POLYSEMY OF CASE MARKERS: A TYPOLOGICAL STUDY

Kazuyuki Yamaguchi

University of New Mexico

0. Introduction

One linguistic phenomenon which has attracted many linguists is what I would call 'conflation of semantic roles in the same case marker, or the same adposition', as in (1)

(1) (a) Mary left *for* Paris.
    (b) Mary worked *for* her children.
    (c) Mary worked hard *for* her exam.
    (d) Mary couldn't sleep *for* pain. [sic.] (Heine et al. 1991: 152)

According to Heine et al., the meanings, or the semantic roles of the English preposition *for* are different. This preposition in (1a) means Allative; (1b), Benefactive; (1c), Purpose, and (1d), Cause. Our questions, which will be attempted to answer in this paper, are:

(2)(a) What are the possible combinations of semantic roles (for example, can Cause and
      Benefactive be expressed by the same morpheme)?
   (b) How universal are certain combinations of semantic roles. For example, how
     universal is the combination of Allative and Purpose? Is this combination found in more
     languages than the combination of, say, Comitative, Instrumental and Manner?

1. Case forms and Semantic roles

Before proceeding to our main discussion, we briefly consider what are case forms and semantic roles. First, concerning case forms, we will use this term in a broad sense: case forms are *the grammatical forms which mark the relationship between the noun(s) and the verb in a sentence.*[1] This definition allows us to treat as case forms not only grammatical or obligatory case markers[2], but also adjunct or semantic (concrete) case markers, which are typically expressed by peripheral expressions such as prepositions or postpositions (for the justification of this broad definition of 'case form', see Blake 1994 and Ohori 1996). Accepting this broad definition, rather than the traditional definition of case, that is, case is the inflectional nominals which express grammatical or semantic functions[3], may be justified by the fact that the grammatical functions such as Subject and Object, which are expressed by inflectional morphemes in Latin or Turkish, are expressed by peripheral expressions like adpositions in Japanese or by word order in English:

(3)(a) Taro ga   Hanako o ket -ta. (Japanese)
    Taro nom. Hanako acc. kick past
    'Taro kicked Hanako.'

(b) My mother scolded the dog severely.
      sub.   obj.

And we use the term 'case forms', instead of 'case markers', because case forms under our definition do not always *mark* grammatical relations such as Subject or Object.

The next issue to consider is semantic roles. It seems to me that there are (at least) three major assumptions or criteria, based on which semantic roles are determined. The first criterion is that based on verb types (Chafe 1970, Foly and Van Valin 1984, Van Valin 1993, Givon 1984). For example, Chafe (1970: 144-166) suggested that Experiencer is the semantic role required by experiencial verbs such as 'feel', 'hear' and 'learn', while Benefactive is the semantic role required by benefactive verbs such as 'own', 'acquire', and 'find'. The second criterion is that based on the action chain model (Langacker 1991) or the causal chain model (Croft 1991). Their account of how semantic roles can be determined is very similar, partly because they are greatly influenced by Talmy (1985). For this reason, our discussion here is limited to Langacker's model. Langacker (1991: 283) proposed the following action chain:

O==>O==>O==>O   Fig. 1

The initial entity (that is, the leftmost object) is called 'the head', and the final entity, 'the tail'. Based on this model, certain semantic roles can be determined. In usual situations, Agent is head; Patient is tail, and Instrument is, according to Langacker, an intermediate entity. The last criterion is what I would call 'the intuition-based criterion'. Those assuming this criterion believe that there are a finite number of universal semantic roles. This criterion may be found in Fillmore (1968)[4], and those working under Chomsky's Government and Binding theory.

However, these three criteria do not seem to be able to work well for any typological study, because they require native speaker's intuition for the language in question. The first criterion requires "native speaker judgement", as Foly and Van Valin (1984: 28) persuasively argues.

"many of the tests for the verb classification require native speaker judgement about particular constructions."

And the second and third criteria cannot be considered as the appropriate criteria due to their too subjective way of determining semantic roles: without native speaker's intuition, one may not correctly argue which noun phrase corresponds to a head or a tail in the action chain, and therefore one cannot determine whether a noun phrase in question is Agent or Patient. And the fact that there has been no agreement as to how many and what semantic roles are in the universal list[5] seems to indicate that semantic roles cannot be determined based only on our intuition. Instead, we will take the position that semantic roles should be empirically determined (see especially Comrie 1981).

Specifically, instead of using one of these criteria previously proposed, we will determine semantic roles by the following procedures.

(4) Procedure 1: Assume that every case form expresses only one function or one
                 semantic role (for example, *with* => Instrumental)
    Procedure 2: Describe different senses of the same case marker based
                 on one's intuition (*with* => Instrument, Comitative)
    Procedure 3: Examine other languages, and if the intuitively different
                 senses of the same case marker in one language are expressed by
                 different case markers in the other languages, then, consider
                 these different senses as different semantic roles.

As an example, consider the English preposition *for*. As the first step, we assume that this preposition has the only one meaning or semantic role, say, Allative as in (1a). Now our intuition may tell that there are other functions or semantic roles in this preposition, as represented in (1b), (1c) and (1d), and so what we do next is to examine other languages if the (similar) meanings could be expressed by distinct case forms. As for Allative, as in (1a) and Benefactive in (1b), they are expressed by different forms in Inuit, an unaffiliated language (*-mut* and *-mik*), so that they can be appropriately considered different semantic roles. As for Purpose, as in (1c), this semantic role seems to be able to be expressed in a different way from Allative and Benefactive in Margi, an Afroasiatic language (*gà* for Purpose,while *àrá* for Allative and Benefactive). Cause is expressed in different way from Allative, Benefactive, and Purpose in Japanese, an unaffiliated language (*de* for Cause, while *ni* for Allative, Benefactive and Purpose). Through our brief examination, then, we can now claim that Allative, Benefactive, Purpose and Cause are distinct semantic roles.

Following the Procedures in (4), the semantic roles relevant to this paper are as follows:

(5) Semantic Roles used in this Study
    Ablative: He rose **from** the chair.
    Allative: He turned **to** the altar and walked towards it.
    Agent: **The robots** assembled the car.
    Benefactive: She did the shopping **for** her mother.
    Cause: He died **from** starvation.
    Comitative: He talked **with** her.
    Function: I used the stick **as** a club.
    Instrumental: She squashed the spider **with** a slipper.
    Locative: The vase is **on/ under** the table.
    Manner: He did it **with** great skill.
    Path: Taro walks **through** a wood.
    Purpose: He went to the Red Rooster **for** some take-away.
    Reference: I told him **about** the incident.
    Price: I bought it **for** two pounds.

Two points must be mentioned. First, as might be expected, there may be several case forms to express the same semantic role (for example, Cause can be expressed by *from* and *by* in English). In this instance, we will list every case form in question. Second, we will focus on the peripheral case forms rather than grammatical case forms, because (a) the former cases show conflation patterns of more semantic roles than the latter cases (for example,the main semantic role expressed by Object is typically Patient) and (b) peripheral case forms are usually more easily recognizable

than the grammatical case forms.

2. The conflation patterns of semantic roles

Now we are in a position to discuss the conflation patterns of semantic roles expressed by case forms. By 'conflation of semantic roles' expressed by one case form, we mean the following kinds of expressions.

(6) (a) I went to New York **with** John. (Comitative)
    (b) He opened the door **with** a crowbar. (Instrumental)
    (c) He swims **with** ease (Manner) [emphasis, K.Y.] (Croft 1990: 9)

The question to consider here is: why does the preposition *with* express (at least) three different semantic roles? Is there any motivation for one case form to express these certain conflation of semantic roles? What is very interesting concerning this issue is that this specific conflation of semantic roles is not limited to *with* in English: this conflation can be found in Hausa and Classical Mongolian (see Croft 1990), in Modern Greek, and the partial combination (i.e.the pair of Comitative and Instrument or that of Instrument and Manner) can be found in many languages (Margi, !Kung and Japanese, for example). Other, often found conflations of semantic roles in the same case forms are, to mention a few, that of Locative/ Allative and Purpose (Thulung: Allen 1975: 149, 381), of Ablative and Cause (Lahu: Matisoff 1982), Ergative-Instrumental and Causal (Tauya: Macdonald 1988: 1990). This may lead us to wonder what other comflation patterns of semantic roles in the same case form can be found. The next section introduces some previous approaches regarding this issue.

3. Previous Research

After her careful examination of the languages of the Bodic branch of Tibeto-Burman, Genetti (1986, 1990) discovered the polysemic patterns of case forms in (7).

(7) Locative > If/ Although, When/While/After[6]
    Ablative > When/ While/ After, Because
    Allative > Purpose
    Ergative[Agent: K.Y]/ Instrumental > Because, When/ While/ After
                                   (partial representation of Genetti 1986)

According to her, we can expect that conflation of Locative, When-sense, and If-sense is more likely than for instance, that of Purpose and If, and this seems true in many languages.

In a similar, but more thorough fashion, Heine et al. (1991), through their research on Ik and Kanuri, both of which belong to Nilo-Saharan, found the polysemy of cases in Fig.1.

Source	Target
Allative	Locative, Benefactive, Dative, Purpose, Time, Cause Manner, clause subordination
Ablative	Locative, Agent, Cause, Manner
Path	Agent, Cause
Locative	Time, Cause, Condition, Manner, clause subordination
Comitative	Instrumental, Means, Conditions, Manner
Benefactive	Dative, Purpose, Cause, Time
Instrumental/ Means	Manner
Time	Cause, Condition, Manner

---   Fig. 1-----------

(Heine et al. 1991: 257)

The meanings or semantic roles under Source are the basic, or original ones, while the meanings under Target are the derived ones (for example, Allative is the original meaning, while Benefactive is the derived meaning from Allative). Although Heine et al. investigate languages in the same family, they imply that their findings are also applicable to other languages.

So far, we have implied the universality of the polysemic patterns proposed by Genneti and Heine et al. But this does not mean that their findings suggest or conclude some universal facts. This is because their findings are based on a limited number of languages, and we are far from being in a position to answer the questions posited in Introduction, which are repeated here as (8):

(8)(a) What are possible combinations of semantic roles (for example, can Cause and Benefactive be expressed by the same morpheme)?
   (b) How universal are certain combinations of semantic roles. For example, how universal is the combination of Allative and Purpose? Is this combination found in more languages than the combination of, say, Comitative, Instrumental and Manner?

The main purpose of this paper is to answer these questions. For this purpose, we need a language sample, with which some universal claims can be made.

4. Language Sample

This section discusses how to sample languages. I suggest that an appropriate framework, or a language sample, on which universal claims will be made, must meet the following criteria in (9).

(9)(a) Every languages must have an equal chance of selection (random sample).
   (b) Languages in the sample must be stratified (stratified sample).
   (c) The sample must include many languages.
   (d) The sample should not include too many languages.

   (9a) appears obvious. The assumption of random sampling is necessary to infer from the

sample the properties of the population. (9b), a stratified sample, is necessary in order not to introduce any obvious genetic bias into the language sample: each stratum is determined with reference to genetic affiliation, and languages are carefully chosen from each stratum. This guarantees that the sample avoids introducing any obvious genetic bias. (9c) is necessary in order to exclude what Croft (1990: 18) calls "pure accident". As an example, consider the fact that the English possessive verb, 'have' often corresponds to the existential verb, 'be' in Japanese.

(10)(a) I have a fever.
    (b) Netu ga aru.
        fever nom.exist (Lit. 'Fever exists.' or 'There is fever with me.')

Based on this fact, one might propose that this would be reduced to the frequently made claim that the Japanese speakers and the English speakers conceptualize the world in different ways (see Hinds 1986). This conclusion becomes doubtful, however, once one notices that the same linguistic phenomenon can be found in African languages. This kind of fallacy can only be avoided if more languages are examined. As for (9d), because more than a certain number of languages would include the languages which are genetically proximate enough to introduce bias in the sample, the sample should have as few languages as possible on the condition that the sample provide enough data for one's hypothesis[7].

For the purpose of providing the language sample which fulfills the above criteria, the Gramcats sample is used. The Gramcats sample was created by Joan Bybee and her colleagues, and is based on the list of 5000 languages made by Voegelin and Voegelin (1978). I have used their sample because it fulfills the conditions in (9): the 76 languages in the Gramcats sample have been carefully chosen to be maximally unrelated to each other. The only difference between the Gramcats sample and mine is that I have chosen only 26 languages from the Gramcats sample. These carefully chosen languages are maximally unrelated languages from their sample. Despite the fact that my sample does not provide as much information as the Gramcats, I chose this sampling method with a small number of languages, because this allows me to examine each language thoroughly. The languages examined here are as follow:

----------------Language sample used in this paper---------------------------
Inuit (Unaffiliated languages), Margi (Afroasiatic), Cheyenne (Macro-Algonquian), Chacobo (Andean-Equatorial), Alyawara (Australian), Koho (Austroasiatic), Motu (Austronesian), Papago (Aztec-Tanoan), Abkhaz (Caucasian), Guaymi (Macro-Chibchan), Kui (Dravidian), Abipon (Ge-Pano-Carib), Karok (Hokan), Modern Greek (Indo-European), Yagaria (Indo-Pacific), !Kung (Khoisan), Slave (Na-dene), Mwera (Niger-Kordofanian), Bari (Nilo-Saharan), Palantla Chinantec (Oto-Manguean), Zuni (Penutian), Shuswap (Salish), Lahu (Sino-Tibetan), Dakota (Macro-Siouan), Buriat (Ural-Altaic), Tok Pisin (Creoles)
--- Fig.2

5. As for data sources

All twenty-six languages are , as mentioned above, taken from the Gramcats library. This consists of descriptive grammars of individual languages and articles. And some languages provide not only descriptive grammars, but also texts and even dictionaries. On the other hand, many grammar books are far from satisfactory as references. Due to the insufficient data of many languages and various background of authors for these grammar books, I faced primarily the following two problems:

(11)(a) The authors of the languages in our sample use different labels for the items
        relevant to our discussion.
     (b) Many sources do not offer enough data relevant to our discussion.

As for (11a), in order to make every language in my sample comparable, I used only the definition of case forms and of semantic roles made in Section One. And to minimize the effect of (11b), I consulted as many accessible sources as possible. Although more data was required for many languages, I chose to follow only what grammar books explicitly state, and not to make any apparently plausible inferences based on them. For example, the grammar of !Kung states that the case form -l xwa has Comitative and Instrument. Looking at this conflation pattern and other languages, one may well suggest that Cause meaning may be expressed by the same case form although the grammar on !Kung, because of the lack of information in it, does not tell us. However we will not take this position and follow only what grammar books tell us.

6. How to Decide the Diachronic Order of Semantic Roles?

This section discusses how to determine the diachronic order of semantic roles in the same case form. As an example, imagine how the semantic roles, such as Allative, Benefactive, and Purpose in the English preposition *to* have developed. This is not an easy question to answer, but it becomes even more difficult for other less well documented languages, because there are no available historical documents. Due to the lack of relevant information, we need some criteria to re-establish the diachronic development of semantic roles. As far as I know, the only criteria to determine the historical development of semantic roles (without historical documents) were proposed by Heine et al. (1991: 156). For this reason and another reason that they respect every historical evidence available to them, we will use their criteria as follows:

A."A grammatical category is more grammaticalized than another category if it is
    etymologically derived from the latter.
B."If two case functions differ from one another only in the fact that one has a spatial function
    whereas the other has not, then the latter is more grammaticalized."
C."If two grammatical categories differ from one another only by the fact that one typically
    implies some human participant whereas the other implies an inanimate participant, then the
    latter is more grammaticized."
D."A category referring to a concept that has potentially three physical dimensions is less
    grammaticalized than one referring to a concept that has only one possible dimension, which
    again is less grammaticalized than one whose referent does not show any physical
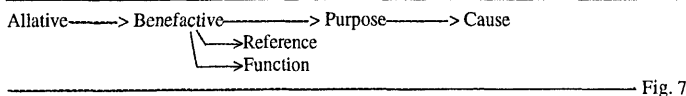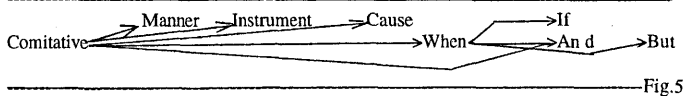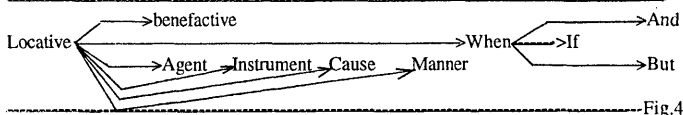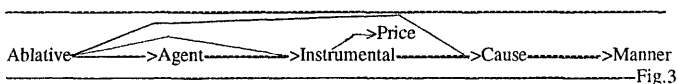
dimensionality. "

E. "If two categories differ from one another only in the fact that one expresses a temporal relation whereas the other expresses some "logical" relation, then the latter is more grammaticalized. "

F. "If two categories differ from one another only in the fact that one is more inclusive, that is, may include the other in certain contexts then the more inclusive is the more grammaticalized one."

Consider exactly how historical development of semantic roles can be determined based on the above criteria by some examples. As shown in (6), the English preposition *with* expresses Comitative, Instrumental, and Cause. According to the criteria A, Comitative precedes Instrumental: "wherever there is historical information available, it turns out that the COMITATIVE use preceded the INSTRUMENTAL use in time, not vice versa" (Heine et al. 1991: 158). And according to the criterion F, INSTRUMENTAL precedes CAUSE in time.

6. Result

Our language sample of twenty-six languages were examined according to the methodology discussed above. Our data of case markers of twenty-six languages in our language sample lead to the following generalization in the following Figures, and our data also suggest how often the diachronic development from one semantic role to another occur is:

The Diachronic Paths of Semantic Roles

Fig.3

Fig.4

Fig.5

Path————————> Cause

Fig.6

Allative--------> Benefactive————————> Purpose————> Cause
                        \————>Reference
                         \————>Function

Fig. 7

Frequency of Minimal Pairs of Semantic Roles on Diachronic Paths

----

1. [Purpose=>Cause](14)   17.[Locative=>Cause](2)
2. [Benefactive=>Purpose](10)   18.[Comitative=>And](2)
3. [Allative=>Benefactive](9)   19.[Locative=> Instrument](2)
4. [Instrument=>Cause](7)   20.[Locative=>Agent](2)
5. [Time Locative(When)=>If](5)   21.[Benefactive=>Function](2)
6. [Benefactive=>Reference](5)   22.[Manner=>Cause](1)
7. [Instrument=>Manner](4)   23.[Ablative=>Instrument](2)
8. [Instrument=>Manner](4)   24.[Comitative=>Time Locative(When)](1)
9. [Comitative=>Instrument](3)   25.[Time Locative(When)]=>And](1)
10. [Locative=>Time Locative](3)   26.[Benefactive=>Cause](1)
11. [Agent=>Instrument](2)   27.[Locative (Path)=>Cause](1)
12.[Ablative=>Cause](2)   28.[Locative=>Manner](1)
13. [Locative=>Benefactive](2)   29.[Instrument=>Purpose](1)
14. [Comitative=>Manner](2)   30.[Time Locative=>But](1)
15. [Ablative=>Agent](2)   31.[Time Locative=>Cause](1)
16. [Agent=>Cause](2)   32.[Comitative=>Cause](1)
---Fig.8----------------------

7. Conclusion

The aim of this paper was to answer the questions in (2), and we suggested that our findings based on twenty six genetically unrelated languages provide the answers for these questions. Our findings showed that there are only a finite number of conflation patterns of semantic roles in the same case form, and certain conflation patterns of semantic roles are much more frequent than others. These findings may make us wonder, for example, what motivate the conflation of Purpose and Cause much more frequent than other conflation patterns of semantic roles, and this will be our future research.

Note

(1) As might be implied from our definition of case, we will not discuss what is usually called the genitive case.

(2) For the distinction between grammatical and concrete (semantic) case, see Kurylowicz 1964: 179.

(3) For the traditional definition of case, see especially Comrie 1986, 1991.

(4) "The case notion comprise a set of universal, presumably *innate* , concept which identify certain types of judgements human beings are capable of making about the event..." (emphasis mine Fillmore 1968: 24)

(5) "The most frequent criticism is also the most serious-- that neither Fillmore nor his any of his followers has succeeded in producing a definitive list of the underlying cases [semantic roles: K.Y]." (Janda 1993: 31)

(6) Although our focus is on case forms, we also need to take clause linkage markers into our consideration because they have typically developed from case forms, and then they do not have any clear boundary to be divided. For this reason, clause linkage markers will be considered wherever appropriate.

(7) Joan Bybee in personal communication suggested that more than 100 languages may re-introduce genetic bias in the sample.

Reference

Blake, Barry. (1994). *Case*. Cambridge: Cambridge University Press.

Comrie, Bernard. 1981. *Language Universals and Linguistic Typology*. Oxford: Basil Blackwell.

------------------. 1986. On delimiting cases. In *Case in Slavic,* ed. by Richard D. Brecht and James S. Levine, 86-106. Columbus, Ohio:Slavica.

-------------------. 1991. Form and function in identifying cases. In*Paradigms: The Economy of Inflection.* ed.by Plank, Frans. Mouton de Gruyter.

Croft, William. 1990. *Typology and universal.* Cambridge: Cambridge University Press.

----------. 1991. *Syntactic categories and grammatical relations.* Chicago: University of Chicago Press.

Chafe, Wallace L. *Meaning and the Structure of Language.* Chicago: Chicago University Press.

Fillmore, Charles. 1968. The case for case. In *Universals in linguistic thoey,* ed. Emmon Bach and Robert T. Harms, 1-90. New York: Holt, Rinehart & Winston.

Foly, William A., and Robert D. Van Valin, Jr. 1984. *Functional syntax and universal grammar.* Cambridge: Cambridge University Press.

Genetti, Carol. 1986. The development of subordination from postposition in Bodic languages. *BLS 12: 387-400.*

-------------. 1991. From postposition to subordination in Newari. In Elizabeth C. Traugott & Bernd Heine (ed.) 2: 227-255.

Givon, Talmy. (1984). *Syntax: A functional-typological introduction. vol.1: 140.* Amsterdam/ Philadelphia: John Benjamins.

Heine, Bernd. 1988. The dative in Ik and Kanuri. In Croft, Denning, and Kemmer 1990: 129-49.

------------------ et al. 1991b. *Grammaticalization.* Chicago and London: The University of Chicago Press.

Kurylowicz, Jerzy. 1964. *The inflectional categories of Indo-European.* Heidelberg: Winter.

Langacker, Ronald. 1991. *Foundations of cognitive grammar, vol.2: Descriptionve application.* Stanford: Stanford University Press.

Macdonald 1988. Subordination in Tauya. in John Haiman & Sandra A. Thompson (eds.): 227-246.

Matisoff 1982. *The grammar of Lahu.* Berkeley and Los Angeles: University of California Press.

Ohori, Toshio. 1996. Case markers and clause linkage: Toward a semantic typology. In Eugene H. Casad (ed.) *Cognitive Linguistics in the Redwoods: 693-712.* Berlin: Mouton De Gruyter.

Talmy, Leonard. 1988. The relation of grammar to cognition. In *Topic in cognitive linguistics.* ed. Brygida Rudzka-Lstyn. 165-205. Amsterdam: John Benjamins.