【Forum: Workshop on *The World Atlas of Language Structures* and Typological Analysis】

Diversity of Cases: Using *The World Atlas of Language Structures**

Masahiko Nose

(Tohoku University/Japan Society for the Promotion of Science)

Key words: WALS, typology, case, locative, linguistic diversity

1. Introduction

The World Atlas of Language Structures (*WALS*) (Haspelmath *et al.,* 2005) is a convenient tool not only for typologists but also for all linguists of all persuasions. *WALS* is a database of more than 2000 languages in the world containing information on phonological, morphological, syntactic, and other features (a total of 142 features). This paper deals with a specific morphological feature highlighted in *WALS*, namely the number of cases (Iggesen 2005a). Selecting languages with more than 10 cases, I examine the kind of cases the languages have, and then consider the reasons why these multiple-case languages do indeed have so many cases. Finally, I discuss whether the languages with rich case systems have other common typological features as described in *WALS* (Dryer 2005a, b).

First, "case" should be defined. According to Blake (1994: 1), "[c]ase is a system of marking dependent nouns for the type of relationship they bear to their heads. Traditionally the term refers to inflectional marking, and typically, case marks the relationship of a noun to a verb at the clause level or of a noun to a preposition, postposition or another noun at the phrase level." Following Blake's definition of case, Iggesen (2005a) examined 261 languages for the number of cases, the results of which are shown in Table 1.

Table 1: Number of cases in 261 languages by Iggesen (2005a:202–205)

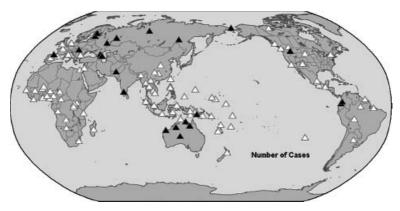|	
No morphological case-marking	100
2 case categories	23
3 case categories	9
4 case categories	9
5 case categories	12
6–7 case categories	37
8–9 case categories	23
10 or more case categories	24
Exclusively borderline morphological case-marking[1]	24

In Table 1, the two notable features are: (i) there is no morphological case-marking in 100 languages including Arabic, Swahili, Mandarin, Thai, etc.; and (ii) there are 24 languages that have more than 10 cases, which include Awa Pit, Basque, Brahui, Chukchi, Epena Pedee, Estonian, Evenki, Finnish, Gooniyandi, Hamtai, Hungarian, Hunzib, Ingush, Kayardild, Ket, Lak, Lezgian, Martuthunira, Mordvin (Erzya), Nez Perce, Nunggubuyu, Pitjantjatjara, Toda, and Udmurt, and the number of cases are shown in Table 2.

---

1) There are 24 languages of "exclusively borderline morphological case-marking." This feature refers to the languages that have overt marking only for concrete (or peripheral) case relations, such as locatives or instrumentals (Iggesen 2005a, b). This type is represented by Plains Cree (Algonquian; Saskatchewan, Alberta), in which the only case-inflecting device used is the locative suffix.

Table 2: The 24 languages with more than 10 cases and their number of cases

Language	Language Family	Area	Number of Cases
Awa Pit	Barbacoan	South America	12
Basque	Basque	Europe	15
Brahui	Dravidian	Eurasia	11
Chukchi	Chukotko-Kamchatkan	Eurasia	13
Epena Pedee	Choco	South America	15
Estonian	Finno-Ugric	Europe	14
Evenki	Altaic	Tungusic, Eurasia	12
Finnish	Finno-Ugric	Europe	15
Gooniyandi	Australian, Bunuban	Australia-New Guinea	13
Hamtai	Trans-New Guinea, Angan	Australia-New Guinea	10+1?
Hungarian	Finno-Ugric	Europe	18
Hunzib	Nakh-Daghestanian	Eurasia	11
Ingush	Nakh-Daghestanian	Eurasia	8+5?
Kayardild	Australian, Tangkic	Australia-New Guinea	13
Ket	Yeniseian	Eurasia	12
Lak	Nakh-Daghestanian	Eurasia	35
Lezgian	Nakh-Daghestanian	Eurasia	18
Martuthunira	Australian, Pama Nyungan	Australia-New Guinea	13+8?
Mordvin (Erzya)	Finno-Ugric	Eurasia	13
Nez Perce	Pnutian, Sahaptian	North America	13
Nunggubuyu	Australian, Gunwinyguan	Australia-New Guinea	11+2?
Pitjantjatjara	Australian, Pama Nyungan	Australia-New Guinea	13
Toda	Dravidian	Eurasia	13
Udmurt	Finno-Ugric	Eurasia	15

Figure 1 shows a geographical map generated by *WALS*, where white triangles represent the languages with no case marking and black ones represent the languages with more than 10 cases. White triangles are densely distributed in Africa and South Asia, and black ones are scattered in Eurasia and Australia. In this paper, I focus on the black ones, which comprise 24 languages with multiple cases, and further I discuss the distribution of cases and their relations to other grammatical features in *WALS*.

△ 1. No morphological case-marking [100]
▲ 8. 10 or more cases [24]

Figure 1: The languages with no case-making and with more than 10 cases by Iggesen (2005a) "Number of Cases" (feature 49, *WALS*: 202–205)

2. Collecting and classifying cases

I have collected examples of cases from descriptive grammars of individual languages with more than 10 cases. The exact (or approximate) number of cases is shown in Table 2. The numbers of the cases in several languages depend on the descriptions of each language and their numbers are different from the books. The numbers in Table 2 are partly different from Iggesen's (2005a, b) findings.

There are two problems in collecting and classifying the cases. First, it is difficult to count the exact number of cases, although I have carefully followed the definition of case (cf. Borin's (1986) study of counting cases in Hungarian[2]). For example, Martuthunira, Nunggubuyu, and some other languages have semantically peripheral suffixes or postpositions other than cases, so it is hard to distinguish cases from such non-case

2) Borin (1986) discusses the exact number of cases in Hungarian. According to his study, the case number in Hungarian varies widely from 5 to 27 or more. The definition of the case is different from those in previous studies or descriptive grammar. This study finally decides on 18 cases, but Iggesen (2005a) claims that Hungarian has 21 productive cases.

markers. This paper follows Blake's (1994) definition and carefully counts the number of cases. Nevertheless, the case numbers of Martuthunira and Nunggubuyu shown in Table 2 are not exact (see also Hamtai and Ingush).

A second problem is that of assigning different case names to a single meaning, or the same case name to different meanings. For example, Finnish, Estonian, and Mordvin have an *abessive* case indicating "without" or "lacking," whereas Chukchi, Kayardild, Martuthunira, and Pitjantjatjara have the *privative* case indicating the same meaning. Moreover, Gooniyandi employs the *deprivative* case, and Ket and Udmurt the *caritive* case for almost the same function. On the other hand, Estonian and Finnish have an *allative* case indicating "onto," and Basque and Chukchi also have an *allative* case, but their meaning is "to" (including the dative function) or "towards." Although the case name is allative, it has a slightly different function in each of the languages concerned.

To solve these problems, I have assigned several functional descriptions to each case observed in the 24 languages. For instance, Awa Pit has 12 cases (Curnow 1997), and the relevant functional descriptions have been added to each of them (see Table 3).

Table 3: Cases and functional descriptions in Awa Pit (Curnow 1997)

Case names	Case forms	Functional descriptions
1. Nominative	φ	[subject], [grammatical]
2. Accusative	*Ta*	[object], [grammatical]
3. Locative-allative 1	*Ta*	[in], [place]
4. Locative-allative 2	*pa*	[in], [around], [place]
5. Locative-allative 3	*mal*	[to], [place]
6. Locative	*ki*	[at], [place]
7. "Until"	*kima*	[until], [termination], [destination]
8. Possessive	*Pa*	[possession], [part-whole]
9. "With"	*Kasa*	[with], [instrumental], [comitative]
10. "Because"	*Akwa*	[because], [reason]
11. "Like"	*Patsa*	[like], [equative]
12. Semblative	*Kana*	[similarity], [figurative]

By assigning several non-strict functional descriptions to cases, it is possible to classify the cases in terms of syntactic and semantic functions. Finally, we summarize the results obtained from the analysis of the 24 languages, where 25 universal functions are identified.

3. Results and discussion

Due to the assigning of several functional descriptions, we obtain four major kinds of functions: grammatical, locative, adverbial, and others. Furthermore, there is no "rare case" in the 24 languages, and the cases observed in this paper are usually different types of locative cases. Some "rare cases" and their functional descriptions are observed in (1), but the same or almost the same case is usually observed in other languages also. The cases in Hamtai, Papua New Guinea are quite peculiar.[3] For example, there are three emphasis cases: *-hga, -ma, -na*;, one for thought, one for person and one for thing.

(1)   Some rare cases
        Perlative (Chukchi), Prosecutive (Ket): [place], [along]
        Dissimilarity (Epena Pedee), Comparative (Ingush): [comparison], [than]
        Emphasis on thought, person, thing (Hamtai): [only]
        Proprietive: (Kayardild, Martuthunira), Relative (Nunggubuyu): [having], [relational]

   Table 4 summarizes the 25 "universally identifiable case functions" of all the cases observed. These functions are selected and classified from the functional descriptions. The 25 case functions are grammatical, locative, adverbial, or others in terms of the functional descriptions. For example, in Table 3, we have already shown the cases in Awa Pit, and its case dis-

---

   3) The data in Hamtai is taken from a descriptive grammar written by Oates & Oates (1968), according to which, Hamtai does not have any nominative/accusative or ergative/absolutive case. However, the *WALS* data on Hamtai shows that it is an accusative-type of language.

Table 4: Universally identifiable case functions

Grammatical:	Locative:	Adverbial:
[1] Nominative	[7] General-locative	[17] Abessive
[2] Accusative	[8] In: Inessive	[18] Causal
[3] Ergative	[9] Into: Illative	[19] Comparative
[4] Absolutive	[10] From inside: Elative	[20] Terminative
[5] Dative	[11] On: Superessive	[21] Essive
[6] Genitive	[12] Onto: Sublative	[22] Instrumental-comitative
–	[13] From ontop: Delative	[23] Translative
–	[14] Adessive	Others:
–	[15] Allative	[24] Relational-concerning
–	[16] Ablative	[25] Vocative

Table 5: Case distributions in Awa Pit from universally identifiable case functions

Case functions	Exists or not	Specific case examples
[1] Nominative	Yes	Nominative
[2] Accusative	Yes	Accusative
[3] Ergative	No	–
[4] Absolutive	No	–
[5] Dative	No	–
[6] Genitive	Yes	Possessive
[7] General-locative	Yes	Locative
[8] In: Inessive	Yes	Locative-allative 1, 2
[9] Into: Illative	Yes	Locative-allative 3
[10] From inside: Elative	No	–
[11] On: Superessive	No	–
[12] Onto: Sublative	No	–
[13] From ontop: Delative	No	–
[14] Adessive	No	–
[15] Allative	No	–
[16] Ablative	No	–
[17] Abessive	No	–
[18] Causal	Yes	Because
[19] Comparative	Yes	"Like", Semblative
[20] Terminative	Yes	"Until"
[21] Essive	No	–
[22] Instrumental-comitative	Yes	"With"
[23] Translative	No	–
[24] Relational-concerning	No	–
[25] Vocative	No	–

tributions may be described in terms of universally identifiable case functions, as in Table 5.

In Table 5, a checklist of case functions, and description of the case functions of the 24 languages are given. Below, in Figure 2, an attempt is made to visualize the results of the checklist by using a bioinformatics software (SplitsTree4; Huson & Bryant 2005). Figure 2 demonstrates similarities between the languages in the distribution of the universally identifiable functions. When two languages share many case functions, they are visualized as closer, and thus we can observe linguistic distances in terms of case functions.

There are several interesting observations to be made from Figure 2. There are geographically groups and also genealogically related groups, i.e., language families, such as Finno-Ugric (Hungarian–Estonian–Finnish–Udmurt), Australian (Martuthunira–Kayardild–Pitjantjatjara), Caucasus (Lezgian–Lak–Hunzib), and Dravidian (Brahui–Toda), and
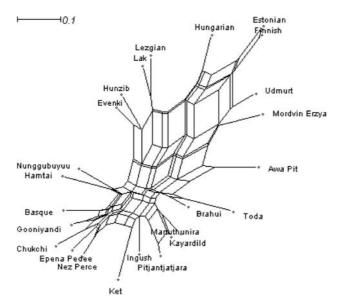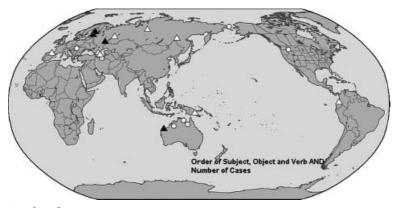

Figure 2: Visualization of relationship of case marking

deviant languages: Awa Pit, Basque, Hamtai, Ingush, and Ket. Moreover, there are several peculiar pairs in the visualization: Ingush–Ket, Epena Pedee–Nez Perce–Chukchi, Hamtai–Nunggubuyu and Hunzib–Evenki.
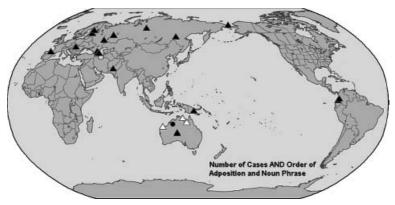
Finally, I contrast the number of cases (feature 49) with other *WALS* features, and argue that there are other features common among the 24 languages with more than 10 cases (cf. Nichols 1992). First, we analyze the combination of the number of cases and word-order patterns (WALS feature 81) in Figure 3.

Word-order pattern SVO is found in four languages (Estonian, Finnish, Martuthunira and Mordvin), and no dominant order in six (Chukchi, Gooniyandi, Hungarian, Kayardild, Nez Perce, and Nunggubuyu). SOV, meanwhile, is the typical order in nine languages (Awa Pit, Basque, Evenki, Hamtai, Hunzib, Ingush, Ket, Lezgian, and Udmurt). Second, we analyze the combination of the number of cases and adpositions in Figure 4.

△ 1. [9:9:7] 10 or more cases AND SOV
○ 2. [6:6:4] 10 or more cases AND No dominant order
▲ 3. [4:2:2] 10 or more cases AND SVO

Figure 3: Combined features: Iggesen (2005a: 202–205) "Number of Cases" (feature 49) (more than 10 cases) and Dryer (2005a: 330–333) "Order of subject, object, and verb" (feature 81)

▲ 1. [18:15:11] 10 or more cases AND Postpositions
△ 2. [3:3:1] 10 or more cases AND No adpositions
● 3. [1:1:1] 10 or more cases AND Inpositions

Figure 4: Combined features: Iggesen (2005a: 202–205) "Number of Cases" (feature 49) (more than 10 cases) and Dryer (2005b: 346–349) "Order of adposition and noun phrase" (feature 85)

Among the languages with more than 10 cases, there is none that has prepositions. Instead, they predominantly have postpositions (18 languages). Inpositions are used in Gooniyandi in Australia. Furthermore, Kayardild, Martuthunira, and Nunggubuyu do not have any adposition and such languages seem to use many suffixes, instead of adpositions.

4. Conclusion

This paper has focused on one of the *WALS* features, "number of cases" and has aimed to demonstrate a practical usage of *WALS*. The languages with more than 10 cases in *WALS* have the following characteristics:

First, case distributions are not always consistent with genealogical or geographical distributions.

Second, it has turned out to be clear that when a language has a rich case system, it has many locative cases and adverbial cases. There are varieties of locative meanings: [on], [in], [by], [below], [under], [along], etc.

Finally, as a result of contrasting case with other *WALS* features, it

has been established that the languages discussed in this paper have other frequent grammatical tendencies, SOV word order, and postpositions. Gooniyandi has no dominant order but employs inpositions.

References

A. Descriptions of sample languages

Awa Pit

Curnow, Timothy J. (1997) A grammar of Awa Pit (Cuaiquer): An indigenous language of western Colombia. Unpublished doctoral dissertation. Australian National University.

Basque

Hualde, José Ignacio and Jon Orlitz de Urbina (eds.). (2003) *A grammar of Basque*. Berlin: Mouton de Gruyter.

Brahui

Andronov, Mikhail S. (1980) *The Brahui language*. Moscow: Nauka.

Chukchi

Dunn, Michael John (1999) A grammar of Chukchi. Unpublished doctoral dissertation, Australian National University.

Epena Pedee

Harms, Philip Lee (1994) *Epena Pedee Syntax*. Arlington: Summer Institute of Linguistics and University of Texas.

Estonian

Oinas, Felix J. (1966) *Basic course in Estonian*. Bloomington: Indiana University Press.

Evenki

Nedjalkov, Igor (1997) *Evenki,* Descriptive grammar series. London: Routledge.

Finnish

Karlsson, Fred (1999) *Finnish: An essential grammar*. London: Routledge.

Gooniyandi

McGregor, William. (1990) *A functional grammar of Gooniyandi*. Amsterdam:

John Benjamins.

Hamtai

Oates, William J. and Lynette F. Oates (1968) *Kapau pedagogical grammar,* Pacific Linguistics, series C-10. Canberra: Australian National University.

Hungarian

Keresztes, László (1995) *A practical Hungarian grammar*. Debrecen: Debreceni Nyári Egyetem.

Hunzib

van den Berg, Helma (1995) *A grammar of Hunzib*. München: Lincom Europa.

Ingush

Nichols, Johanna (1994) Ingush. In: Smeets Rieks (ed.). *The Indigenous Languages of the Caucasus* 4, 77-145. Delmar/ New York: Caravan Books.

Guerin, Francoise (2001) *Desciption de l'Ingouche: parler du centre nord du Caucase*. München: Lincom Europa.

Kayardild

Evans, Nicholas (1995) *A grammar of Kayardild, with historical-comparative notes in Tangkic*. Berlin: Mouton de Gruyter.

Ket

Vajda, Edward J. (2004) *Ket.* München: Lincom Europa.

Lak

Zhirkov, Lev I. (1955) *Lakskij jazyk, fonetika i morfologija*. Moscow: Izdatel'stvo Akademii Nauk SSSR.

Friedman, Victor A. (1992) Lak substantival declension: 40 cases or 50? http:// humanities.uchicago.edu/depts/slavic/papers/Friedman-LakCase.pdf.   University of Chicago.

Lezgian

Haspelmath, Martin (1993) *A grammar of Lezgian*. Berlin/ New York: Mouton de Gruyter.

Martuthunira

Dench, Alan Charles (1995) *Martuthunira: A language of the Pilbara Region of Western Australia*, Pacific Linguistics, Series C-125. Canberra. Australian National University.

Mordvin (Erzya)

Keresztes, László (1990) *Chrestomathia Morduinica*. Budapest: Tankönyvkiadó.

Nez Perce

Rude, Noel (1985) Studies in Nez Perce grammar and discourse. Unpublished doctoral dissertation. University of Oregon.

Nunggubuyu

Heath, Jeffrey (1984) *Functional grammar of Nunggubuyu*. Canberra: Australian Institute of Aboriginal Studies.

Pitjantjatjara

Bowe, Heather (1990) *Categories, constituents, and constituent order in Pitjantjatjara, an aboriginal language of Australia*. London: Routledge.

Toda

Emeneau, Murray B. (1984) *Toda grammar and texts*. Philadelphia: American Philosophical Society.

Udmurt

Winkler, Eberhard (2001) *Udmurt*. München: Lincom Europa.

B. In-text references

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

Borin, Lars (1986) Is Hungarian a case language? *Fenno-Ugrica Suecana (Journal of Finno-Ugric Research in Sweden)* 8: 1–33.

Dryer, Matthew S. (2005a) Order of subject, object and verb. In: Haspelmath *et al.* (eds.): feature 81 in *WALS*: 330–333.

Dryer, Matthew S. (2005b) Order of adposition and noun phrase. In: Haspelmath *et al.* (eds.): feature 85 in *WALS*: 346–349

Haspelmath Martin, Matthew S. Dryer, David Gil, and Bernard Comrie (eds.) (2005) *The World Atlas of Language Structures*. Oxford: Oxford University Press.

Huson, Daniel H. and David Bryant (2005) Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution* 23(2): 254–267.

Iggesen, Oliver A. (2005a) Number of cases. In: Haspelmath *et al.* (eds.): feature 49 in *WALS*: 202–205

Iggesen, Oliver A. (2005b) *Case-Asymmetry: A world-wide typological study on lexeme-class-dependent deviations in morphological case inventories*. München: Lincom Europa.

Nichols, Johanna (1992) *Linguistic diversities in space and time*. Chicago: University of Chicago Press.

Abstract

Case markers, as a nominal morphological feature, indicate the functions of NPs in a clause. The number of cases in individual languages is diverse, ranging from no case-marking in Chinese and Arabic to more than 10 cases in Hungarian and Nez Perce. In *The World Atlas of Language Structures* (*WALS*), this morphological feature is highlighted under the heading of "number of cases." Selecting languages with more than 10 cases, I examine what kinds of cases each of them has, and then consider the reasons why these multiple-case languages do indeed have so many cases. Finally, I discuss whether the languages with rich case systems have other common typological features as described in *WALS*. The languages with no case marking are densely distributed in Africa and South Asia, and the languages with more than 10 cases are scattered in Eurasia (Basque, Finnish, and Evenki) and Australia (Gooniyandi, Martuthunira, and Kayardild). Cross-linguistic comparison shows that the languages with rich case systems are in fact rich in locative cases but poor in terms of the variety of cases. Finally, as a result of contrasting case with other *WALS* features, it is established that the languages discussed in this paper have other frequent grammatical tendencies, SOV word order, and postpositions.

《要　旨》

格の多様性について
―言語構造のワールドアトラス(WALS)を使用して―

野　瀬　昌　彦

（東北大学／日本学術振興会特別研究員）

　本論文では，「言語構造のワールドアトラス」（WALS）を使用し，名詞の形態論的特徴である「格の数」を特に取り上げた．「格の数」について WALS を使用して世界中の言語を観察すると，まったく格標示を持たない中国語やアラビア語のような言語から，格が豊富なハンガリー語やネズパース語まで，世界で多様な分布を示す．本研究は，WALS の「格の数」の特徴で観察された，格の数が 10 以上ある 24 言語を対象にした．それらの 24 言語について，いくつの格を持つか，どのような格を持つか，格の数の豊富さと他の WALS の特徴に関連性があるかを検証した．その結果，格を豊富に持つ言語は地域的には欧州からアジア，北南米，オーストラリア，そしてパプアニューギニアと広く分布することが判明した．また，観察される格としては，場所格と副詞格を多く持ち，さらに格を豊富に持つ言語は，SOV 語順が支配的で，後置詞を好む傾向がある．