INFLECTIONAL PARADIGMS

Byron W. Bender
University of Hawai'i at Mānoa

I. INTRODUCTION

Inflection and paradigms go hand in hand. When each member of a given lexical class has different forms with differing grammatical properties, these forms should be capable of presentation in rows and columns in an array that has traditionally been called a paradigm. One such paradigm can be chosen to represent each differing pattern of inflection, so that there will be as many exemplary paradigms as there are patterns to the inflections. Even as simple a system as that of English *count nouns* has its inflectional paradigms:[1]

COMMON NOUNS

	weak nouns				strong nouns		
BOOK:	SG.	PL.		MOUSE:	SG.	PL.	
	book	books			mouse	mice	etc.

Spanish verb inflection presents paradigms of only slightly greater complexity:

PRESENT INDICATIVE

	verbs in *-ar*				verbs in *-ir*		
AMAR: 'love'	SG.	PL.		VIVIR: 'live'	SG.	PL.	
1ST	amo	amamos			vivo	vivimos	
2ND	amas	ama'is			vives	vivi's	etc.
3RD	ama	aman			vive	viven	

Not only is inflection capable of paradigmatic presentation, but there are compelling reasons why it must be so conceptualized if it is to be properly understood. For part of the exercise of placing the inflections of a given noun properly within their paradigm is the identification of the grammatical categories (and the properties within those categories) that should label each row and column of the paradigm. The two English paradigms given above have just one dimension, that of number, while the Spanish paradigms have two, number and person. Fuller Spanish paradigms, of which the above present indicative forms are only portions, will include tense (*imperfect* and *future* vs. *present*) and mood (*subjunctive* vs. *indicative*), and thus have a total of four dimensions. In a paradigm properly formed, there is a one-to-one agreement between grammatical categories and dimensions, with the properties of each category displayed along a separate dimension. This, in our view, forms part of the definition of inflection: whether inflected forms can be fit into a paradigm with a separate grammatical category identified for

---

[1]For purposes of this discussion, we ignore the question as to whether these English paradigms should have two additional cells for the genitive singular and plural.

each dimension. We have not completed the determination of inflection until this task has been fully accomplished.

A. CITATION FORMS.

Several other points can be noted concerning the two simple inflectional paradigms above. One form of each word whose paradigm is displayed is chosen to represent that word as lexeme. For English nouns it is the singular form from within the paradigm that serves as label of the lexical item displayed. For Spanish verbs it is the infinitive form, a form that happens not to occur in the paradigms at hand. These are the forms we would expect to find as headwords for each of these lexemes in a dictionary. Which form is chosen to serve as headword or citation form is primarily a matter of convention, although where the markedness of the forms is obvious we expect to find a form that is less marked chosen, rather than one that is more marked. (See the discussion of markedness in the next section of this paper.) For example, with English nouns we expect to see the form not marked by the *-s* serve as citation form, rather than the form with the *-s* (or its analog, for words like mouse).

B. GRAMMATICAL PROPERTIES.

Each of the forms within an inflectional paradigm is more than just a word-form. As the occupant of a cell within a paradigm it is the bearer of all the grammatical properties that label the cell, and subject to any developments that may be occasioned by them. For example, **mice** is more than m-i-c-e or the sequence of phonemes /mays/. It is also the *plural* of **mouse**, and in the future if it should happen to change to **mouses**, it would not be because of the operation of regular sound changes on the sequence /mays/, but because it was reassigned to the *plural* cell within the paradigm of **book** instead of its present paradigm. Similarly, **amamos** is more than a string of six Spanish phonemes; it is the *1st person present indicative* form of **amar**.

C. ARRANGEMENTS.

The arrangements of the cells in these paradigms is arbitrary and determined in large part by convention. These conventions may or may not be well motivated, and are worthy of reexamination to ensure that they are optimal for the purposes of the description of which they are a part. As exemplified in the above paradigms, Western grammatical tradition generally presents singular forms before plural forms, either above them, or to their left. Is there any motivation for this other than convention? On the other hand, the ordering of the persons in the Spanish paradigms would seem to proceed naturally from their labels. But is there anything other than convention (or egocenteredness!) to support the choice of these particular labels for the grammatical properties involved? One of the techniques we recommend for determining the optimal arrangements of paradigms is to suspend for a moment the traditional labels and the places usually accorded them, and to permute the rows and columns of the paradigm in other possible arrangements while focusing on the forms that fill the cells.[2] This will be the subject of a later section of this paper.

D. LABELS.

Labels and the interrelations of the properties they represent may be either equipollent (symmetrical and equal in significance) or privative (wherein the meaning of one term is

---

[2]This is a technique to our knowledge first recommended explicitly by Kenneth Pike in Theoretical Implications of Matrix Permutation in Fore (New Guinea) *Anthropological Linguistics* 5 (8):1-23, November 1963.

altered from positive to negative in another).[3] The labels *singular* and *plural* are equipollent, whereas, for example, the substitution of *plural* and *nonplural*, or conversely *singular* and *nonsingular*, would be privative. The traditional person labels are also equipollent; a privative alternative sometimes proposed is the following, where + and - are used to indicate whether a term is positively marked or is unmarked with respect to a particular semantic feature:

	EQUIPOLLENT	PRIVATIVE
FEATURES:	'INCLUDES SPEAKER'	'INCLUDES ADDRESSEE'
1ST PERSON	+	-
2ND PERSON	-	+
3RD PERSON	-	-

Questions that arise concerning the labeling of grammatical properties include not only the choice between equipollent and privative, but among the privative choices, which term should be viewed as positively marked, and which should be viewed as unmarked. (The latter decision will affect the name chosen for the privative term, *nonsingular* vs. *nonplural*, for example.) More importantly, what are the criteria, if any, that can be used to answer such questions in a principled way. Before proceeding further in a general discussion of inflectional paradigms, let us first turn to a brief review of markedness theory as it applies to grammatical and semantic categories.

II. MARKEDNESS IN INFLECTIONAL PARADIGMS.

Inflectional paradigms provide fertile ground for testing whether the hypothesis of marked and unmarked categories is revealing of universal tendencies in language phenomena.[4] As we examine inflection from language to language we will want to ask ourselves whether and to what extent the hypothesis is applicable and revealing of such general tendencies. The inflectional entities that we will be designating as marked and unmarked are the members of each grammatical category--the grammatical properties within each grammatical category rather than the categories themselves. For example, we will be designating properties such as *singular*, *plural*, *dual*, etc. within the category of number rather than the category of number as such.[5]

---

[3]P. H. Matthews, *Morphology: an introduction to the theory of word structure,* Cambridge University Press, 1974, pages 43, 150-53.

[4]See especially the review article "Language Universals" prepared by Joseph H. Greenberg for the Linguistic Institute of the Linguistic Society of America held at Indiana University in 1964 and published in *Current Trends in Linguistics 3: Theoretical Foundations*, Thomas A. Sebeok, ed., The Hague: Mouton, 1966, pages 61-112, which summarizes and relates in a general framework the work of linguists such as Trubetzkoy in phonology, Jakobsen and Hjelmslev in grammatical categories, and some preliminary work of his own on kinship terminology.

[5]The term "category" has a usage in the literature on markedness theory that is at variance with its usage in the term "grammatical category." To avoid confusion, we will speak of marked and unmarked "members" or "terms" rather than marked and unmarked "categories." Thus we will refer to the labels for the properties of each of the rows and columns of an inflectional paradigm as "members" or "terms," and reserve "category" for its traditional usage as in "grammatical category." The latter corresponds to the dimensions of a paradigm--the dimensions of its rows, its columns, or (for paradigms of more than two dimensions) its strata, slopes, or clines--to mention a few of the analogs which may prove helpful in

A. MEANING CHARACTERISTICS

According to the hypothesis, marked and unmarked terms are distinguished by characteristics of both form and meaning. Characteristics of meaning include the following:

1. Facultative expression.

A characteristic of meaning or usage, this is the name given to a pervasive tendency in the nature of human thinking to treat one member of an oppositional category (the unmarked member) ambiguously as either the representative of the entire category, or (in its *par excellence* interpretation) as the opposite of the marked member.

Thus, in Jakobson's terms, 'woman' states the presence of the marked category, 'feminine', while 'man' is used chiefly but not exclusively to indicate the absence of 'feminine.' 'Man' thus has two meanings, to indicate the explicit absence of 'feminine' in the meaning 'male human being' but also to indicate 'human being' in general. It is this ambiguity of 'man' which is exploited by Shakespeare when he says, "No! Man delights not me, nor woman neither, though by your smiling you seem to say so." In the earlier part of his speech Hamlet had described his reactions to nature so that the opposition first in mind is nature vs. man, but immediately the possibility of the opposition man vs. woman also occurs to Hamlet.[6]

An example from inflection is the English present tense. The simple past tense is usually considered the marked member in opposition to the present or nonpast, whose ambiguity as unmarked member can be seen in its use in the "historical present" of narratives

> Behind the arras, hearing something stirre,
> He whips his rapier out, and cries: a rat, a rat,
> And in his brainish apprehension killes
> The unseene good old man....[7]

where it serves in lieu of the marked member as part of its timeless sense, in contrast with examples such as the following where it is definitely *(par excellence)* nonpast in sense:

> Here comes the bride.

> We are gathered together ....

> I hereby pronounce you man and wife.

2. Dominance.

This characteristic, closely related to that of facultative expression, refers to instances in which one member of a heterogeneous collection is chosen to represent the entire collection, as in Sanskrit **ahani** DAY AND NIGHT (literally THE DAYS (DUAL)), or Spanish **los padres** THE PARENTS (literally THE FATHERS). An inflectional example is Spanish

> El hijo y la hija son buenos.
> The son and the daughter are good.

---

visualizing multidimensional paradigms. Properties such as *singular* and *plural*, as members of the category number, will label the rows or columns of a dimension devoted to number.

[6]Greenberg 1966, page 72.

[7]As quoted by Otto Jespersen (*A Modern English Grammar on Historical Principles*, Part IV, Syntax (Third Volume), London: Bradford & Dickens, 1931 (reprinted 1954)) from Shakespeare's Hamlet (IV. 1.9).

in which the adjective **buenos** GOOD agrees in gender with **hijo** SON rather than **hija** DAUGHTER.[8] The term that appears in such instances where one or the other must be chosen--in this last example the masculine gender--is the unmarked term. Similarly, that we say in English

> a ten-gallon hat
> a five-gallon can
> a two-dollar bill

is evidence that singular (nonplural) is the unmarked term with respect to plural.

B. CHARACTERISTICS OF FORM

1. Zero expression.

The unmarked member may be unmarked in form as well as in meaning—that is, marked by zero. Since *plural* is generally considered to be the marked term in comparison with singular, it is not surprising that the *singular*, *nonplural* form of English nouns should have zero expression. Conversely, it sometimes happens that marked forms will have more phonetic substance than their unmarked counterparts, that they will be lengthier due to the space within the form occupied by the overt form of the mark. It would certainly go against expectations if, other things being equal, a set of marked forms were each shorter than their unmarked counterparts. As an example, compare the *singular* and *plural* suffixes for person and number of Spanish verbs in the *present indicative* (where -# stands for a zero suffix):

	NONPLURAL (unmarked)	PLURAL (marked)
1ST	-o	-mos
2ND	-s	-'tis
3RD	-#	-n

Not only is each *nonplural* form shorter than its *plural* counterpart, but the *3rd person* forms in each column are briefer than the *1st* or *2nd person* forms, the sort of correlation to be expected if 3rd person is less marked than either *1st* or *2nd*. These forms thus support the privative analysis given for the traditional three persons under the discussion of labels above.

2. Syncretism.

Syncretisms provide us with a double set of information. First of all, they tell us something about the term within which they occur--namely, that it is marked. Distinctions existing in the unmarked member are often neutralized (syncretized) in the marked member. Good examples are to be found in Latin case inflections, for example, those of *3rd Declension masculine* and *feminine* nouns.

	NONPLURAL	PLURAL
NOM	-s	-e:s
ACC	-em	-e:s
DAT	-i:	-ibus
ABL	-e	-ibus
GEN	-is	-um

---

[8] The Arabic and Spanish examples are from Greenberg 1966, page 75.

Two instances of syncretism are enclosed in boxes. The distinction between the *nominative* and *accusative* cases made in the *nonplural* (unmarked) member is neutralized in the *plural* (marked) member. *Dative* and *ablative* undergo parallel syncretism in the *plural*. Each of these instances of syncretism is evidence that *plural* is the marked member, the member in which other distinctions are neutralized.

Secondly, syncretisms tell us something about the terms between which they occur--namely, that they are closely related and minimally distinguished. The Latin syncretisms above are evidence that the distinctions between *nominative* and *accusative* on the one hand, and *dative* and *ablative* on the other, are minimal, probably differing from each other by only one feature.

Interestingly, almost any resemblances between forms in a paradigm can be seen as partial syncretisms, with the same implications for their relative markedness.

3. Deviance.

A closely related formal point is that marked terms have a lesser degree of morphological irregularity. If we look at the dative and ablative endings of all five Latin declensions, the following picture emerges:

Declension:	1ST	2ND	-----3RD-----		-----4TH-----		-----5TH-----	
Nmbr. **Case**								
SG DAT	-ae	-o:		-i:	-ui:	-u:	-e:i:,	-ei:
ABL	-a:	-o:	-e,	-i:		-u:	-e:	
PL DAT	-i:s	-i:s	-ibus		-ibus,	-ubus	-e:bus	
ABL	-i:s	-i:s	-ibus		-ibus,	-ubus	-e:bus	

Several observations can be made. While the *dative* and *ablative* are everywhere syncretized in the *plural*, in the *singular* they are fully syncretized only in the *2nd declension*, and partially in *3rd* and *4th*. More to the point with respect to deviance, no two declensions mark either their *datives* or *ablatives* identically in the *singular*, while in the *plural 1st* and *2nd* do, and *3rd* and *4th* do partially. Further, the *3rd* through the *5th* have the common element **bus** in the *plural*. The greater deviance of the *singular*, manifest in all these ways, is evidence of its unmarked nature in comparison with the *plural*.

6. Frequency.

The unmarked member is characteristically more frequent in running text. This follows at least in part from the dual role it plays in both facultative expression and dominance, although it may also result from an economy of language design and/or an economy of effort in language usewhereby more frequently recurring entities are given simpler marking. Note that it is the frequency of language use as recorded in texts that is being referred to, not that of grammatical lists or the lexicon.

III. PARADIGM PERMUTATION

As noted in our earlier discussion of arrangements and labels, there is merit in temporarily suspending traditional arrangements and labels and independently permutating the rows and columns while focusing only on the forms themselves in an effort to find an optimal

arrangement. If that arrangement should turn out to be the original one, its optimality will at least have been confirmed.

A. OPTIMAL ARRANGEMENT OF ROWS AND COLUMNS IN A PARADIGM

1. juxtapose instances of syncretism

2. juxtapose similarities of form

Similarities of form that correlate with similarities of meaning should be juxtaposed in such a way as to facilitate their highlighting.

3. juxtapose common elements of meaning

Common elements of meaning should be juxtaposed in logical order in those instances where similarities of form do not dictate otherwise, in order to suggest and permit insights into markedness relations and possibly lead to privative labeling of rows and columns.

Following is a paradigm of only moderate complexity from Arabic verb inflection that can serve to illustrate some of these points.[9] Traditional equipollent labels are retained for the time being. Note first of all the two enclosed syncretisms, that of *dual* and *plural* in the *1st person*, and that of *masculine* and *feminine* in the *2nd person dual*. This arrangement also focuses on any correlations of form with person, with a column devoted to each of the three persons in the knowledge that vertical alignments are more salient visually than horizontal ones. Thus it is easy to see that *1st* and *2nd person* forms have an initial dental consonant, and that *3rd* person forms have no initial consonant.[10] This latter lack of phonetic substance has been made overt by use of a null symbol, and the other person correlations in initial consonant have been put into bold face for highlighting. Also to be noted along the vertical axis is an **-um-** formative that correlates with *2nd person* except in the *singular*. We can conclude that these data support the hypothesis that *3rd* is the least marked of the persons.

	1ST	2ND	3RD	
SINGULAR	-**t**u	-**t**a	-#a	MASCULINE
		-**t**i	-#at	FEMININE
DUAL	-**n**â	-**t**umâ	-#__â	MASCULINE
		-**t**umâ	-#atâ	FEMININE
PLURAL	-**n**â	-**t**um	-#û	MASCULINE
		-**t**umna	-#_na	FEMININE

(mn > nn)

Here follows another arrangement which uses privative labels and puts each of the numbers in a column, with *dual* to the right of *plural* as a specially marked *plural.* The consistent mark of

---

[9]These data are taken from Bruce L. Pearson's *Workbook in Linguistic Concepts* (New York: Knopf, 1977), Ex-8.10a, page 62, where 15 inflected forms are given for each of two verbs, as for example:

ðahabtumâ	YOU TWO *(masc.)*WENT	ðarastumâ	YOU TWO *(masc.)* STUDIED
ðahabtumâ	YOU TWO *(fem.)* WENT	ðarastumâ	YOU TWO *(fem.)* STUDIED
ðahabtunna	YOU *(fem. pl.)* WENT	ðarastunna	YOU *(fem. pl.)* STUDIED

[10]This statement holds only if the **n** of the *3rd feminine plural* is interpreted as being medial, which finds some justification in the parallelism between the **na** formatives of the *plural feminine* forms.

7

the *dual* is the long **a** (circumflex **a**), which has been extended to the *plural* in the *1st person* through syncretism. This extension is counter to the principle of dominance, which leads us to expect the unmmarked form--here it should have been the *plural* form--to appear in such instances, where one of two possibilities is chosen.[11]

Everything else about the syncretisms supports our marking hypotheses, however. The fact that the one syncretism takes place within the *1st person* is evidence that *1st person* is a marked member (with respect to the feature 'speaker involved'). That it takes place between *dual* and *plural* is evidence that they are separated by just one feature--plus or minus *dual*. That the syncretism in the *2nd person* between *feminine* and *masculine* takes place within the *dual* is evidence that *dual* is the marked member of its opposition with *plural*. That it is between *feminine* and *masculine* confirms that they are minimally separated--plus or minus *feminine*.

			PLURAL:	–	+	+
			DUAL:	–	–	+
SPKR	ADDR	FEM				
+	–	–	-tu	-n__**â**	-n__**â**
–	+	+	-ti	-tunna	-tum**â**
–	+	–	-ta	-tum	-tum**â**
–	–	+	-#at	-#__na	-#at**â**
–	–	–	-#a	-#û	-# **â**

The other formative correlating with number is -na- (highlighted as -NA-), marking *plural* in the *feminine*. When we search for formal correlations with gender, as in the array below, this same formative is the only one to be found, from this point of view marking *feminine* in the *plural*.

			MASCULINE	FEMININE
1ST	SG		-tu	-tu
	DL		-n__**â**	-n___**â**
	PL		-n__**â**	-n___**â**
2ND	SG		-ta	-ti
	DL		-tum_**â**	-tum__**â**
	PL		-tum	-tunna
3RD	SG		-#a	-#at
	DL		-#___**â**	-#at__**â**
	PL		-#û	-#__na

[11]Yet there seems to be a tendency among languages in this direction. See, for example, the use in Polynesian languages of earlier trial forms as duals. This matter needs more study.

In this array we have repeated the *1st person* forms to emphasize something that has been suppressed until now--the complete syncretism of *masculine* and *feminine* in the *1st person*--further evidence of its marked nature as over against *3rd*.[12]

Although this by no means exhausts the permutations of this paradigm we could engage in, it should be sufficient to give an idea of the potential of this approach and to illustrate some of its techniques. We have employed a number of devices which can be used to discover as well as to display form-meaning correlations in paradigms.

B. HIGHLIGHTING

Most of the devices can be seen as means of highlighting (or alternatively, suppressing) forms or parts of forms.

1. Use of overt markers for zero elements.

2. Enclosure of similar or identical forms within boxes.

Our usage has been to reserve this device for full syncretisms. However, there is no reason why it could not also be used for portions of forms that are identical.

3. Substitution of typefaces.

Boldface, capitals, and other more prominent typefaces can be used to highlight resemblances, and those less prominent to suppress extraneous material.

4. Vertical alignment.

Where different portions of a form are simultaneously aligned in different ways with material above or below, thereby creating gaps, the parts on either side of the gap can be tied together with underlining as a reminder of the form's unity.

5. Complete suppression.

This has been applied to the stems of all Arabic words in the foregoing arrays, as an aid to focusing in on the relevant portions of each form. However, there is a danger that in doing this exclusively we may lose sight of the fact that in word-and-paradigm theory inflectional paradigms consist of word-forms, not just affixes. Languages with concomitant stem variation may help remind us of the need to include and deal ultimately with entire word-forms.

A remaining question concerns which array we might choose for some general purpose such as a descriptive grammar--an array that would incorporate as many as possible of the insights we have gained from our investigations of this Arabic paradigm without overdoing it and defeating the purpose of our description. We have given the Arabic data in three different arrays, one each with the categories of person, number, and gender placed in turn along the horizontal dimension, to test each with the advantage of vertical alignment. Which category has the most formal correlates to reveal in using this advantage? If there is a clear answer to this question, it should be capitalized upon in our final display. Clearly gender in the third array has the least to be revealed. It is less clear whether person in the first array or number in the second has the most. In order to build as much as possible on the past experience and the expectations of our readers, we would probably tip the balance in favor of the second array on

---

[12]This syncretism is also suppressed in Pearson's original problem, where two separate but identical forms, differing only in their glosses (*masc.* vs. *fem.*) are not given in the *1st person* as they were in the *2nd* (see Note 9 for examples of how parallel *2nd person* forms are treated).

the basis of tradition, which usually puts person on the vertical dimension and number on the horizontal.

Arabic Verb Inflection

		--------------------NUMBER--------------------		
PERSONS INVOLVED:		NONPLURAL	DUAL	PLURAL
SPEAKER		-tu	-n__**â**	
ADDRESSEE	(FEMININE)	-ti	-tum**â**	-tun**na**
	(NONFEMININE)	-ta		-tum
NEITHER	(FEMININE)	-#at	-#at**â**	-#**__na**
	(NONFEMININE)	-#a	-#__**â**	-#û

The labels chosen here are meant to exemplify a middle ground between the traditional and the purely privative, replete with pluses and minuses. The use of boxes for each cell permits a slightly different treatment of syncretism, with only one copy of the form filling the center of such enlarged cells. Although the term *singular* is less clumsy than *nonplural*, and *2nd masculine* than *addressee (nonfeminine)*, there may be value in reminding ourselves of the facultative nature of such unmarked terms. Ultimately, most of the decisions represented in this final array are debatable, and largely a matter of taste.