive prototypicality was operationalized as the degree to which a meaning is used in sentences generated by the subjects in a context-free environment. Thus, if a subject often generated sentences in which BREAK meant the *destruction* of something physically, for example, then sense group 1 would be regarded more typical than other meanings for that subject. (See Table 4. for results).

2.4.1. Subjects

The same as those used for the receptive prototypicality judgment test except the number was reduced to 26.

2.4.2. Instrument and procedure

2.4.3. Analysis

The sentences produced by the subjects were classified according to the seven sense groups. Then a weighted score was given to each sentence based on the reversed order of production: the first sentence, the second sentence, and the third sentence were given 3 points, 2 points, and 1 point, respectively. The sum of the weighted scores for each sense group was then given to each respective sense group. The total sum for each sense group was used as its productive prototypicality judgement index.

3. Results

3.1. BREAK as a whole in the Brown Corpus

Table 1 shows the total number of "inflected forms" of the verb BREAK which were initially extracted from the Brown Corpus for this research. Although not surprising, Table 1 and Table 2 (below in the next section) indicate that the most frequently used form among the inflected forms of the verb BREAK is "broke," or the past time form. The number of occurrences of "broke" is greater than that of "broken", which includes both adjectives and past participles. Type-token ratios and other related information are also provided in Table 1.

Table 1. Summary of Frequencies of BREAK in The Brown Corpus Including Type / Token Ratio

Variations	Freq.	Total	vocab.(%)	words(%)	Freq.(%)
BROKE*	1	1	16.67	0.38	0.38
BREAKS	12	13	33.33	5.00	4.62
BREAKING	25	38	50.00	14.62	9.62
BROKEN	63	101	66.67	38.85	24.23
BROKE	71	172	83.33	66.15	27.31
BREAK	88	260	100.00	100.00	33.85

Type/Token Ratio		0.023077	Total Words Picked		260

Toru Kinoshita and Tomohiro Sakai

Note. Freq. = Words in Frequency; Total = Word Total; Vocab.%=% of vocabulary; Words(%)=% of Words; Freq.(%)=% of Words in Frequency; BROKE* is an adjective meaning not having money.

3.2.Results of the reduction procedure.

Each category of BREAK is shown below in Table 2. Underscored categories indicate those actually used for the analysis of the meaning of BREAK for each sentence. As a whole, cases were moved to the next step of analysis in terms of classification of meaning (discussed below) and in terms of the genres in which they were used.

Table 2 Breakdown of BREAK in the Brown Corpus

Category	Frequency[9]	Category	Frequency[10]
Break-Noun	23	Break-Phrasal	32
Break	30	Breaks-Noun	3
Breaks-Phrasal	4	Breaks	4
Broke-Phrasal	47	Broke	19
Breaking	25	Broken-Phrasal	16
Broken	41		

3.3. The classification of the meaning of BREAK.

The total number of instances of BREAK in the corpus was 260 and 116 of them were used. Based on the reduction procedure explained in the method section, each observed use of BREAK (N=116) was identified using the adapted Longman system.Table 3. in Appendix A given the cross-tabulations of sense group and the genres in which they appeared in the Brown Corpus.

The columns of Table 3 contain the seven sense groups (plus an eighth, for "none of the above") and the rows indicate the genres specified by the Brown Corpus. A descriptive label for each genre is given in Appendix B.

The row total (See the right most figure next to cell A8) reports 12 actual observations of genre A. The column total (See the bottom of the first column under cell R1) shows 57 (49.1 %) observed cases for sense group 1, destruction. Looking at the column totals for each sense group and their percentages, it is clear that the absolute

majority of the meanings of BREAK used in the Brown Corpus belong to the first sense group, i.e., 'destruction'.

Table 4 below summarizes the genre frequency of each sense group, the receptive and productive prototypicality index, and the absolute frequency of each sense group.Genre frequency refers to the number of genres in which a particular sense group was used in the Brown Corpus. For example, the first row of Table 4. shows that this meaning was used 56 times in the corpus altogether, and that it appeared in 11 genres.

As mentioned before, Table 4. indicates that the 'destruction' sense acounts for a clear majority of the occurrences of BREAK. However, whether one can say that this sense is more widely used than other meanings interms of genre is not so clear.It is certain that this sense group appears in more genres than any other meaning since the genre frequency of this sense group is higher than that of any other group. However, this result could be due to the absolute numbers. In other words, other meanings did not appear in more genres simply because their absolute frequencies were not large enough.[11] On the other hand, it seems quite possible to say that the fact that this sense group accounts for the majority of tokens (and appears in more genres than any other sense group) means that English writers feel safe in using it in more genres than others. Further investigation using the prototypicality of each sense will help sort out these conflicting views.

Table 4. Results of absolute frequency, genre frequency and prototypicality

Sense group	Absolute Freq.	Genre Freq.	Productive proto- typicality	Receptive proto- ypicality
1 (destruction)	56	11	100	4.23
2 (damage)	8	7	5	4.57
3 (exposure)	8	7	0	2.73
4 (unexpected)	9	9	18	3.53
5 (overpower)	12	6	0	2.95
6 (interruption)	7	5	6	1.89
7 (change of existence)	12	8	10	2.34

[11] Unfortunately, the data of the present study potentially might when these conditions for the use of

---

As a supplement to the current study, 86 sentences containing BREAK and its variations were selected using COBUILD Direct (a partial corpus of the Bank of English). Those sentences were then classified into the same seven sense groups as in Table 4, following the same procedures as have been explained so far. The Spearman rank order correlation between the rank order of the absolute frequencies of the seven sense groups obtained from COBUILD Direct and the rank order obtained from the Brown Corpus as shown in Table 4 turned out to be 0.778 ($p < 0.001$). The relatively strong correlation between the two corpora added another piece of evidence for the justification of using the Brown Corpus for the present study.

3.4. The prototypicality of each sense group

Table 4 summarizes the results of two prototypicality judgement tests, as well. The productive prototypicality and the receptive prototypicality measured and computed as explained in Section 2.3. are reported in the fourth and fifth columns of Table 4, respectively. For example, the first row of the fourth column indicates that the sum of the weighted scores given to the generated sentences containing the meaning of *destruction* for BREAK was 100, while the second row of the fourth column shows that the sum of the weighted scores for *damage* was 5. According to the fourth column of Table 4., the sense group characterized as *destruction* was regarded as much more "typical" than the sense group characterized as *exposure* or the sense group characterized as *overpower*.

Similarly, the fifth column (receptive prototypicality) shows subjects' evaluation of the prototypicality of contextualized tokens of each sense group for BREAK in the receptive situation. The computation of these scores is explained in Section 2.3. The last column of the first row shows that the *destruction* group received 4.23, which is the second highest among these seven senses. This means that *'destruction'* was judged by the subjects to be the second most typical meaning among the seven meanings, while Sense Groups 6 and 7 were judged the least and the second least typical meanings of BREAK.

3.5. The relationship between genre frequency and prototypicality

Table 4 is the most comprehensible summary of results. In order to closely examine the relationships among the genre frequency based on each sense group and the

'destruction' sense is used in eleven genres of the Brown Corpus while the 'interruption' appears in only five genres in the same corpus. Figure 2 shows the results of the computed productive prototypicality for each sense group. The horizontal axis of Figure 2 refers to the Productive Prototypicality, computed according to the procedures described in the methods section. The vertical axis of this figure shows the seven sense groups. It should be noted that the order of the sense group in Figure 2 is arranged not by magnitude of productive prototypicality but by order of genre frequency of each sense group (as in Figure 1).

Similarly, Figure 3 is a graphic representation of the results of the prototypicality judgement for each sense group. The horizontal axis provides the measured receptive prototypicality. The order of the sense groups in this figure is, again, not that of genuine receptive prototypicality but that of genre frequency for each sense, as in Figures 1 and 2.

Figure 4 attempts to capture the relationship between the two types of prototypicality used in the present research. The horizontal axis of Figure 4 indicates the receptive prototypicality, like Figure 3. The sense groups in Figure 4 are arranged in accordance with the descending order of productive prototypicality. Thus, if there were perfect agreement between these two types of prototypicality measures, the higher the vertical position of a sense group, the longer its bar would be. In this chart, it is possible to examine this agreement to some extent. Nevertheless, it is obvious that there is some discrepancy between the two measures of prototypicality. Particularly 'damage' and 'interruption' seem to be problematic if we believe that there is only one prototypicality attached to each sense of BREAK in the human brain.

4. Discussion and Conclusion

The present study attempted to answer the following three research questions:

(1) What is a working categorization of the various meanings of the polysemous verb BREAK?

(2) What is the prototypicality of each sense of BREAK?

(3) What is the relationship between the prototypicality and intralinguistic transferability of senses of BREAK?

> The more 'core-like' or 'prototypical' a meaning is, the more widely it is used across various genres, i.e., the higher intralinguistic transferability it has.

With regard to the first research question, the categories of the seven sense groups modified from the *Longman Dictionary* proved useful. In only a very few cases was it difficult to determine a categorization.

The second research question was partly answered by the two tests measuring the productive and the receptive prototypicality associated with each sense group. As Table 4. and Figure 2 indicate, the prototypicality measured by the production test showed that the 'destruction' group was clearly the most prototypical sense group. The table and figure also indicates that the second most prototypical sense group was 'unexpected', i.e., 'go against expectation' (sense group No.4), although there is a big gap between the first and the second senses. The same table and figure also show that the 'exposure' group and 'overpower' group turned out to be the least prototypical. Neither occurred among the three freely-generated sentences produced by the subjects in the research. The same test also indicated that, in terms of productive prototypicality, the second (or third) least prototypical group was the sense of 'damage'.

Table 4. and Figure 3 demonstrate that the 'damage' and 'destruction' groups were, respectively, the most and the second most prototypical meanings of BREAK according to the receptive prototypicality test. The same test determined that 'interruption' of activity, and 'change of existence, etc.' groups were considered the least prototypical sense groups.

Judging from the results of the two types of the prototypicality tests, it seems safe to say that the 'destruction' sense is one of the most prototypical among the various meanings of the verb BREAK, because it was this sense that ranked as the first and second most prototypical meaning on both tests. There are at least two additional reasons to assume that the meanings included in 'destruction' constitute the core meaning of the word. One such reason is that Kellerman's (1978) multidimensional scaling[12] indicated that the sentences he used with "break one's leg" or "break the cup" are more "core"-like in meaning than others, such as "break the code" or "to break the surface". Another reason is the way in which dictionaries arrange the major meanings of the word BREAK. In *The Longman Dictionary of Contemporary English, The Oxford Advanced Learner's Dictionary of Current English, Collins COBUILD English Language Dictionary,* and

General Service List by West, the meaning that Group 1 represents is treated as one of the most important or the most basic (thus "core"-like) meanings of the word BREAK, by listing it as one of the first three meanings.

However, the results obtained from the productive prototypicality test and receptive prototypicality test did not yield clear-cut answers to the original research questions. That is, the research question, consciously or unconsciously, assumed that it would be possible to say that the prototypicality of a certain sense group is one thing, while that of another sense group is another. However, Figure 4 now allows us to visualize the existence of a disagreement between the two types of measurements, so we now have to ask which type of prototypicality is presumed in this question.

Of course, what is suggested by the disagreement itself is ambiguous. One possible interpretation would be that there is in fact only one theoretical trait named prototypicality here, and it is quite natural to see some discrepancy given that the two types of measurement must include some amount of measurement error. The second interpretation is that there is indeed no single entity called prototypicality. Instead, there are at least two different types of prototypicality. Although it is true that a certain amount of error must be involved, the two tests yielded different results because they were actually reflecting the existence of the two different kinds of prototypicality.

One explanation is that the subjects may have thought that the receptive prototypicality test was asking them to imagine which meaning they think other people would most typically use. This interpretation of the prototypicality test would transform it into a kind of productive prototypicality test. Similarly, it could mean they used the same information as for the productive prototypicality even if there was another set of information for receptive prototypicality.

On the other hand, there seem to be several reasons to assume two sets of prototypicality. One reason to believe more than one entity of prototypicality for each word comes from the size discrepancy between one's passive and active vocabulary. It is quite common for people to recognize more words in a receptive or passive task than they actually produce spontaneously. This suggests that there are words for which only the receptive side of the information exists in the lexicon. Further, this could mean that the same kind of phenomenon exists for each nuance/separate meaning of a word. Only a

Either interpretation could be valid in terms of general research design. The present research cannot determine which interpretation is more suitable based on objective evidence. (These interpretations could be translated into corresponding models and would be statistically testable if it were appropriate to use some sophisticated statistical modeling techniques, such as structural equation modeling.)

The answer to the last and central research question of this study, i.e., identifying the relationship between the genres and the sense group with regard to prototypicality, was supplied by Figure, 2 and 3. Taking the discrepancy discussed in the results section related to the second research question, if productive prototypicality is used as an index of prototypicality, Figure 2 shows a relatively clear correspondence between intralinguistic transferability and prototypicality. Sense groups are arranged vertically based on descending order of the genre frequency of each sense. While 'interruption' appears as somewhat of a misfit and an exception, the other six sense groups show clear agreement between the productive prototypicality and the genre frequency associated with each sense group found in the Brown Corpus data. Consequently, the hypothesis was supported in that the more (productively) prototypical a sense group, the more widely that particular sense group was actually used in various genres in the Brown Corpus, i.e., the more intra-linguistically transferable it was.

The relatively clear relationship between intralinguistic transferability and prototypicality, however became somewhat less salient when receptive prototypicality is used instead of productive prototypicality. It may well be possible to recognize some kind of agreement between the order of receptive prototypicality and genre frequency as well. Nevertheless, such a relationship seems to be somewhat weaker than that of productive prototypicality, with 'damage' and 'change of existence' as the two outstanding exceptions to the pattern.

One immediately emerging question is why the productive prototypicality showed a stronger agreement with genre frequency than the receptive prototypicality did. One possible explanation of this could be as follows. Comparatively speaking, productive prototypicality obtained by asking subjects to generate sentences with BREAK should more directly reflect actual use of the word in a large scale corpus than receptive prototypicality. This is because producing several sentences is, in fact, nothing but an actual use of them. On the other hand, asking subjects to rate how prototypical each sense

Intralinguistic (Genre) transferability and prototypicality of meaning of polysemous words: the case of BREAK.

In other words, not only the subjects in the present research produced the sentences with the verb BREAK, but also the writers in the Brown Corpus produced each meaning of BREAK in a similar manner. Neither group chose meanings based on the given senses. In contrast, subjects were given the full range of various senses when they were asked to choose which sense of BREAK looked most or second most prototypical in the receptive prototypicality judgement.

Interestingly, the relatively strong agreement between productive prototypicality and genre frequency for each sense could potentially contradict one of the conclusions of Shirai (1990) vis-a-vis the verb PUT. The productive prototypicality operationalized in the present study was also used in Shirai. Furthermore, Shirai also used the Brown Corpus as one of his two major sources of data. In addition, both studies used TESL/ Applied Linguistics students at UCLA as subjects. Yet, Shirai reports that the sentences most typically produced by his subjects are not in accordance with those actually found in the Brown Corpus, though his study says this tendency is stronger in speech data taken from the UCLA Oral Corpus. In other words, productive prototypicality is not a good predictor of the absolute frequency of the sense groups of PUT used in the Brown Corpus.

One possible account for the gap between the productive prototypicality judgement and that of the Brown Corpus data is that the productive prototypicality judgement test leads subjects' attention to the full range of grammatical forms. Unlike these subjects, the authors of the samples in the Brown Corpus do not seem to have paid special attention to each form of the tense/aspect system of the verb.

It is, of course, quite legitimate to attribute the different findings of the two studies to the idiosyncratic nature of these studies. Indeed, we used different words (BREAK vs. PUT) as targets, different computational methods for determining the productive prototypicality (weighted and non-weighted scoring for each elicited sentence based on their order of occurrence from each subject), and, above all, had different research focuses (genre frequency vs. absolute frequency).

In particular, regarding the last difference between Shirai (1990) and the present study, it is noteworthy that Table 4. demonstrates that the present study data also shows a gap between the productive prototypicality of sense groups and their absolute frequencies.

Consequently, on the one hand, the present study confirms the findings of Shirai (1990). On the other hand, this study suggests that (productive) prototypicality and genre frequency may not be as different from each other as are productive prototypicality and absolute frequencies. That is, productive prototypicality may not predict how often a particular category of sense is actually used in corpus, but it may be able to predict how widely that category is used across various genres. Whatever the reason for the contradictory findings, the existence of the difference suggests that further studies should be conducted.

As for theoretical and methodological problems and future improvements, first, a more systematic way for defining each meaning of BREAK should be sought. For this purpose, a cluster analysis using various features, such as [+/- movement], [+/- agent], etc. might be promising, although the problems associated with feature analysis have been discussed in Section 2.

One related problem is to seek out some methods for incorporating the idiomatic and / or phrasal expressions which were excluded from the current study. Their exclusion was, in one respect, necessary in order to reduce the potential ambiguity of the meaning of the verb BREAK in each sentence following conventions adopted by Kellerman (1978). One undesirable side-effect of these conventions is, however, that they greatly reduce the sample size. Thus, since a larger sample size is more desirable, some procedures should be sought which might enable us to handle sentences including at least BREAK and some prepositions. Another alternative to increasing the sample size of the data would be to use much larger corpus databases or to combine them provided this would not cause other undesirable side-effects.

Second, a more systematic way of establishing a prototypicality scale is necessary. As discussed before, it might be the case that different measurements tap different types of prototypicality and / or different aspects of one prototypicality associated with each sense group. Larger-scale multidimensional scaling, a larger sample elicitation experiment of typical sentences using BREAK, and some indicator such as a card-sorting task[13] should be considered for the improvement of studies such as the one reported here.

Third, in order to see the effects of 'authors' idiosyncrasy and the text from which the samples were taken to form the Brown Corpus, loglinear modeling should be

considered. The loglinear modeling procedure could circumvent the potential violation of using chi-square tests discussed before (E. Hatch, M. Shall, personal communication).

Fourth, to enhance the reliability of a similar study, elicitation techniques should take into consideration the effects of grammatical forms (e.g., tense/aspect) of BREAK in the instructions to the task as well as cultural and schematic context. One possibility is to present all five forms of BREAK (break, breaks, breaking, broke, will break) in the instructions and ask subjects to generate sentences freely for each of the five forms. Different tense and aspect might produce quite different results, just as the context and discourse would do (M. Celce-Murcia, personal communication).

In conclusion, the present study partially confirmed the central hypothesis in that the productive prototypicality of each sense group indeed showed an agreement/ association with the pattern of genre frequency although receptive prototypicality showed the same pattern to a much lesser extent. Thus, the more (productively) prototypical a sense of BREAK is, the more widely it is used in various genres, i.e., the more intralinguistically transferable it is.

Finally, the present study has not resolved all the theoretical and/or practical issues treated here. Among these issues are methodological limitations and possible improvements on them, partly different findings from related studies such as Shirai (1990), the potential extension of the external validity by using alternative corpus data, and any other hidden pit-falls, or even potentially valuable insights. All of these certainly necessitate further study.

References

Bell, D. M.(1998). Cancellative discourse markers: A core/periphery approach. *Pragmatics.* 8, 4, 515-541.

Berglund, Y. (1997). Future in present-day English: Corpus-based evidence on the rivalry of expressions, *ICAME Journal* 21, 7-19.

Burgess, C. & Livesay, K. (1998). The effect of corpus size in predicting reaction time in a basic word recognition Task: Moving on from Kucera and Francis. *Behavior*

*round table on language and linguistics (*pp.181-206). Washington, D.C.: Georgetown University Press.

Celce-Murcia, M. (1980). Contextual analysis of English: Application to TESL. In D. Larsen-Freeman (Ed.) *Discourse analysis in second language research* (pp. 41-55). New York: Newbury House.

Celce-Murcia, M. (1980). Data-based language analysis and TESL in J. E. Alatis, (Ed.) *Linguistics, language teaching, and language acquisition: The independence of theory, practice, and research Proceedings of the 1990 Georgetown University Round Table on Languages and Linguistics.* (pp. 245-259). Washington, D.C.: Georgetown University Press.

Coleman, L. & Kay, P (1981). Prototype semantics: the English word *lie. Language, 55*, 26-44.

Collins (1987). *COBUILD English language dictionary.* London and Glasgow: William Collins & Sons Co Ltd.

Cunningham, T. H. & Graham, C. Ray (2000). Increasing Native English Vocabulary Recognition through Spanish Immersion: Cognate Transfer from Foreign to First Language *Journal of Educational Psychology*, 92, 1, Mar, 37-49.

Francis, W. N. & Kuçera, H. (1982). *Frequency analysis of English Usage.* Boston: Houghton-Mifflin.

Fries, U., Muller, V., & Schneider, P. (Eds.), (1997). *From Aelfric to the New York Times: Studies in English Corpus Linguistics,* Amsterdam: Rodopi.

Givón, T. (1984). Deductive vs. pragmatic processing in natural language. In W. Kintsch, J. Miller, & P. Polson (Eds.), *Method and tactics in cognitive science.* Hillsdale, N. J. : Lawrence Earlbaum Associates.

Grabowski, E. & Mindt, D. (1995). A corpus-based learning list of irregular verbs in English, *ICAME Journal* 19, 5-22.

Hatch, E. & Lazaration, A. (1991). *The research manual: Design and statistics for applied linguistics.* New York: Newbury House Publishers.

Hennoste, T., Koit, M., Roosmaa, T., & Saluveer, M. (1998). Structure and usage of the Tartu University corpus of written Estonian, *International Journal of Corpus Linguistics,* 3, 2, 279-304.

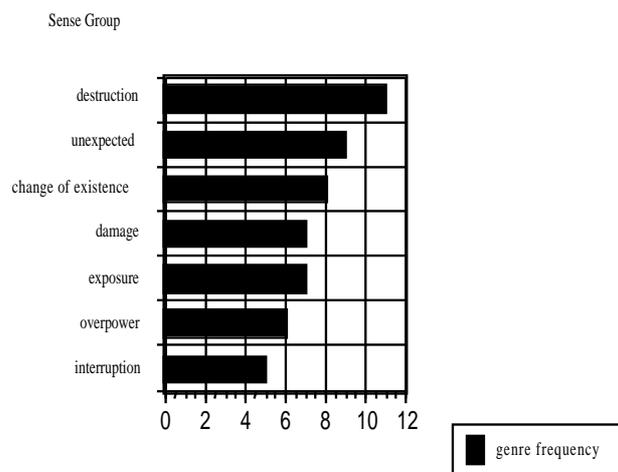Hornby, A. S. (Ed.) (1974). *Oxford advanced learners' dictionary of current English.*

Kellerman, E. (1978). Giving learners a break: Native language intuitions as a source of predictions about transferability. *Working Papers on Bilingualism 15,*, 59-92.

Kellerman, E. (1986). An eye for an eye: constraints of the L2 lexicon In E. Kellerman & M.Sharwood Smith (1986) *Crosslinguistic influence in second language acquisition.* (pp.35-48). New York, NY: Pergamon press.

Michaelis, L. A. (1996) Cross-world continuity and the polysemy of adverbial STILL. in Fauconnier, G., & Sweetser, E. (Eds.), *Spaces, Worlds, and Grammar* (pp. 179-226). University of Chicago Press,

Miller, G.(1978). Semantic relations among words. In M. Halle, J. Bresnan, & G. Miller (Eds.).*Linguistic theory and psychological reality*., Cambridge, Massachusetts: MIT Press.

Nelson, K. (1985). *Making sense: The acquisition of shared meaning.* New York: Academic Press

Pause, P. E., Botz, A. & Egg, M. (1995). Partir c'est quitter un peu: A two-level approach to polysemy. in U. Egli, P. E. Pause, C. Schwarze, A. von Stechow, & G. Wienold (Eds.) *Lexical Knowledge in the Organization of Language.* (pp. 245-282).Amsterdam, The Netherlands: John Benjamins Publishing Co.

Procter. P. (Ed.) (1978). *Longman dictionary of contemporary English.* London: Longman Group Ltd.

Rosch, E. H. (1973). On the internal structure of perceptual and semantic categories. In T. E. Moore (Ed.), *Cognitive development and the acquisition of language* (pp.111-144). New York: Academic Press

Shirai, Y. (1990). Putting PUT to use: Prototype and metaphorical extension. *Issues in Applied Linguistics. 1,* 1, 78-97.

Sinclair, J. (1995). Corpus typology: A framework for classification, in Melchers, G. & Warren, B. (Eds.).*Studies in Anglistics.* (pp. 17-33) Stockholm: Almqvist & Wiksell International,

Sociological Abstract Inc. (2001). *Linguistics and Language Behavior Abstracts (LLBA).* San Diego, CA: Sociological Abstract Inc.
    (2001)
                                                                    (pp. 161-

imiron: koa to purototaipu [Lexico-semantics of English basic verbs: Exploration into lexical core and prototype. (pp 23-41).] Tokyo, Japan: Sanyusha

Tanaka S. & Abe, H. (1985). Conditions on interlingual semantic transfer. *On TESOL '84.* 101-120.

Tanaka S. & Kawade, S. (1989). Kihon doushi no imi bunseki: Koa, parame-ta-, shiten. [Semantic analysis of basic verbs: Core, parameter, & perspective] *Ibaraki Daigaku Kyoyoubu Kiyou* [Ibaragi University College of Arts and Science Bulletin.] *21*, 385-418.

Ueda, N. (1998a) Mental lexicon and understanding word meanings from the perspective of prototype effects: A case study. *Journal of Japan-Korea Association of Applied Linguistics*. 1, 135-158.

Ueda, N. (1998b) Prototype effects in understanding word meanings. *Journal of Japan-Korea Association of Applied Linguistic*s. 2, 37-56.

Uhlenbeck, E. M. (1996) Some remarks on homonymy and polysemy in Partee, B. H., & Sgall, P.(1996) (Eds.), *Discourse and Meaning: Papers in Honor of Eva Hajicova.(pp.* 119-126.*)* Amsterdam, The Netherlands: John Benjamins Publishing Co.

Viberg, A. (1999) Polysemy and differentiation in the lexicon: Verbs of physical contact in Swedish. in J. Allwood & P. Gardenfors (Eds.), *Cognitive Semantics: Meaning and a Cognition*.(pp. 87-129). Amsterdam, The Netherlands: John Benjamins.

West, M. (Ed.). (1953). *A general service list of English words.* London: Longmans, Green and Co Ltd.

Figure 1. Genre Frequency of Sense Group

Intralinguistic (Genre) transferability and prototypicality of meaning of polysemous words: the case of BREAK.

Sense Group

Sense Group	genre frequency
destruction	11
unexpected	9
change of existence	8
damage	7
exposure	7
overpower	6
interruption	5

genre frequency

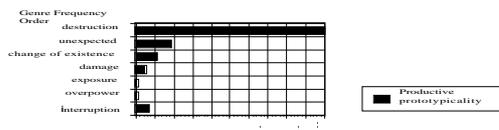Figure 2. Productive Prototypicalty in Genre Frequency Order of Sense Group

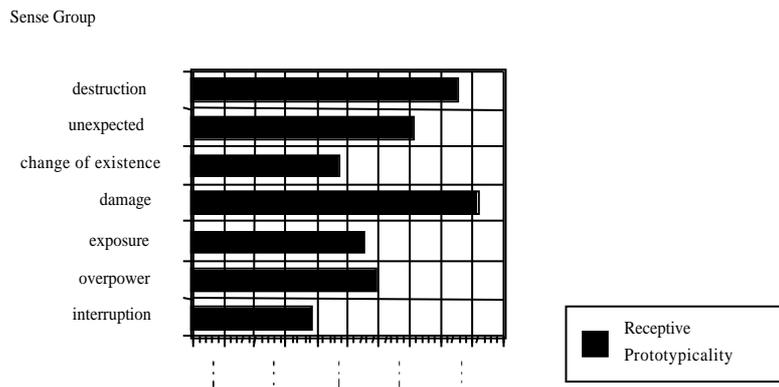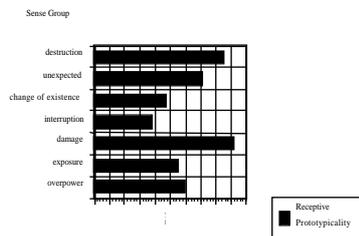Figure 3 Receptive Prototypicality in Genre Frequency Order by Sense Group

Figure 4. Receptive Prototypicality in Productive Prototypicalty Order

Intralinguistic (Genre) transferability and prototypicality of meaning of polysemous words: the case
of BREAK.

Appendix A

Table 3. Genre by Sense Group: Frequency Crosstabulation

_____

                                    Sense Group

_____

GENRE	1	2	3	4	5	6	7	8	Row Total	(%)
A	4	0	1	1	3	2	0	1	12	10.3
B	5	0	0	0	0	0	0	0	5	4.3
C	0	0	0	1	0	0	1	0	2	1.7
D	0	0	0	0	0	1	1	0	2	1.7
E	2	1	0	0	2	0	0	1	6	5.2
F	8	1	1	1	2	0	2	0	15	12.9
G	5	1	1	2	2	0	2	0	13	11.2
H	0	0	1	0	0	0	1	0	2	1.7
J	6	1	0	0	1	2	0	0	10	8.6
K	4	2	1	1	0	1	0	0	9	7.8
L	4	1	3	1	0	0	2	0	11	9.5
N	5	0	1	0	2	1	2	0	11	9.5
P	10	1	0	1	0	0	1	0	13	11.2
R	4	0	0	1	0	0	0	0	5	4.3

_____

| Column TOTAL | 57 | 8 | 9 | 9 | 12 | 7 | 12 | 2 | 116 | |
| (%) | 49.1 | 6.9 | 7.8 | 7.8 | 10.3 | 6.0 | 10.3 | 1.7 | 100 | |

Toru Kinoshita and Tomohiro Sakai

Appendix B

Topic	Number of Samples	Number of words
BROWN CORPUS	500	1136854
Informative Prose	374	846361
A. Press: Reportage	44	98898
Political	14	31149
Sports	7	15908
Society	3	6862
Spot News	9	20207
Financial	4	8945
Cultural	7	15827
B. Press: Editorial	27	60528
Institutional	10	22310
Personal	10	22627
Letters to the Editor	7	15591
C. Press: Reviews (theatre, books, music, dance)	17	39756
D. Religion	17	38591
Books	7	15766
Periodicals	6	13594
Tracts	4	9231
E. Skills and Hobbies	36	81391
Books	2	4566
Periodicals	34	76825
F. Popular Lore	48	108398
Books	23	51799
Periodicals	25	56599
G. Belles Letters, Biography, Memoirs, etc.	75	169505
Books	38	85787
Periodicals	37	83718

Intralinguistic (Genre) transferability and prototypicality of meaning of polysemous words: the case
of BREAK.

Industry Reports	2	4434
College Catalog	1	2302
Industry House Organ	1	2215
J. Learned	80	179803
Natural Sciences	12	26912
Medicine	5	11365
Mathematics	4	8963
Social and Behavioral Sciences	14	31333
Political Science, Law, Education	15	33275
Humanities	18	40323
Technology and Engineering	12	27632
Imaginative Prose	126	290493
K. General Fiction	29	66567
Novels	20	45983
Short Stories	9	20584
L. Mystery and Detective Fiction	24	55173
Novels	20	45919
Short Stories	4	9254
M. Science Fiction	6	13841
Novels	3	6946
Short Stories	3	6895
N. Adventure and Western Fiction	29	66813
Novels	15	34473
Short Stories	14	32340
P. Romance and Love Story	29	67280
Novels	14	32173
Short Stories	15	35107
R. Humor	9	20819
Novels	3	7019
Essays, etc.	6	13800

Appendix C

(Receptive prototypicality judgement sheet)

Direction: Could you evaluate the following sentences including the verb BREAK in terms of how typical, central, or core-like, the particular meaning of BREAK in each sentence is ? Please use 5 point scale with 5 for the most typical meaning and 1 for the least typical one.

		least typical		most typical				
e.g.	Give me a break.	3?		1	2	3	4	5

1.   She broke the cup.

2   .Nobody could break the witch's spell . (destroy the power of)

3.   Her radio's broken. (not function)

4.   It took a few drinks to break the ice. (ease the tension)

5.   Her fall was broken by a tree. (reduced in force)

6.   When will the weather break? (change)

7.   She broke his heart. (hurt emotionally)

8.   He broke wind. (caused wind to escape violently)

9.   They broke the enemy code. (cracked, solved)

10. The man who broke the bank at Monte Carlo. (rendered penniless)

11. They are always breaking promises. (failure to keep)

12. Who's going to break the news? (announce)

13. She broke her journey in Delhi. (interrupted)

14. The police broke the enemy resistance. (ended)

15. The waves broke on the shore. (disintegrate)

16. The Americans today broke yet another record. (bettered)

17. She broke her leg.

First language   (English or _____)

Gender   (Male   Female   )

Thank you for your corporation.