Collocation and Extended Meanings

In-Dong Song

Collocation can contribute to economic and systematic processing of significant linguistic topics, *i.e.*, metaphor, metonymy and anomaly. Generalizing the application as well as the concept of collocation, this paper shows an efficient computational theory of collocation is feasible.

An over-all picture of collocation can be captured through a view of collocation as continuum, *e.g.*, a frozenness cline. Frozenness entails two kinds of expectation: rule-based and use-based expectation. The frozenness continuum is partitioned in terms of a ‘collocational space’ and partial functions which enable even the composition of meanings of frozen expressions.

Derailment from collocational expectation results in anomaly or figurative extension of meanings. Metaphor, a paradigmatic extension of meaning, was treated basically in terms of rule-based approaches. But, dead or frozen metaphors can be efficiently processed in terms of collocational expectation, in which novel metaphors and anomalies can be distinguished better than in former approaches which were entirely dependant on rules.

Metonymy, a syntagmatic extension of meaning, was relatively more neglected than metaphor in linguistic studies. We show an example of calculating metonymy by the collaboration between default and compositional processing, or the cooperation between use-based and rule-based inferences.

1. Introduction

Collocation is the company a word keeps with other words therein and a significant linguistic fact to be described or explained systematically. By collocation, word meanings are significantly determined and understood; in

---

1 This paper reorganizes part of Song (1994) under the same framework with some revisions. The aim of reorganization, on the whole, is to focus on the contribution of our collocation theory to the computation of extended meanings. I am glad to express my deepest thanks to Professors Myong-Won Cho, Kiyong Lee, Nahm-Sheik Park, Ik-Hwan Lee and Gyonggu Shin for their comments on Song (1994), part of which this paper newly organizes.
other words, part of the meaning of a word can be dealt with according to their tendency to co-occur within texts. The psychological motivation of collocation is evidenced by various psychological experiments or observations. Evidence from lexicographers' experience to theoretical arguments supports a distinct status of collocation within a linguistic system.

An overall picture of collocation, however, has been missing, so we provide a brief one that views collocational phenomena as continuum. With that continuum, we particularly focus on the calculation of extended meanings (i.e., meanings extended from literal ones). In doing this, we base our argumentation on the meaning schema M2(U)M1, which reads simply as 'an extended meaning (Meaning2) requires the use (Use) that in turn requires a basic meaning (Meaning1).'

Meanwhile, an adequate way of calculating extended meanings is still needed, for there has been no clear way to distinguish between novel metaphors and semantic anomalies, no efficient systematic computational procedure of metaphoric expressions, or almost nothing about the computation of metonymic expressions. With a theory of collocation and the meaning schema mentioned above, we show the 'compositionality' of meaning holds even in the cases of extended meanings.

2. Collocation as Continuum

2.1. Evidence for Collocation

The study of collocation has usually been neglected partly because of the syntax-driven modern grammars and partly because of the flux of views on collocation. Even those who recognize the term 'collocation' often regard it as old-fashioned or as a term used without any theoretical motivation in language teaching. In other words, collocation has been treated merely as a kind of meaning theory once studied during the Structuralist era along with Structural semantics or Componential Analyses. This paper, however, shows collocation is a linguistically significant fact, the description of which is psychologically and linguistically well motivated. First, let's see the psychological aspects, considering children's use of language.

As for children, recognition and use of collocation such as icecream come before those of constituent parts, e.g., ice and cream. The constituent parts
Collocation and Extended Meanings

of ‘icecream’ have lost their functions as compositional parts here and ‘icecream’ now functions as a whole. Icecream is neither necessarily a cream since it can be made from a vanilla, chocolate or strawberry, nor consists of ice in contrast to ‘iceberg,’ ‘ice box,’ ‘ice cap,’ ‘ice rink,’ etc. Instead, it refers to a soft sweet cold mixture containing cream or its substitutes, though it is distinct from ‘cold cream.’ But with all this, no one refuses to use ‘icecream’ for that sense. This word has an inseparable meaning which cannot be gathered directly from its constituents.

Let’s see a more systematic example for the psychological motivation of collocation. In his/her acquisition of the first language, a child simply collocates the expressions that make sense (Matthews 1981: 187):

(1) mummy-come, big-cake, etc.

Even in acquiring syntax he/she uses a simple collocational schema. Such strings as (1) become longer by extending collocations, e.g.,

(2) a. mummy-come-yesterday
    b. mummy-want+wear-hat = mummy-want-wear-hat.

Grammatical categories of ‘construction,’ ‘sentence’ and so on are not appropriate to the schemata with which learning begins (Matthews 1981: 194).

Similarly, after the first stage through which the child passes in his/her acquisition of language is that of babbling, the second stage starting towards the end of the first year is typically that of holophrastic speech. The following stage in the course of the second year is that of a simple two-word or three-word telegraphic collocation. Such a collocation typically bears no overt grammatical markings such as present vs. past tense, singular vs. plural, and so on (Lyons 1977: 89-90).

The above observations are limited to simple adjacent collocations which are dominant in children’s language learning. Ending with naive observations, they do not mention any controlled experimental research. Experimental evidence for psychological reality of collocation—e.g., on the collocational links’ priority in children’s language learning—has been presented in some research. In word association experiments, young children

\(^2\) We treat a compound as a kind of frozen collocation.
tend to respond to ‘table’ with eat, to ‘dark’ with night, to ‘send’ with letter and to ‘deep’ with hole, while adults typically respond to such words with chair, light, receive and shallow.\(^3\)

In the mental lexicon, the ties between collocates tend to be retained even after serious brain damage or tend not to deteriorate in old age:\(^4\) when young and old subjects were presented with two simultaneous strings of letters and asked to judge whether both were words, the older group, average age 70, took longer to respond than the younger group, average 28, but if the strings were collocations (e.g., rain-wet), then the responses of both groups were equally speeded up. Experiments show that aphasics also preserve collocational links well:\(^5\) when aphasics were asked to squeeze a rubber ball if they recognized a relationship between pairs of words, they responded fastest if these words were collocates or co-ordinates, and their error rate was exactly the same as that of normal subjects.

These findings are strengthened by evidence from common tongue-slips, where people sometimes begin with one phrase and then get ‘derailed’ on to a familiar collocation such as

(3) a. Hungarian restaurant

    \(for\)

    b. Hungarian rhapsody. (Aitchison 1987: 78)

This evidences how collocational links are powerful and long-lasting in the lexicon and linguistic processing.

Such permanence in memory shows the existence and primacy collocational links have in the mental lexicon. In other words, collocation is a sure phenomenon which cannot be refuted by the fact that it cannot be captured through syntax or syntax-governed morphology. Therefore, colloc-

---


cation, with such a status in the mental lexicon, should be described or explained by an appropriate theory. Any grammar which neglects this task cannot be said to be complete or adequate since we live by collocation in our linguistic life.

The term ‘collocation’ has been employed mostly in pedagogical works and also deemed to belong to the past European Structuralist theory of language (e.g., Celce-Murcia & Larsen-Freeman 1983; Sung-Shik Cho et al. 1990: 217). So now we discuss the linguistic raison d’être of collocation — i.e., a unique status of collocation within a linguistic system.

In their practical works, lexicographers often recognize the existence and importance of collocation which cannot be attributed to any existing syntactic or semantic formulations:

(4) a. hard luck, hard facts, hard evidence
    b. *strong luck, *strong facts, strong evidence
    c. tough luck, *tough facts, *tough evidence, etc. (Sinclair 1987: xvii)

In the examples of hard luck and its near-synonymy tough luck, hard means ‘unfortunate,’ while in hard facts and hard evidence, it means ‘unlikely to be proved wrong,’ which is almost synonymous with strong in strong evidence. These patterns of collocation, however, show that the near-synonyms tough and strong go only with luck and evidence respectively. So those patterns of collocation are not determined by meaning. Furthermore, we can capture no distinctive syntactic generalization in those examples in (4): hard co-occurs with a noun, singular or plural, in the same manner as strong or tough (these can also co-occur with plural, e.g., strong men or tough guys). They are not just anomalies to a syntactic rule like ‘NP→(Det) (Adj) N’ because, given a context, the collocates are expected from each other. The reason is that they co-occur frequently in a language system.

We argue that there is a collocational level and it occupies an area interfacing some parts of grammar and we assume the methodological approach to collocational phenomena should be basically lexical, not syntactic, semantic, or what not. Halliday (1966) calls such an independent level as lexis and our assumption will not necessarily contradict his argument since we consider the relation between the collocational level and the other parts of
grammar, especially syntax, to be gradient rather than discrete.\(^6\)

There are questions which cannot be solved within any traditional level of grammar.

(5) a. make up\(_1\)
   
   She made up her face
   
   Your complexion needs a different makeup.

b. make up\(_2\)
   
   She made up her team
   
   Your committee needs a different makeup.

Grammatically, the primary distinction is the differentiation between *made up* as a verb, *makeup* as a noun, etc. The grammar (here, the syntax) cannot handle the distinction between *make up\(_1\)* and *make up\(_2\)*.

On the syntagmatic or collocational axis of lexis, there is a lexical item which does not have any defined syntactic status and is neither identified as a morpheme, a word nor a word group: *e.g.*, the single but discontinuous sequence such as *let···in···for* in

(6) He let me in the other day for a lot of extra work.

Thus, the lexical component is not secondary to syntax but stands as a distinct level. It is not a relation between the categories in a grammatical structure—*i.e.*, it is neutral to category or constituency; thus, collocation cannot be described or explained on the syntactic level. Morphological analyses, mainly dealing with the word formation, cannot see the whole picture of the collocational phenomena, either. Collocation occurs within or beyond a sentence.

The level of lexis describes the tendencies of lexical items to collocate with each other. Such tendencies cannot be expressed, as we see in (7), in terms of small closed sets of grammatical categories. The first two words in (7a) occur in the same paragraph in Roget's *Thesaurus*, sharing some semantic similarity, but *cruelty* would be odd among them (Sinclair 1966: 410

(7) a. tome, paperback, cruelty

While showing no special tendency to co-occur, the words, tome and paperback, share collocations with other words like those in (7b) which bear no direct relation to cruelty (ibid.).

This kind of lexical patterning forms a cline of 'frozenness' according to the strength of the pattern. This strength can serve as a parameter for a collocational continuum which will be discussed soon. The study on such strength can be done in terms of frequency or probabilistic corpus analyses. Likewise, Halliday (1991), viewing language as inherently probabilistic, stresses the need to investigate frequencies in texts to establish probabilities in the grammatical system. Along this line, we can posit that the collocational study, though inherently probabilistic, can constitute a part of linguistic modeling. For a complete grammar, the probabilistic analysis of collocation should be included along with the intuitive aspects.

Further arguments will be possible for the motivation of a collocation theory, including possible distinctions among the frequently confused concepts, i.e., collocation, selection(al) restrictions, co-occurrence restrictions and sortal incorrectness. Selectional restrictions which work with syntactic rules have been thought to be syntactic within the Transformational Grammar since Chomsky (1965). Since syntactic rules are stated in terms of categorial or phrasal terms, selectional restrictions are an inherently category-bound relation: e.g., a verb 'sleep' requires an 'animate' subject noun. From a purely semantic point of view, selection(al) rules are from the semantic hierarchy. From the lexical point of view, selectional restrictions concern the paradigmatic choices of words.

Collocation is different from selectional restrictions in that it operates even below the word level like (8), beyond syntactic categories, and often beyond clausal or sentential boundaries.

(8) The horsewoman clapped her spurs to the horse's flanks.

\footnote{Chomsky (1965: 95-96) views 'selectional rules' as expressing selectional restrictions or restrictions of co-occurrence.}
Grammatically, in contrast to the category- or constituency-related selectional rules, collocation permeates into any category and bonds the words as ‘friends.’ Different from the phrase- or clause-bounded selectional restrictions, collocation is fundamentally textual and contextual. Being corpus-based and coming into existence by probabilistic knowledge, collocation can go even beyond text to become cross-textual or cross-registral. Semantically, collocation is a syntagmatic extension of lexical meaning, while selectional restrictions concerns part of the paradigmatic extension of lexical meaning. In contrast to the deductivity of selectional restrictions, collocation operates inductively in the course of processing, since collocation is based on probabilistic tendency and realized by syntagmatic accompaniment. Our point here is that collocation is distinct from selectional rules and sortal incorrectness which is to be discussed below.

Next, the notion of sortal correctness stemmed from the tradition of logic and has been used to explain normality or anomaly of a proposition by referring to the sortal range of a predicate. Linguistically, it can be seen as a semantic device which can explain an anomaly. Thus, an anomalous sentence *The table is laughing* can be explained as sortally incorrect because ‘the table’ falls outside the sortal range of ‘laugh.’ The ill-formedness of this sentence can also be predicted by selectional restrictions operating within a grammar in such a way as the selectional feature of ‘laugh’ are in conflict with the feature of ‘table.’ Whereas selectional restrictions treat the allowable combination between two words, especially between a predicate and its argument, and a head and its modifier, sortal incorrectness usually deals with the semantic relation between a predicate and its argument.

But collocation can be related to any sequence of words. It is not a yes/no question like sortal incorrectness or selectional restrictions, but probabilistic since collocatability means the co-occurrence tendency of words. In short, collocation may cooperate with selectional or semantic restrictions as will be exemplified later and those three kinds of relations or, at least, collocation and either of the other two relations, *mutatis mutandis*, can be cooperative even to account for extended meanings.

To sum up, collocation is not an old-fashioned and once-studied obsolete notion, but a prevalent linguistic fact to be described or explained systematically. Its psychological motivation is evidenced by various experiments or observations. Its strength and prominence in the mental lexicon
and children's language acquisition provides one of strong incentives for a theoretical study of collocation. In addition, linguistically, there are ample examples to support the existence of the collocational level.

2.2. Contextual Theory of Meaning

A collocation theory relates to a contextual theory of meaning. Collocation is a linguistic term which began to be developed systematically by J. R. Firth, the founder of the London School of Linguistics. It is akin to the use theory of meaning in philosophy introduced by Wittgenstein (1953 [1945]: 20), who underlined the use in determining the meaning of an expression:

(9) For a large class of cases—though not for all—in which we employ the word “meaning” it can be defined thus: the meaning of a word is its use in the language.

However, the use theory is said to be subject to circularity (Ik-Hwan Lee 1985: 52, Findlay 1974 [1961]): i.e., the concept of ‘use’ presupposes the meaning of the expression concerned. This will be critical if the use theory insists that the meaning of a word can be determined only by its use. Meanwhile, the ‘meaning by use’ can be thought to be based on the ‘use by meaning’. The apparent circularity of this relation can be avoided if the ‘based-on’ relationship is indexed or counted like (10a), which can be restated schematically like (10b):

(10) a. meaning2-by-use1 is based on use1-by-meaning1
    b. meaning2 by use by meaning1

8 Ik-Hwan Lee (1985: 31-45) classifies the theories of meaning roughly into seven types: referential, synthetic, truth-conditional, mentalistic, behavioristic, and pragmatic (or use-is-meaning) semantics. Since the referential semantics, formalizable in the form of truth-conditional semantics, is subsumed under the synthetic semantics, and mentalistic and behavioristic semantics can be refuted (cf. Ik-Hwan Lee 1985), now there remains only one type of theory, the pragmatic semantics. We call the reference-based synthetic theory simply the ‘referential theory,’ and the pragmatic or use-is-meaning theory simply the ‘contextual theory’ or ‘use theory’ which seems to be fit to comprise the meaning theory of the European, especially London, school of linguistics.

9 The merits or demerits of the contextual theory are pointed out by Ik-Hwan Lee (1985), and Song (1994) suggests a possible recuperation of the handicaps of this theory.
c. M2(U)M1

d. EXTENSION(Meaning1, Meaning2, Use1):-Use1(Meaning1, Meaning2).10

The process described by (10b) can be abbreviated like (10c) and also represented in a Prolog notation roughly like (10d) as the process of extension—paradigmatic or syntagmatic—from a basic literal meaning. The notation of (10d) implies the extension is a general or default phenomenon derived from ‘meaning1’ of every lexeme. The synthetic constraint of (10c) predicts such common semantic phenomena as metaphor, i.e., the paradigmatic extension of basic meaning, and collocation, i.e., the syntagmatic extension.

Now the vicious circularity can be avoided which is considered to be inherent in the theory of ‘meaning by use’, ‘meaning by collocation’, or the contextual theory of meaning. The reason is that the defining meaning1 and the defined meaning2 are on the typological hierarchy through the medium of language use or the function of use which maps meaning1 into meaning2 and the extension process is distinctively bounded owing to implicit intra- or extra-factors of human cognition relating to the use and processing of language.11

Besides, the schema ‘M2(U)M1’ desirably integrates the referential semantic tradition and the contextual theory of meaning. It also explicitly excludes any overgeneralization about Firth’s (1957) ‘meaning by collocation’, since collocation, for Firth, refers to the syntagmatic relations into

10Song (1994) employs the term ‘ELEVATION’ for ‘EXTENSION’. But this paper uses the latter as a general term comprising both the paradigmatic and syntagmatic shift of lexical meaning, and the former merely as a special metaphor for the paradigmatic extension. Meanwhile, the syntagmatic shift, ‘EXPANSION’, describes the metonymic or collocational addition of semantic quality.

11We think a habitual or conventional constraint operates on the possibly infinite recursion of the typological steps: informally,

any recursion upward meaningN (N>2) in the meaning type cannot occur

because any habitual occurrence of meaning2 will be reduced down to the meaning1 and enter into the dictionary of a language. Thus, what is complementary to this extension by use is the reduction by a ‘freezing’ force in the synchronic use as well as diachronic change of a language. This argument is indirectly supported by the mathematician’s type-theoretical resolution of paradoxes: no meta-language higher than ‘meta-meta-meta’ typological level is needed and the infinite raising of type is blocked.
which lexical items habitually enter and is a part of the meaning of the lexical items concerned: *e.g.*, (11) 'one of the meanings of *night* is its collocability with *dark*, and of *dark*, of course, collocation with *night*'.

So the theory of meaning by collocation does not clashes with existing traditional referential theory of meaning, but presupposes it as the M2(U)M1 schema expresses. Though some literature on semantics sharply divides the use theory from the referential theory of meaning, this schema captures a more integrated view of them. This idea will be represented in a frame to calculate the extended meanings later.

### 2.3. Degree of Frozenness

In terms of M2(U)M1, the basic meaning of every lexeme is extensible and the mode of extension is paradigmatic (*e.g.*, metaphor) or syntagmatic (*e.g.*, collocation, metonymy, *etc*.). These two axes of extension crisscrosses each other, forming the continuum of frozenness. This continuum stretches itself from the most freely combining expressions to the most frozen ones. The mental lexicon shows a wide range of a collocational continuum, *i.e.*, the frozenness cline of syntagmatic extension (*Aitchison 1987: 78*): at the one end of the spectrum words associate with other ones more productively (*e.g.*, *rude adolescents*) than the words in the habitual connections or clichés (*e.g.*, *bright and early*) which overlaps with idioms or frozen expressions (*e.g.*, *fall into place*).

The collocational continuum reflects the gradation of expectation: words in the customary co-occurrence combinations can be said to predict the occurrence of the other(s). We call this kind of prediction 'use-based prediction' in contrast to 'rule-based prediction' which exploits rule-based inference like selectional restrictions.

*Kjellmer (1991)* lists the types of phrases which are differentiated according to their predictiveness:

(12) a. 'fossilized phrases' which are bidirectionally predictive:

*Anno Domini, be-all and end-all, etc.*

b. 'fossilized phrases' which are unidirectionally predictive:

*Morse code, ball-point pen, open sesame, from afar, etc.*
c. 'semi-fossilized phrases' in which a word predicts several words very limitedly: billy can, billy goat; by and by, by and large; go/grant/jump/stand bail, etc.
d. 'variable type phrases' which are more frequent than any type above and in which a word predicts the other(s) more loosely in that the phrase is less frozen: glass of water, classical music, feel comfortable, a number of, etc.

The three types, i.e., fossilized, semi-fossilized and variable ones, show the gradience of prediction or cohesive power. In other words, the three types can be said to be different in their frozenness.

Kjellmer's taxonomy of set phrases in terms of the 'scale of idiomaticity', however, is not sufficient to capture the whole picture of the collocational phenomena. His classification is limited to phrases, especially to idiom-like ones, while our theory of collocation covers the expressions from sentence-like framework to encapsulation and compounding, or from free combination to frozen expressions. Thus, Kjellmer's taxonomy cannot capture and explain the novel metaphors, for instance.

To show the degree of frozenness, Chungmin Lee (1992) classifies the semantic representations of frozen expressions into two types: the lexically filled as in (13a) and the lexically open as in (13b):

(13) a. kick the bucket; spill the beans
    b. let alone; the -er... the -er.

(13a) allows only one argument, i.e., an external one, and shows a higher degree of frozenness. Chungmin Lee's (1992) analysis shows that even idiomatically frozen expressions can be different in their degrees of frozenness. More generally, we think every linguistic expression has some degree of frozenness between the two ends, i.e., nil and full degrees, of the same continuum. This knowledge is formally represented in Section 2.4.

By the way, Fraser (1970) defines the frozenness hierarchy like (14) in terms of five kinds of operations, i.e., Adjunction (A), Insertion (I), Permutation (P), Extraction (E) and Reconstruction (R), which are not quite the same as their syntactic transformations.
(14)\(^12\) L6: no restriction on the application of operations  
L5: A,I,P,E,R apply  
L4: A,I,P,E apply  
L3: A,I,P apply  
L2: A,I apply  
L1: A apply  
L0: no operation applies  

Between each two contiguous levels in (14), the part-of relationship holds. Fraser puts 'kick the bucket' into the last but one level, i.e., 'quite frozen' in our terms, because it only allows Adjunction. Thus, (15a) is allowed:

(15) a. your friend's kicking the bucket  
b. your friend's kicking the large bucket  
c. your friend's kicking the bucket down.

However, while (15b) is the case of 'Insertion' (in Fraser's terms) and unacceptable as the sense of 'die,' (15c) has an Adjuncted (in Fraser's term) item 'down' but the reading is not idiomatic. In sum, the explanation of frozenness solely based on the rules is likely to be mistaken since it always seeks for the discrete cases. Therefore, Fraser's syntactic definition of frozenness should be more constrained or the notion of frozenness should be defined in some other way.

Unlike Fraser's (1970) definition here, we think of frozenness as a composite notion that requires not merely syntactic information but semantic and contextual, i.e., pragmatic factors. The emphasis here is that the syntactic studies of idioms are not a proper explanation of idioms and the syntactic or formal factor is just one aspect of the frozenness continuum which results from the semantic or pragmatic strategies of the speakers of a language.

2.4. Partition of Collocation

Now we introduce a more comprehensive and delicate view of collocation as continuum between productive vs. frozen expressions. The collocational continuum comprises a variety of extremes forming gradient continua that

crisscross the frozenness cline: *e.g.*, the cline between adjacency vs. discontinuity (or between encapsulated single-word sequences vs. multi-word ones), between the most frequent vs. the rare or null occurrences, between literal vs. extended meanings, *etc.* In this paper, however, we mainly consider the literalness cline between literal vs. novel metaphors.

The knowledge of such gradient parameters crisscrossing the collocational cline of frozenness can be delicately represented by a rather formal notation. Among those parameters, literalness interwoven with frozenness is represented in the form of a partition graph. This representation ranges from the least frozen area, novel metaphors, literal expressions, variable expressions, set phrases, and lexicalized (frozen) expressions, consecutively forming the gradience of frozenness and reflecting the fact that the meanings of the lexicalized expressions can be literal or extended.

According the meaning schema $M_2(U)M_1$, every lexeme has the potentiality for the meaning extension ($M_2$). Technically, a frame model for the lexical meaning ought to admit the expansibility for a collocation list and the calculability for the paradigmatic meaning shift (or elevation): *i.e.*, the syntagmatic shift, *e.g.*, the collocational meaning, is retrieved from the expanded space of a semantic frame, *i.e.*, 'CSPACE', and the paradigmatic extension across semantic fields (or frame fields) is calculated from the basic literal information (or the original frame information). Before the calculation of an extended meaning, let's see a frame representation of a collocational meaning.

The CSPACE can be represented in a frame form. For example, (17) is a brief frame for the collocational triplet 'a piece of' in (16) which ranges in its meaning from the literal like (16a) to the figurative like (16b):

(16) a. a piece of bread  
   b. a piece of information  

(17)$^{13}$ a. piece1:

\[
\begin{array}{l}
\text{(piece1} \text{[CFRAMEWORK}[G1\text{ a1}][NUM\ a1]]) \\
\text{[N1 piece1] [UNIT piece1]}
\end{array}
\]

$^{13}$In this frame, CFRAMEWORK represents a collocational framework and includes the grammatical subcategorization relating to the word concerned. The SEM slot contains the meaning resulting from the mapping by the partial function $F_1$: *i.e.*, $F_1(a,\text{piece1},o_f)=$'One_Piece' (or 'a-piece-of'), which expects N2 and thereof $F_1(a,\text{piece1},o_f)(N2)=$ 'One_Piece(bread/cake/meat/toast)' is deduced.
These examples show that the single lexeme 'piece' can have a range of meaning from the literal to the figurative, and its CSPACE specifies its meaning by use, i.e., [CSPACE [SEM list]] together with its frequent collocates, i.e., [CSPACE [N2 NONSPLIT list]]. The word 'piece' can have an almost literal sense as in (17a); and a figuratively extended meaning as in (17b). In the case of 'piece2,' it is different from 'piece1' in that its meaning cannot be computed compositionally just within its basic sense frame and thus has a collocational space, i.e., CSPACE, which specifies the collocational list and the meaning of 'piece2' thereof.

The preliminary of a frame like (17) may seem to have redundancy in that the N2 slots have attributes like 'EDIBLE-SPLIT' and 'NONSPLIT' in its CSPACEs and only with those features one seems to be able to specify the collocation list. Those features, however, are never exhaustive in specifying their values or instances since we can easily identify the cases which are the members of the attributes but do not belong to the collocate class of the lexeme concerned. That is, 'apple' goes against the frame information in (17a) and 'air' can go against the information to be described in (17b).14

The frame can be more elaborated to comprise more fine-grained cells representing the form, grammatical information, 'genus' information, 'differentia' information and a collocation list of a lexeme. We will use some of these to calculate metaphors.

14Similar problems are met in an analysis wholly depending on selectional restrictions (Allen 1987: 256–257).
The information in a frame can be reinforced by stochastic arrangements. One can get the best or typical translations of source language compounds by simply searching a target language corpus and selecting the construction which is most frequent.\textsuperscript{15} If a system has a frame which carries default collocations based on statistics, the system need not search the corpus everytime it meets candidates for collocation. Moreover, the frame is continually renewed through a corpus statistics about its CSP information. The collocate list arranged according to frequency will be of use to reduce time especially in searching a fairly long list.

Now, with the EXPECTED\_LIST of CSPACE and some OPERATORs, collocations can be partitioned; the results are rare collocations, variable collocations, set collocations and lexicalized collocations. The operators working on a collocational continuum are as follows:\textsuperscript{16}

\begin{equation}
\text{(18) operators on collocational continuum}
\begin{align*}
\text{NOT FROZEN:} & \quad F0 \Rightarrow \text{zero or rare collocations} \\
& \quad \text{(e.g., novel metaphors)} \\
\text{SLIGHTLY FROZEN:} & \quad F1 \Rightarrow \text{variable collocations} \\
& \quad \text{(e.g., 'a piece of bread')} \\
\text{QUITE FROZEN:} & \quad F2 \Rightarrow \text{set collocations} \\
& \quad \text{(e.g., dead metaphors, idioms, phrasal verbs)} \\
\text{MOST FROZEN:} & \quad F3 \Rightarrow \text{lexicalized collocations} \\
& \quad \text{(e.g., compounds, encapsulation)}
\end{align*}
\end{equation}

\textsuperscript{15}For instance, since the German compound 'Oppositionsgruppe' may be translated to either 'opposition group,' 'group of opposition,' 'oppositional group' or 'opposition’s group,' Rackow et al. (1992) choose the most frequent NN option, i.e. 'opposition group,' which occurs 89 times in the corpus, by consulting a corpus of 40 million words of The Washington Post articles.

\textsuperscript{16}Upon the gradient continuum down to the frozen area, the Principle of Least Effort works. We also think Nahm–Sheik Park’s (1992b) 'Law of Conservation' operates on this continuum as a principle compensating for the loss of information which can appear during the freezing or reducing process through the Principle of Least Effort. That is, a speaker tend to reduce effort by his/her compact use of an expression to the effect of the domain reduction in information processing. As for the hearer, he/she tends to demand maximal perceptual information for the ease of interpretation to the effect of process reduction in information processing, through the Principle of Least Effort. The Law of Conservation balances these two tendencies and provides compromise between speaker and hearer in information processing.
The operators, F0 through F3, are partial functions like the purely semantic functions in Gazdar et al. [hereafter GKPS] (1985: 238). GKPS's approach is purely semantic and goes against such purely syntactic views of idiomaticity as Fraser (1970). The basic assumption of the GKPS approach is that the parts of idioms should be assigned interpretations, unlike the Generative Grammar tradition in which idioms are treated as elements which are possibly assigned internal syntactic structure without any internal semantic structure.17

Partial functions are the relations satisfying the uniqueness condition for a function but perhaps fail the domain condition.18 With this function, the normal principles of compositionality is true of idiomatic interpretation for ‘spill the beans’:19 (‘E’ is ‘elevation’, ‘F’ ‘partial function’)

\[
\begin{align*}
(19) \ a. \ E(\text{spill'}) &= \$ \text{spill} ['\text{divulge'}]; \ E(\text{the'}) &= \text{the'}; \ E(\text{beans'}) &= \\
&= \$ \text{beans} ['\text{information'}].\20
\end{align*}
\]

17 An expression like *keep tabs on*, for example, is partially analyzable despite its lack of full compositionality. As for *tabs*, in particular, we can construe it as something like ‘surveillance’ or ‘contact’ (Langacker 1987: 477; GKPS 1985: 239).

18 A relation \( R \) from \( A \) to \( B \) is a function if and only if it meets both of the following conditions: (1) Uniqueness condition – each element in the domain is paired with just one element in the range; (2) Domain condition – the domain of \( R \) is equal to \( A \). When the function mapping an ordered pair of real numbers \(<a, b>\) into the quotient of ‘a’ divided by ‘b’ is not defined when ‘b=0,’ but this function is single-valued and thus meets the uniqueness condition, it is a partial function. Thus, a partial function is a total function on a subset of the domain. (Partee, ter Meulen & Wall 1990: 30–32)

19 Even frozen idiomatic elements like *strings* in the following can be emphasized through topicalization, coherently collocated with quantifiers or modifiers, and go through VP-ellipsis: Those strings, he wouldn’t pull for you (Topicalization); Pat got the job by pulling strings that weren’t available to anyone else (Modified); pull a string or two (Quantified). This fact indicates that some interpretation is being assigned to those idiom chunks which are lexically filled (GKPS 1985: 237–8); in other words, the meaning of those frozen expressions can be compositional.

20 The exact argument of ‘E’ will be the literal meaning. Hence, here the literal meaning “spill” serves as an input for ‘E’. “$ spill” is an elevated or extended interpretation of the word ‘spill’ in a context. ‘E(\text{the'})’ returns a default literal value without any different extension. This fact can be partly explained by an observation that closed, *i.e.*, grammatical, words tend to resist being elevated; hence, we can call them ‘closed’ against elevation.
b. $F(\text{the}', \text{beans}') = \text{the}'-\text{beans}'$, $F(\text{the}', $ \text{beans}) = $ \text{the}-\text{beans}$

['information']; $F(\text{spill}', \text{the}'-\text{beans}') = \text{spill}'(\text{the}'-\text{beans}')$,

$F(\$ \text{spill}, $ \text{the}-\text{beans}) = \$ \text{spill}(\$ \text{the}-\text{beans})$.\(^{21}\)

However, the operators in (19) do not coincide with GKPS's (1985) purely semantic partial function in that they comprise not just semantic but formal and contextual aspects as well. In other words, collocational frozenness in our model captures the fixed property of form, meaning and context together in contrast to GKPS's focusing solely on semantic aspects of idiomatization. The basic assumption on collocational gradience of frozenness is thus multi-variational. Especially, we underline the textual or contextual impact on a collocation being condensed or frozen to get a solid status down on the collocational cline.

To represent technically the gradience of frozenness in (18) which reflects the degree of fixedness of the form, meaning and context of a linguistic sequence, an expression of the (slightly or not) frozen sequence has a CSPACE partially filled with the use meaning and the prediction for the expression expected in collocation. We note the family resemblance, i.e., the relation sharing CSPACE, between the varied crisscrossing groups belonging to a collocational continuum.

Respecting the idea argued for so far, we can express the knowledge in (18) on a graph or network which is pipe-lined into the use vs. rule contrast along the continuum as in (20),\(^{22}\) which will in turn contribute to embodying instances abounding in a collocational continuum, e.g., as in (21). In the diagram of (21), we focused on the information of SEQ1 node to represent the compound's information among the collocation sequences. Lack-

\(^{21}\)In (19b) two kinds of readings for 'spill the beans' are returned by 'F': literal and figurative (i.e., idiomatic). Our formulation captures this fact more explicitly than GKPS (1985: 238-89), who do not show the difference between the two kinds of meaning of 'the' in their example (119) on p. 239 therein.

\(^{22}\)A principle of CSPACE information composition can be set up for a consistent combination of CSPACE information: e.g., The sframe [...] [CSPACE] [...] value of a given phrase is the unification of the sframe [...] [CSPACE] [...] values of its daughters.

This principle requires that all of the CSPACE information associated with any part of an utterance will be inherited as part of the set of CSPACEs associated with the utterance as a whole. This will make certain of the compositionality of CSPACE calculation.
ing this kind of comprehensive view on collocation, traditional studies on compounds has treated them separately from collocational phenomena, so missed the generalization cutting across the syntactic classes.

(20) Major groups of collocational sequences

![Diagram of collocational sequences]

- **Operator**
- **CSPACE**
- **SEM**
- **Expected_List***

- **F0, F1, F2, F3**
- **[ ] [ ... ]**

- **{novel {literal expressions} {variable expressions} {set phrases} expressions}**
- **{lexicalized expressions}**

- **{zero or rare {variable expressions} {set collocations} {lexicalized collocations}}**

- **(FROZENNESS)**

- **(collocational continuum)**

- (rule-based interpretation by categorial expectation: zero lexical expectation)
- (use-based interpretation by lexical expectation through F1, F2, F3)

* default collocates
  
  [ ]: null list; [ ... ]: non-null list
(21) Expectation pattern of collocational sequences

SEQ1
(= lexicalized)
- single element
- double element
  (shortening)
- recoverable
  - recoverable
  - clipped acronym
  - sound change & sound spelling
    - change
      - 'gospel'
      - <good message>
        - 'lady'
        - <loaf-kneader>
          - 'lord'
          - <loaf-ward>
            - 'forecastle'
            - [fouksl]
            - 'blackguard'
            - [blaegə:rd]
            - 'topsail'
            - [topsl]
            - 'NATO'
            - 'UN'
            - ...
            - 'canary(bird)'
            - 'return(ticket)'
            - 'oil(painting)'

SEQ2
(= pair)
- adjacent element
- discontinuous element
  (in-phrase out-phrase)
- 'buy cheap'
- 'heavy smoker'
- ...
- attached
  - detached
    - (lexical form)
      - (phrasal form)

SEQ3
(= multi-word sequence)
- adjacent
- discontinuous
  (template)
  (template)
- 'kick the bucket'
- 'a piece of'
- 'the -er-
soon'
- 'an agony of'
- 'leave-
well alone'
- 'black comedy'
- 'white'
- 'get -’s foot
well alone'
- 'once in
the door'
- 'make -’s
day'
- 'in the door'
- 'make -’s
day'
- 'in the door'
- 'make -’s
day'
- 'in the door'
- 'make -’s
day'
- 'in the door'
- 'make -’s
day'
- 'in the door'
- 'make -’s
day'
- 'in the door'
- 'make -’s
day'
- 'in the door'
- 'make -’s
day'
- 'in the door'
- 'make -’s
day'
- 'in the door'
- 'make -’s
day'
- 'in the door'
- 'make -’s
day'
- 'in the door'
- 'make -’s
day'
- 'in the door'
- 'make -’s
day'
- 'in the door'
- 'make -’s
day'
- 'in the door'
- 'make -’s
day'
- 'in the door'
- 'make -’s
day'
- 'in the door'
- 'make -’s
day'
- 'in the door'
- 'make -’s
day'
- 'in the door'
- 'make -’s
day'
- 'in the door'
- 'make -’s
day'
- 'in the door'
- 'make -’s
day'
- 'in the door'
- 'make -’s
day'
- 'in the door'
- 'make -’s
day'
- 'in the door'
- 'make -’s
day'
- 'in the door'
- 'make -’s
day'
- 'in the door'
- 'make -’s
day'
- 'in the door'
- 'make -’s
day'
- 'in the door'
- 'make -’s
day'
- 'in the door'
- 'make -’s
day'
- 'in the door'
- 'make -’s
day'
- 'in the door'
- 'make -’s
day'
- 'in the door'
- 'make -’s
day'
- 'in the door'
- 'make -’s
day'
- 'in the door'
- 'make -’s
day'
- 'in the door'
- 'make -’s
day'
- 'in the door'
- 'make -’s
day'
- 'in the door'
- 'make -’s
day'
- 'in the door'
- 'make -’s
day'
- 'in the door'
- 'make -’s
day'
- 'in the door'
- 'make -’s
day'
- 'in the door'
- 'make -’s
day'
- 'in the door'
- 'make -’s
day'
- 'in the door'
- 'make -’s
day'
- 'in the door'
- 'make -’s
day'
- 'in the door'
- 'make -’s
day'
- 'in the door'
- 'make -’s
day'
- 'in the door'
- 'make -’s
day'
- 'in the door'
- 'make -’s
day'
- 'in the door'
- 'make -’s
day'
- 'in the door'
3. Calculation of Extended Meanings

3.1. Manipulated Collocation and Semantic Deviance

A decisive characteristic of collocation is the predictable nature of their collocates: the presence of one of them will predict the presence of the other(s). The continuum of collocations ranges from well-established phrases to weakly cohesive sequences. In this spectrum, at the use-driven end of total predictability there is little or no room for maneuver by the language user. At the other rule-driven end of a variable predictability a creative or productive part comes into play, conditioned by rules such as selectional restrictions.

Such a property of collocational continuum implies the degree of appropriateness, which is related to the manipulation of a rather stable collocation. Collocational expressions show the gradience from literal usage to figurative extension. Most words are related to a wide range of objects and situations and their meanings are figuratively extended to any appreciable degree of appropriateness.

Speakers of a language very largely exploit prefabricated set expressions that make it possible to move quickly through a discourse and the listeners expect the speakers to behave likewise. It is a kind of use-based expectation occurring on the collocational continuum. Any manipulation of such expectation leads to some unexpected effect like surprise (Aitchison 1987: 9).

However, deviations from a collocational norm can be used intentionally to a favorable purpose. Such intentional manipulations are mostly used for humorous effects and seem to be easily found in novels, mass media, ads, etc. (Kjellmer 1992: 123):

(22) a. Run for your Wife
   (an advertisement for a West End farce in the mid-1980s)
   a'. Run for your life or run for one’s life

(23) a. A ticket every day is money down the tube
   (A slogan for 1985 season tickets in the London Underground)
   a'. money down the drain
(24) a. comprehensive and systematic compilation of “word of mouse”
(Macintosh information in an advertisement in MacUser; 1986/9:7)

a’ [by] word of mouth.

In each pair of (22) through (24), the reader’s familiarity with the normal collocations, like the a-primes, keeps the collocation stable and predictable, which means the solid position of collocations for an English-speaker’s fluency or easy handling of his/her own language. The effect of deviation from expected collocation can be startling or comical. Such manipulative substitution of the unexpected for the expected is the very art of comicalities.

The deviations from the rule-governed area as well as from the relatively habitual lexis-governed zone imply that there is a gradient continuum between them. Of course, there are in-between examples of collocational deviations that are from the collaboration of collocational manipulation and the violation of selectional restrictions. Unlike the intentional manipulation of collocations, those deviations are not supported by the situations concerned.

If any use of a word cannot be accepted as neither literal nor figurative, it leads to an incurable anomaly since the meaning of every word is either literal or figurative. This occurs all throughout the gradient slope of the frozenness continuum: i.e., literalness forms another continuum crisscrossing the frozenness slope. Traditionally, such an incurable deviance has usually been studied under various labels like selectional restrictions, co-occurrence restrictions, sortal incorrectness, semantic restrictions, etc., which are all considered to be rule-based operations. But since the collocational continuum is the result of co-operation or division of labor between rule-based and use-based inferences, the processing of such deviance should be done under the collaboration of collocational and rule-based inferences.

In the least frozen area where rule-based prediction is practiced, the calculation of appropriateness or interpretability is done by rule-based inference. Any sentence suspected to be deviant would be calculated as deviant if no appropriate context should be provided to compensate for its violation of selectional restrictions. In the more frozen area, appropriateness computation of literal or figurative expressions is rather carried out by use-based inference, i.e., by using the CSPACE list. This calculation is a question of
degree, not a yes/no problem.

Contrary to this approach, traditional treatments of anomalies or deviant sentences filter out the sentences blocked by the rules like selectional restrictions without consideration of partial information. Our daily life, however, sees a lot of expressions that are deviant or nearly deviant through collocational manipulation or the violation of semantic restrictions: e.g., ‘? electric roll’ (cf. ‘electoral roll’), ‘*the table saw Tom,’ ‘the fifth day saw them on the top.’ Thus, meaning composition from partial information even in nearly deviant expressions is needed.

Now, let’s see how the meaning composition works on semantic anomalies. To calculate and judge semantic deviance from partial local meanings to their whole global meaning status (i.e., undefined owing to the breakdown of composition), the system accesses the default processing procedure first and then the compositional processing. The content of the CSPACE illustrated before represents the default value of lexical use; thus, if the computation of meaning by using this default information fails, the processing goes over to resort to the compositional rules. For the calculation of extended meanings, we introduce Collative or Frame Semantics in which the frames consist of sframe and cframe. The cframe which carries collocational information is similar to the CSPACE list and sframe represents semantic information originating from lexical fields (see (38)).

Then, such a processing procedure can be summarized as follows:

(25) a. default processing: collocation- or use-driven processing which accesses CSPACE of a frame in a frame dictionary (or in a cframe dictionary) and then simply combines the meanings of the parts concerned.

b. compositional processing: rule-driven processing which accesses the (first) SEM of a frame in an sframe dictionary, whose path in a frame already has come up in memory in accessing CSPACE: i.e., such processing accesses sframe in default of any default processing and the whole meaning related is computed compositionally.

In the following examples of (26) which are different in their violations, (26a) contains extremely rare collocations in contrast to the strong collocations (e.g., blue–music, music–moving) in (26b). So (26a) goes through rule-
based processing, while (26b) is accepted as a collocationally good expression with little access to the rule-driven processing.

(26) a. Colorless green ideas sleep furiously
    b. The blue music is moving.

Anyway, in terms of selectional restrictions, there are at least five direct violations in the common sense interpretation of (26a), whereas (26b) violates only once (or possibly two): (‘#’ represents a pair which suffers from selectional clash which is called here simply as ‘contradiction’ in a limited sense.)

(27) a. # (colorless, green), # (colorless, ideas), # (green, ideas),
    # (ideas, sleep), # (sleep, furiously)
    b. # (blue, music); and possibly # (music, moving).

Any parser should calculate the degree of these violations or appropriateness. But theoretically speaking, even such calculation is not sufficient because (26b) may receive a figurative interpretation only if an adequate context provided. Thus, a parser, according to our theory, seeks for clues to introduce a proper context for an appropriate interpretation. Alternatively, after it scores the violation degree(s) in an edge in the chart related to the given suspected ‘deviant’ expression, it finds extensible meanings of each words in the string to avoid featural clashes and matches them to calculate a new violation committed even with an extension. In such a way, (26b) may have an metaphorical meaning if ‘blue’ is elevated to ‘sad.’

To return the best interpretation, a good parser, human or machine, will not stop here till its parsing about extension is complete. For the case like (26a), each extension is checked to avoid immediate clashes, for instance, like (28): (‘E’ means an elevation, ‘ε’ an expansion)\(^\text{23}\)

\(^{23}\)Though we listed every possible extension, they can be used by a parser conjunctively or procedurally. For each M2 meaning, the parser check the collocation list while checking the featural matches. Thus our parser’s operation will be use-based as well as rule-based. If an intensifier’s elevation is of no need, elevation of ‘furiously’ would be vacuous and such a fact can be stated as a redundancy rule. If intensifiers are not elevated from their original meanings in the text, we need not include them in our calculation of metaphorical extension by analyzing the patterns of metaphorical elevations in a large corpus.
In (28), there are two kinds of extensions of lexical meaning, which are in complementary opposition and each can be determined by the other. They are metaphoric and metonymic extensions. Thus, the figurative or rhetorical extensions of M1 basically appear on these two levels of extension, which fact is well shown in the example (28) above. After such a checking process relating to extension as (28), the least violated interpretation is returned: in this case still with one clash, i.e.

(29) #(sleep, furiously).

Meanwhile, to explain such a sequence as ‘Colorless green ideas sleep furiously,’ Oh & Seo (1992) propose a sortal restriction with a ‘concept catalog’ and a lexicon:

(30) [FURIOUS]-X->(OBJ)-X->[SLEEP]-X->[EXPR]-X->[IDEA:{*}]->-----
    FURIOUSLY <ACTION MANNER
    SLEEP [SLEEP]- <<<<<Sam sleeps quickly>>>>
    (EXPR) ⇒ [ANIMATE],
    furiously adverb; FURIOUSLY
    sleep intransitive verb; SMILE.

Oh & Seo well summarize the researches on sortal incorrectness and propose important generalizations including an abstraction that there is a sortal restriction involving an adverb as a predicate. However, (30) cannot block the combination of ‘sleeps furiously’ with its current information and meets a deviant sequence. The parser based on (30) does not notice the mismatch between ‘furiously’ and ‘sleeps’ with their present framework and cannot test any elevation of meaning because the possibility of such elevation is blocked by the given conceptual catalog. In addition, like any other system treating this kind of anomaly, this system only returns binary replies, yes or no, which are unlikely to human parsers.

Even with an anomalous expression, human parsers try to collect some meanings from it by default however it may be partial. If we respect the principle of least effort which has been rehabilitated by Horn (1988), we
can generalize that a parser expects a smallest maximum degree of specification for a proper interpretation by exploiting smallest maximum perceptual strategy simultaneously. Thus, even a suspected ‘deviant’ sentence would receive a partial interpretation at its most. An interpretation by a human parser, in general, is not an all-or-nothing matter.

Our approach through collocational continuum respects this tendency and procrastinates its final judgment till it gets the most feasible candidate of interpretation. If (26a) may not be globally interpreted to its end, it receives partial local interpretations. With a context or situation provided, a parser will select a best interpretation and return a final degree of appropriateness judged from contexts provided.

This procedure implies that semantic anomaly can be cured by some contextual manipulation: e.g.,

(31) The table saw Arthur.  

is a semantic anomaly since no metaphorically extended meaning can save it from deviation, but it can be improved as a somewhat ‘ironic hyperbole’ through a contextual manipulation (Cruse 1986: 7):

(32) Arthur is paranoiac. He believes all his accidents are due to a cosmic conspiracy. No doubt the table saw him, computed his path across the room, and placed itself just where he would trip over it!

Sentences of this kind are easily found in English as an instance of ‘grammatical metaphor’ in Halliday’s (1985: 319ff) term. Halliday divides metaphors into lexicogrammatical and grammatical ones. The former arises from lexicogrammatical selection of words, while the latter relates to a strong grammatical effect in rhetorical transference. An example of the latter case is (1) ‘The fifth day saw them at the summit,’ which can be paraphrased with some grammatical manipulation as (2) ‘They arrived at the summit on the fifth day.’

But there are borderline cases where it is uncertain whether they contain such grammatical metaphors or metonyms: e.g., (3) ‘The eighteenth century saw many reforms.’ If ‘the eighteenth century’ relates to a grammatical metaphor, we can restate it, by analogy with (2), as (4) ‘They witnessed many reforms in the eighteenth century.’ But a discrepancy lies between (2) and (4), resulting from the following asymmetry: (2’) ‘The people of the fifth day arrived at the…’ [≠ (2)]

(4’) ‘The people of the eighteenth century witnessed…’ [≠ (4)]

That is, (4’) explicitly has the potentiality to become a metonym because the underlined part is contiguous in syntagm, while (2) is not.
Such a semantic oddity could be normalized by the contexts of a fairy-tale, science-fiction or personification in literature. Hence, contextual manipulation is a fairly common device human imagination resorts to. Thus, even (26a) could be fully interpreted if it could be contextualized. Devoid of such a context, the expression will be marked as inappropriate.

In sum, there are two kinds of deviation along the collocational continuum and a parser tries to understand such derailed expressions through two kinds of processing strategies, i.e., default and compositional processing. In the cases of incongruity from violating selectional restrictions, the speaker intends the hearer to make sense of his words by elevated or metaphorical interpretations. Like the violations of rule-based expectation, to breach collocations is not normally tolerated, but the hearer tries to find any situational (or contextual) support to that breach to understand the speaker's intention.

3.2. Metaphor

With an idea of continuum, the habitualness of collocation co-variates with rule-based analogy. As seen in 3.1, such a notion of collocational processing can provide metaphoric or deviant expressions with relevant descriptions. As for novel or dead metaphors, learners, teachers or translators of a language should be able to properly understand those of a target language which seem to be deviant at first. The metaphors in a language are distributed in a gradient continuum, i.e., from live ones down to the frozen dead metaphors which the speakers of that language often use without knowing whether they are figurative or even violating the selectional restrictions.

The metaphoric phenomena range from phrases to composite words: e.g.,

(33) white hope, blue laws, white magic, smart money, category killer, etc.

For instance: “The producers of ideas have all dull-colored faces and are inexperienced. They are never attractive to their boss and always warned to be fired with slaps on their back. So they often gather to drink together to be good and loaded. When they fall asleep, some guys yell and yell, others roll from pain in their stomachs and hearts, still others, somnambulists, wander around the room drunken and with their faces in fury……”
Survivals among the novel metaphors become established in a language to enter the lexicon as set phrases, while obsolete clichés or lexemes drain away from the lexicon unnoticed. Thus, the metaphoric cline crisscrossing the collocational continuum is one of the pipelines filling up the lexicon with new lexemes.

Among the efforts to explain metaphoric phenomena, Lakoff & Johnson (1981: 289) try to give metaphor a conceptual explanation. They claim that the metaphors in (35) can be interpreted on the basis of the ‘conceptual metaphor’ in (34):

(34) ARGUMENT IS WAR
(35) a. Your claims are indefensible
    b. He attacked every weak point in my argument
    c. His criticisms were right on the target
    d. I demolished his argument
    e. I've never won an argument with him
    f. If you use that strategy, he'll wipe you out
    g. He shot down all my arguments.

In these examples, the semantic co-relation between ‘war’ and ‘argument’ enables us to understand those expressions. Semantically, this conceptual model captures similarities among different terms, which converge on one basic or underlying term like ‘ARGUMENT.’ Thus, perhaps the examples in (35) are all related to one primitive term as their focal sense or their best representative. If one understands those metaphors by analogy to (34), such recognition will be rather indirect. Indirectness in recognition implies a conceptual hierarchy and some distances therein between derived metaphors and their underlying primitive-distances to be filled by an inference.

If a primitive is not too deep to be accessed, it will be similar to the notion of genus in our collative semantics which implies a semantic network and calculates metaphors. But our system may take an alternative to this rule-based inference as well: i.e., since the metaphors in (35) are set or nearly set expressions, their use meanings and collocational lists are represented in frames and processed by default, not by rules. We can calculate those metaphors in (35) by invoking rules but here we will show a collocational alternative, beginning with the frozenness of the given expres-
sions:

(36) a. In (35a), since ‘indefensible’ selects ‘claims,’ the latter is specified
in the CSPACE of the former: ‘indefensible’ will have ‘ideas,’
‘statements’ or ‘claims’ in its CSPACE list and ‘cannot be justified’
in its (CSPACE) SEM as an extended meaning.
b. Likewise, in (35b) both ‘attack’ and ‘weak’ select ‘argument’
and ‘point’ as collocates and, for instance, ‘attack’ has ‘speak or
write strongly against’ as a meaning extended from its literal one,
‘use physical violence against.’
c. In (35c) ‘criticism’ is selected by a predicative element, ‘right on
the target,’ of an adjacent collocation in which, for example, ‘tar­
get’ has an extended meaning, ‘effect.’
d. In (35d) ‘demolish’ has an extended meaning ‘prove X to be
wrong where X is an idea, argument or belief.’
e. In (35e) ‘win’ predicts as collocates ‘competition, battle, argument,
bet, etc.’ among which ‘competition’ will be a best example.
f. In (35f) with ‘wipe out’ of a meaning ‘punish with stern censure’
extended from ‘rub,’ the basic meaning of ‘strategy,’ i.e., ‘art of
exercising forces of X such that X is typically military,’ is extended
to ‘art of planning to get X done where X is an action’ but the rela­
tion of these two collocates in an indirect relation will have
stronger cohesion if the ellipses are recovered from the context,
e.g., like ‘strategy in your argument to win over him’ and ‘wipe
you out for such argument of yours.’
g. Finally, in (35g) ‘shoot down’ has an extended idiomatic mean­
ing ‘show X to be wrong or foolish where X is a view, a proposal, an
argument or an idea.’

In comparison to such default processing (cf. Yoon & Kim 1993), the
conceptual calculation will cause more time–spatial complexity because it
should compute each of the expressions separately. Our present system cal­
culates novel metaphors by computing interpretations with sense frames
after the default procedure of matching them against the CSPACE infor­
mation finishes undefined. Calculating the distinction between novel meta­
phors like (37a) and anomalies like (37b), we can see how the system
works.
(37) a. The car drinks gasoline
    b. I ate three phonemes for breakfast (Allan 1986: 141-42).

To calculate metaphoric expressions like (37), we will consider the sframe first. This sframe consists of arcs and nodes: the former holds genus information and the latter differentia information:

(38) a. sframe(drink1,
    [[arcs,
      [[[[supertype, consume1]]],
      [node2,
        [[[[agent,
          [preference, animal1]],
        [[object,
          [preference, drinking_liquid1]]]]]]])
    b. sframe(carl,
       arcs
       supertype, vehicle1
       node 0
       it1, use1, gasoline1)
    c. sframe(use1,
       arcs,
       supertype, consume1
       node2,
       agent, vehicle1
       object, gasoline1)26

Collation proceeds on the basis of a semantic network comprising (38):

(39) a. genus information [= subtype relation] check:
    'carl' is not a subtype of 'animal1' which is an 'agent' of 'drink1'; if a subtype relation does not hold, the meaning of (37a) is not literal.
    b. differentia information [= coordinate relation] check:
    as differentia information, 'carl' has 'carl use1 gasoline1' and

26Squared brackets are omitted since indentation disambiguates.
'animal1' has 'animal1 drink1 drinking_liquid1'; now 'use1' and 'drink1' share 'consume1' as genus information; 'gasoline1' is a subtype of 'liquid1' and 'drinking-liquid1' is a subtype of 'liquid1'; so 'car1' and 'animal1' share 'consume1 liquid1'; so a coordinate relation exists between 'gasoline1' and 'drinking-liquid1' as well as 'use1' and 'drink1'; thus 'car1' and 'drink1' is in a metaphoric relation.

If the first subtype relation check for the subject of a sentence fails and the coordinate relation check for the remaining part succeeds, then this defines a metaphoric relation. If both the genus and differentia checks hold, it defines a literal-meaning relation. If the first subtype relation holds but coordinate relation fails, it defines a metonymic relation, since metonymy is usually so frozen in a context or register that it should be specified as sortal information in the main dictionary or in a temporary dictionary to be properly processed. Finally, if the two kinds of checking both fail, the meaning relation concerned is neither literal, metaphoric nor metonymic: i.e., it could be deviant if it is not supported by an appropriate situation or context. Thus, our framework correctly explains a sentential metaphor in (37a).

<table>
<thead>
<tr>
<th>(40) semantic relation</th>
<th>subtype check</th>
<th>coordination check</th>
</tr>
</thead>
<tbody>
<tr>
<td>literal meaning</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>metaphor</td>
<td>✗</td>
<td>0</td>
</tr>
<tr>
<td>metonymy</td>
<td>0</td>
<td>✗</td>
</tr>
<tr>
<td>anomaly</td>
<td>✗</td>
<td>✗</td>
</tr>
</tbody>
</table>

In the case of (37b), however, there still remains a problem: when we collate 'eat1 food1' with 'DO-X-WITH phoneme1' where 'DO-X-WITH' is typically related in a semantic network to 'study, use, or postulate, ⋯⋯⋯,' the collation fails unless we go up to the highest node, 'act or do something,' in the action domain of a semantic network. Thus, by default (37b) is not literal. Likewise, the 'theme' of 'eat1,' i.e., 'food1,' and 'phoneme1' do not share a supertype unless we arduously climb up to the top node, i.e., 'thing,' of the related domain; therefore, collation fails. So this sentence should be deviant unless provided with some feasible contextual information.

This is the spot semantic researchers usually stop or hesitate to formulate
any explanation. But with appropriate contextualization or normalization, this sentence will be saved from deviancy. Then let’s see an exemplar context (cf. Allan 1986: 142):

(41) Breakfast was served for participants in a Linguistic Society conference and the breakfast cereal was made in the shape of letters, exciting some witty linguist to utter (37b), thereby referring to the fact that he had eaten a ‘p,’ a ‘t’ and a ‘k’ shape.

With this context, the remark (37b) would be felt normal. Then is it a far-fetched metaphor because sentences tend to be judged anomalous owing to their unfitness for the familiar ordinary world? Allan (1986), for example, does not give an explicit answer.

Even though a relevant context is given, (37b) may not be wholly appropriate, if the speaker omits the relevant element(s) in a syntagmatic combination like ‘phoneme-like cereal’ and causes difficulty in recognition or if he/she mistakenly refers to an ‘alphabet cereal’ as a ‘phoneme cereal.’ In those cases, (37b) can be said to violate felicity condition in terms of Speech Act theory: it may not be totally anomalous but it is still subject to inappropriateness. Therefore, we can distinguish between anomaly and inappropriateness.

3.3. Metonymy

So far, discussing extension of meaning, we have equated elevation with metaphorical extension. Still, there remains another axis of extension of lexical meaning: i.e., the metonymic level. Semanticists usually miss the metonymic relation in the extensions of lexical meaning (cf. Allan 1986). Metonymy is frequently used in an extension of lexical meaning, nevertheless. It links the elements syntagmatically to each other, while a metaphor is a link between the elements from two different paradigmatic sets (cf. Ducrot & Todorov 1979 [1972]: 111).

Briefly speaking, in metonymy an object is designated by the name of another object which pertains to that object, while in metaphor an object is designated by the name of a similar object belonging to a different conceptual field on the semantic hierarchy. Thus, metonymy is similar to encapsulation and related to the shortening of a lexical collocation. Meaning by me-
tonymy is learned by use or habit, when meaning by metaphor can be calculated by inference, given a sound semantic hierarchy. Therefore, if such a dualistic explanation of meaning extension is allowed, possible heuristics or algorithms can be found out to represent or implement every possible lexical meaning.

More technically, metaphor can be defined as an extension or shift in genus information and metonymy as an extension or shift in differentia information of a lexeme. When metaphor implies a sortal change, maintaining similarity in differentia information, and thus even without a context a typical meaning by metaphor does not change, metonymy maintains sortal information in a context, implying shift of differentia information in that context, and without a context such shifted meaning by metonymy would be hard to understand.

Then, let's see the rhetorical meaning of (42) (cf. Levinson 1983, Ch. 3) as what is structured by both metonymy and metaphor on the basis of the above distinctions, though Levinson does not make explicit his judgment about (42), i.e., whether it is a metaphor or a metonymy.

(42) Britain rules the waves.

Since 'rule' literally means 'control or the matters/affairs of,' our system may first return the compositional meaning of 'ruled the waves' with the literal meaning of 'waves': 'control or decide the matters/affairs of the waves.' But it is not the case because, unlike 'people, country or any domain of area,' 'waves' does not belong to the set of frequent collocates of 'rule' and so such a reading calculated above is not a preferred one.

So instead of rendering such a reading, the system seeks another common possibility that 'the waves,' not 'wave,' can be extended metonymically (since 'waves' are not an area but accidents in that area, i.e., the 'sea') to mean 'the sea' or 'the ocean' (strictly speaking this is a synecdoche which is a subtype of metonymy). But in fact, 'rule the waves' is already a set collocation and the system returns the final reading by default. In addition, if we show the structure of the rhetoric of this sentence, 'ruled the waves' is a metaphor since literally there is no genus relation between the themes of 'rule,' i.e., 'people or area' and 'the waves' and so on. Therefore, the figurative meaning of this sentence is structured roughly in such a way as
(43) rhetorical meaning of (42) =
METONYMY (the waves \rightarrow \text{the sea}) + METAPHOR ((\text{rule, the sea}) \rightarrow (\text{secure the command of maritime affairs})).

4. Conclusion

Collocation can contribute to economic and systematic processing of significant linguistic topics, \textit{i.e.}, metaphor, metonymy and anomaly. With a generalization of the application as well as the concept of collocation, a theory of collocation treated here proves that an efficient computational theory of collocation is feasible.

An over-all picture of collocation can be captured through a view of collocation as continuum. A typical aspect of the continuum is the cline of frozenness. The term 'frozenness' is more comprehensive than that of former studies and covers the continuum from the rule-governed area to the most frozen area. It entails two kinds of expectation: \textit{i.e.}, rule-based and use-based expectation.

The frozenness continuum is partitioned by a 'collocational space' and partial functions which enable the meaning composition of frozen expressions. Since every word, based on the meaning schema $M_2(U)M_1$, has the potentiality of extension in its meaning and thus has some collocational space (CSPACE), the expressions of a language are partitioned in terms of CSPACE and the partial functions into groups: \textit{e.g.}, rare, variable, set, and lexicalized collocations. This knowledge is formally shown in a graph.

Derailment from expectation results in collocational manipulation, anomaly, or figurative extensions of meanings. As a paradigmatic extension of meaning, metaphor was treated in terms of rule-based approaches. Dead or frozen metaphors, however, are efficiently processed in terms of default collocational processing. Only novel metaphors go through non-default compositional processing. Even in the case of novel metaphors or anomalies, this system recognizes more delicate partial information than former approaches.

Metonymy, a syntagmatic extension of meaning, was relatively more neglected in linguistic studies. In calculating an expression left vague in Levinson (1983), we explicitly show the collaboration in our two-edged strategy exploiting default and compositional processing, or the cooperation
between use-based and rule-based inferences. That is, our analysis shows
the example (42) has come from the collaboration between metaphor and
metonymy.

References

Ducrot, O. & T. Todorov (1979[1972]) Encyclopedic Dictionary of the Sciences of Language (Tr. by C. Porter), Johns Hopkins University, Baltimore.
228, Longman, London.


Department of English
Hoshin University
108 Yangrim-dong, Seo-gu,
Kwajangju, 502-756
Korea