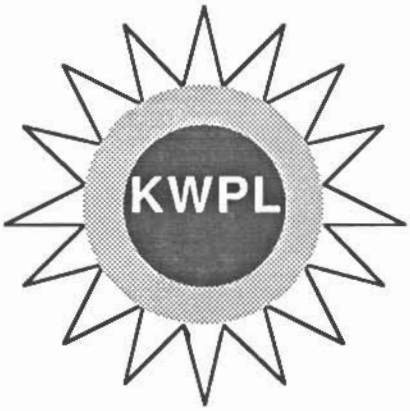
Kansas Working Papers in Linguistics

edited by

Shehdeh Fareh
Yukihiro Yumitani



Partial funding for this journal is provided by the Graduate Student Council from the Student Activity Fee.

Volume 12 1987

Kansas Working Papers in Linguistics

Volume 12, 1987

Jill Brody	Particles in Tojolabal Mayan Discourse	1
Willem J. de Reuse	One Hundred Years of Lakota Linguistics (1887-1987)	13
Antonia Y. Folarin	Lexical and Phrasal Phonology of Yoruba Nouns	43
Noel Rude	Some Klamath-Sahaptian Grammatical Correspondences	67
David Leedom Shaul	A Note on Hopi Consonant Gradation	85
Karen Sundberg	Word Order in Klamath	89
Yukihiro Yumitani	A Comparative Sketch of Pueblo Languages: Phonology	119
Cumulative Contents of Volumes 1-12		141

WORD ORDER IN KLAMATH

Karen Sundberg

Abstract: Klamath, a Native American language of Oregon, has particularly "free" word order. Methods developed by Givón (1983) for the measurement of topic continuity, when applied to Klamath narrative texts, show that the position of NP's relative to the verb is not random, but determined by discourse-pragmatic factors.

Word order in Klamath, a Penutian language of southern Oregon, has been described as "almost completely free" (Barker 1964:341). Barker demonstrates that major arguments within the clause can occur in all logically possible combinations. This paper will demonstrate, however, that in discourse the occurrence of NP's in pre- or post-verbal position is not random, but is heavily influenced by pragmatic factors.[1] This paper examines the effect of the relative topicality of arguments on their position preceding or following the verb. The degree of topicality is determined using five measures of topic continuity developed by Givón (1983) and one additional measure.

The data base consists of seven Klamath texts from Barker 1963: five traditional myths and two ethnographic and procedural texts,[2] altogether comprising 761 usable clauses. In this study, the position of subject, object, and locative noun phrases are analyzed as to the degree of topic continuity encoded in relation to the verb. Unstressed pronouns, though restricted to clause second position, are also analyzed for their topic continuity properties in relation to other categories. The prediction, based on the results reported in Givón 1983, that post-verbal position tends to code more highly continuous topics, and pre-verbal position more discontinuous topics, is confirmed for all noun phrase categories in Klamath. All coding categories, in turn, are seen to rank on a continuum from most continuous, the unstressed pronouns, to the least continuous, the locative NP's, as predicted by the topicality hierarchy proposed by Givón.

Kansas Working Papers in Linguistics. Vol. 12, 1987. pp. 89-118

We will first briefly outline the quantitative methods for assessing topicality used in this study (for a more complete exposition of the methodology and the assumptions underlying and deriving from it, see Givon 1983; there can also be found the previous work producing the conclusions on which the predictions assumed in this paper are based). Succeeding sections will then describe the Klamath constructions which were investigated, the numerical results, and the conclusions which can be drawn about the function of word order in Klamath.

METHODOLOGY

The methodology devised by Givón (1983) for assessing relative topicality of NP's in text involves counting the number of clauses intervening between the NP and the last previous mention and next following mention of its referent in the text, and other measurements which will be described below. Five measurements described in Givón 1983 were applied to Klamath texts: the three anaphoric counts of referential distance, potential referential interference, and same vs. different subject; and two cataphoric counts of persistence and contiguity to major thematic breaks. In addition to these, a count of the number of new referents introduced into the text was made for each noun phrase category.

The texts were first divided into individual constituent clauses, each containing a single finite verb. Relative clauses and tightly-bound complement clauses (i.e. those with subject coreferential to that of the main verb) were not counted as separate clauses; e.g. (1) is counted as a single clause:

(1) coy solwo:lgis san'a:Wawli ?at,
and gather together clothlike obj. want now,

na:nok wic'o:L-as
all fishnet-OBJ

'Then [they] all wanted to gather together the fishnet.' (26:47)[3]

Clauses repeated for emphasis or other stylistic reasons, as (2), were not counted as separate clauses:

(2) coy sa honk ?at gena, gena, doscn'a, doscn'a
and they now go, go, run, run
'Then they went and went, ran and ran.'

Elements within quotations were not measured for topic continuity, but the quotative margin and quoted portion were counted together as a single clause. If no quotative margin was present, the quoted portion counted as an instance of subject zero-anaphora (see below) for the speaker.

Occasional interruptions and background interpolations by the narrator were deleted for purposes of the study, and resumption of the narrative counted as a thematic break except when the interruption occurred clause-medially, as in:

(3) coy honk[4] cew -- dam mat cew q'ay ni
and Antelope -- INT EVID Antelope NEG I

s?aywakta -- kani ?a gatba
know -- someone DEC arrive

'And then Antelope -- did they say Antelope?
I don't know -- someone arrived.' (3:109)

Thus counted the data base totalled 761 clauses.

The measurements which were applied to the texts are as follows:

Referential Distance (RD): To determine the referential distance (RD) between a particular token and its last reference in the preceding discourse, the number of clauses between the token and the last clause with the same referent is counted. The minimum possible value of 1 shows a referent to be maximally topical, while an arbitrary maximum value of 20 clauses shows a referent to be maximally discontinuous. Absence of a referent within direct quotation was not counted as a gap, but occurrence within quoted material was counted as occurrence. Indefinites were not subject to this measurement, as they by definition represent the first occurrence of a referent in the text.

Potential Referential Interference (PRI): The number of referents occurring in the preceding 3 clauses which are semantically compatible with the predication of the referent for which counts are being made is taken as a measure of the potential ambiguity of the referent and the effect it has on the choice of coding strategy. A value of 0 the indicates no potential for ambiguity; a value of 1-3 indicates high potential for ambiguity. (This measure likewise cannot be applied to indefinites).

Persistence: As a measure of importance in the discourse, the number of occurrences of the referent was counted in the 10 clauses following that in which the token occurs. For subjects, only subsequent occurrences as subject were counted; for objects and locatives, occurrence in any role was counted. The minimum possible value of 0 indicates low discourse importance (except for paragraph-final occurrences); the maximum value of 10 indicates a high discourse importance of the referent. Occurrences within direct quotation were counted, while absence was not counted as a gap. As this measurement is cataphoric, it can be applied to indefinites as well as definites.

Same-Subject (SS) vs. Different-Subject (DS): To determine the switch-reference functions of SV and VS patterns, the subject of the preceding clause was determined to be the same or different. And example of a SS pattern occurs in the second clause of (4); an example of DS in the second clause of (5):

(4) coy detdeye:mi-pk-s papakpakkanga,
and (DIST)hungry-DUR-NOM (DIST)(INT)bark around

hok-t w'awk'a
that-REF little coyotes

'Then, being hungry, those little coyotes barked around here and there.' (3:106)

(5) ... hontba, hehji:k'a-lam-ksi.
land, Hehji:k'a-GEN-place.

hehji:k'a cawaltk mna wqepl'aqs-dat
Hehji:k'a sitting on top his summer-house-LOC

'... [it] landed at Hehji:k'a's place. Hehji:k'a had been sitting on top of his summer house.' (8:2-3)

Contiguity to Major Thematic Breaks: The paragraph boundaries indicated by Barker in the texts were taken as major thematic breaks for purposes of this measurement, in addition to breaks occasioned by interpolated comments by the narrator. If a referent occurred in the initial clause of a thematic unit, it was counted as paragraph initial. All other occurrences were classified as paragraph non-initial. However, if two referents were introduced in separate clauses at the beginning of a unit, both occurrences were counted as paragraph initial, as in the beginning of the story of Old Bear and Antelope:

(6) wit'e:m la:b-a wewe:?as gitk.
Black Bear two-OBJ children have

cew l:ab-a wewe:?as gitk.
Antelope two-OBJ children have

'Black Bear had two children. Antelope had two children.' (1:1-2)

Introduction of New Referents: This was a count to determine for each initial occurrence of a referent the particular order in which it was introduced. This is another type of discontinuity which overlaps somewhat with the contiguity at thematic junctures and with the referential distance. This count was applied to all noun phrase categories.

OBJECTS OF INVESTIGATION

The grammatical devices investigated in this study are unstressed third person pronouns, subject and object noun phrases, both definite and referential-indefinite, in pre- and post-verbal positions; and locative noun phrases both preceding and following the verb. As the category of zero anaphora was required in the methodology, a brief description of the phenomenon in Klamath follows.

Zero Anaphora: Zero anaphora is the most frequently used subject-coding device in Klamath, and is also used for objects, though less often. The following exemplifies both subject and object zero anaphora:

(7) coy sl'o:los-dat pnipno:Goga
and elder twig-LOC (DIST) blow into a container
'Then [he] blew [them] into an elder twig.'
(1:141)

Anaphoric zeros were not measured for topic continuity but were counted as referential in determining the continuity of the categories which were measured.

Unstressed Pronouns: Unlike most languages with free word order, Klamath has no obligatory agreement system. Pronouns, when they do occur, occur in clause-second position. The following examples illustrate subject and object pronouns:

(8) coy sa na:nok waytas ge:s?alca
and 3pl every day go to gather ipos
'And every day they went to gather ipos.' (1:3)

(9) coy honk sa-s dom p'as se:w'a
and 3pl-OBJ much food think
'And [she] thought they ate a lot.' (1:49)

Klamath has no noncontrastive third person pronoun as such; instead demonstratives are used with third person anaphoric reference. (No examples of the contrastive third person pronoun occurred in the seven texts examined). In order to distinguish pronominal from other functions of the demonstratives, only demonstratives marked with the "referential" suffix -t were counted as tokens; an example of a third person pronominal object is:

(10) coy honk honk-t poLq'ank,
and that-REF having plucked,

?at no:ql'ank, coy sa p'an
now having roasted, and 3pl eat

'And having plucked it, then having roasted [it], then they ate.' (8:27)

Subject Noun Phrases

Definite Subjects: Klamath also lacks an elaborate case-marking system for nouns. There is no morphological subject marking, and a single object marker which marks datives obligatorily and transitive patients optionally. (Factors governing object marking will be discussed below). Subject NP's may precede or follow the verb:

(11) hehji:k'a cawaltk mna wqepl'aqs-dat
Hehji:k'a sitting on top his summer-house-LOC
'Hehji:k'a had been sitting on top of his summer-house ...' (8:3)

(12) coy be:n hak hottGic'apga hehji:k'a
and again EMPH come running down off Hehji:k'a
'And again Hehji:k'a came running down.'(8:78)

Often subject NP's are preceded and/or followed by a demonstrative or personal pronoun, which may modify the noun or count as a double occurrence. If these occurred on the same side of the verb, as in (13), they were counted in this study as a single occurrence:

(13) coy honk hok-t c'asga:y c'igatk-damna ?ambo
and that-REF Weasel fetch liquid-HAB water
'And that Weasel used to go fetch water ...'
(10:8)

However, if the noun and the coreferential demonstrative occurred on opposite sides of the verb, as in (14), neither was counted, as the position of the NP relative to the verb cannot be determined as one or the other:

(14) ma:ns hok q'e:gi, hok w'ak'a
long that be absent, that Little Coyote
'And he was gone a long time, that Little Coyote.' (3:92)

The small number of occurrences of this latter pattern in the texts (a total of 4) did not warrant the creation of a separate category.

Referential-Indefinite Subjects: Referential indefinite noun phrases include first mention of inanimates, as well as occasional animates which were either preceded by a quantifier, as la:ba wewe:?as in (15), or were not proper names of well-known mythical characters, as Ga:q in (16):

(15) wit'e:m la:b-a wewe:?as gitk
Black Bear two-OBJ children have
'Black Bear had two children.' (1:1)

(16) coy honk Ga:q honk kiko:cn'a
and crow that (DIST) poke w. sharp instr.
'Then a crow speared them with its beak.'(2:27)

All referential-indefinite subjects in the texts precede the verb, as in ex. (16).[5]

Object Noun Phrases

Definite objects: There is only one object marker on nouns in Klamath, occurring always on datives and generally on animate patients of transitive verbs.[6] In this study I have treated all object-marked NP's as belonging to a single category.

Object NP's occur in both pre- (17) and post-verbal (18) position:

(17) coy honk ?at gaba:tis ce:,
and now go to shore upon

sdaynas honk sl'oGi
heart that(OBJ) swallow

'And upon reaching the shore, [he$_i$] swallowed his$_j$ heart.' (3:26)

(18) gmok'am'c sat'wa:ya maqlaqs-as
Gmok'am'c help people-OBJ
'Gmok'am'c helped the Indians.' (26:61)

As is the case with subject noun phrases, so with objects a small number of clauses (4 in the sample) occur in which coreferential elements occur on each side of the verb. Again, these were not counted.

Referential-Indefinite Objects: This category can occur pre-verbally (as in ex. (15)) and, unlike referential-indefinite subjects, also post-verbally, as in:

(19) wit'e:m'am'c c'aGi: Ge:s
Old Bear put handful in mouth ipos
'Old Bear put a handful of ipos into her mouth.'
(1:14)

Locative Noun Phrases: The majority of locative NP's follow the verb, as in:

(20) coy ksembli wqepl'aqs-dal' mna
and take living obj. back summer-house-to his
'Then [he] took [it] back to his summer-house.'
(8:25)

Less often they precede, as in:

(21) coy honk ?at go:s-dat dalmi
and now tree-LOC look up
'Then [he] looked up in the tree.' (3:109)

NUMERICAL RESULTS[7]

Topic Continuity Properties of Subjects

Referential Distance: Table 1 presents the average values of referential distance for the various categories of subject NP:

	N	RD
PRONOUN	132	1.30
VS order	36	3.92
SV order	89	7.82

TABLE 1: Average Referential Distance (in Number of Clauses) for Subjects

The categories rank predictably on a continuum from the most continuous with the lowest referential distance, the PRO category, to the least continuous with the highest average RD value, the pre-verbal full NP subjects. The value for post-verbal full NP subjects ranks approximately halfway between the others, tending to code topics of higher continuity than the SV category, but not so high as the more continuous PRO category.

In Table 2, the variation of tokens within each category is presented:

no. of	PRO		VS		SV	
clauses	N	%	N	%	N	%
1	108	82	18	50	34	38
2	11	08	7	19	9	10
3	5	04	3	08	5	06
4	3	02	1	03	3	03
5	3	02	1	03	1	01
6	1	0.7	--		4	05
7	--		1	03	1	01
8	--		1	03	--	
9	1	0.7	--		1	01
10			--		2	02
11			--		--	
12			--		3	03
13			--		--	
14			--		--	
15			--		--	
16			--		1	01
17			1	03	--	
18			--		--	
19			1	03	--	
20+			2	06	25	30
Totals	132		36		89	

TABLE 2: Percent Distribution of Referential Distance within Subject Categories

82% of all occurrences of PRO subjects are seen to cluster at the 1 clause range; 77% of all VS occurrences in the 1-3 clause range, while the SV category shows clustering at the two extreme ends of the scale, with 64% occurring within 1-7 clauses, and 30% showing maximal RD at the 20+ clause range.

Potential Referential Interference: The average values for the number of potentially interfering referents in the preceding three clauses are recorded in Table 3:

	N	PRI
PRO	132	0.40
VS	36	0.86
SV	89	1.02

TABLE 3: Average Potential Referential Interference for Subjects

These results again follow the predicted pattern, with the PRO category having the least potential for ambiguity, and therefore coding more continuous referents than the VS category, which in turn has fewer interfering referents than the SV category, the SV ordered NP's again coding topics with the lowest degree of continuity.

Table 4 presents the distribution of potential interference across the categories, i.e. the percentage of the total number of pronominal, post-verbal, and pre-verbal subjects for which there could be found in the preceding 3 clauses 1,2, or 3 distinct referents which could plausibly be interpreted as the subject of the clause under consideration:

PIR	PRO		VS		SV	
	N	%	N	%	N	%
0	80	61	10	28	18	20
1	51	39	21	58	53	60
2	1	0.75	5	14	16	18
3					2	02
4+						

TABLE 4: Percent Distribution of Tokens with 1, 2, and 3 Potentially Interfering Referents (PIR) Preceding Three Clauses

We see that the highest percentage of pronouns (61%) have no interfering referents in the preceding clauses; the majority of full NP subjects in both VS and SV patterns have one interfering referent, with the SV category showing a slightly higher percentage of both 1 and 2 interfering referents than the VS category, as well as having the only tokens with 3 such interfering referents.

Persistence: The results for this measurement are given in Table 5:

	N	Persistence
PRONOUN	132	2.13
VS DEFINITE	36	1.36
SV DEFINITE	89	1.98
VS REF/INDEF	0	N/A
SV REF/INDEF	15	0.53

TABLE 5: Average Persistence (in Number of Clauses) across Subject Categories

As a measure of discourse importance, the categories are ranked differently than for the anaphoric measures; pronouns show the highest persistence, the SV/DEF category persists longer than the VS/DEF category, while the REF-INDEF category (all SV) decays most rapidly in the discourse register. The average values obtained for the definite categories reflect their position within thematic units:

pronouns nearly always occur in paragraph medial and final clauses, and, being the most continuous category, (as reflected in the results for RD and PRI), are most likely to be a major component[8] in long equi-subject clause chains. The SV category, more often found in paragraph initial clauses, is most likely to initiate such chains, the two categories together tending to persist throughout the thematic unit. The VS category encodes less continuous referents than the PRO category, and rarely occurs in paragraph initial clauses (which are more typically SV), and so tends to persist for a shorter duration than either the PRO or SV categories.

All referential-indefinite subject NP's occur in the pre-verbal position, this category having by far the lowest persistence. This reflects the tendency in Klamath and other languages (cf. Givón 1983) to introduce major topics as definites or objects, and highly unimportant referents, which decay almost immediately, as indefinites.

The distribution within categories is presented in Table 6:

no. of	PRO		VS		SV		SV-REF/INDEF	
clauses	N	%	N	%	N	%	N	%
0	38	29	17	47	30	34	9	60
1	35	27	8	22	22	25	4	27
2	19	14	4	11	11	12	2	13
3	12	09	3	08	8	09		
4	7	05	1	03	8	09		
5	6	05	1	03	2	02		
6	4	03	1	03	2	02		
7	3	02	--		2	02		
8	2	02	--					
9	--		1	03	--			
10+	6	05			4	04		
Totals	132		36		89		15	

TABLE 6: Percent Distribution of Persistence within Subject Categories

At least 70% of the tokens in each category occur in the 0-1 clause range, with the REF/INDEFINITE category having 100% of its tokens within this range; the VS category having 80%, the SV category 71%, and the PRO category 70%, with the largest number of tokens in the 3-10+ clause range.

Same-Subject vs. Different Subject: The relative distribution of SS and DS occurrences for the definite categories is presented in Table 7:

	SS		DS		TOTAL	
	N	%	N	%	N	%
PRO	91	69	41	31	132	100
VS*	19	53	17	47	36	100
SV*	21	24	68	76	89	100
TOTAL	131	51	126	49	257	100

TABLE 7: Relative Distribution of SS vs. DS Occurrences within Subject Categories *Definite NP's only

As expected, the PRO category shows the highest ratio of SS to DS occurrences, (69% vs. 31%), the VS category an intermediate ration (53% vs. 47%), and the SV category the lowest ratio (24% vs. 76%). As yet another measure of topic continuity, the categories rank in the same order as in the RD and PRI measurements, with pronouns showing the highest degree of continuity to the preceding clause, VS ordered NP's the next highest, and SV ordered NP's the lowest degree, i.e. the most highly discontinuous to the preceding clause.

Contiguity to Major Thematic Breaks: Table 8 presents the relative distribution of occurrences at thematic junctures as opposed to thematic continuations for each category:

	BREAK		CONTINUATION		TOTAL	
	N	%	N	%	N	%
PRO	9	07	123	93	132	100
VS-DEF	7	19	29	81	36	100
SV-DEF	48	54	41	46	89	100
VS-R/I	0		0		0	
SV-R/I	1	07	14	93	15	100
TOTAL	65	24	207	76	272	100

TABLE 8: Relative Distribution of Contiguity to Thematic Break vs. Thematic Continuation within Subject Categories

As noted above, the great majority of PRO occurrences (93%) are in paragraph non-initial clauses; to a somewhat lesser extent this distribution is replicated by the VS category, with 81% of its tokens in non-initial clauses. In contrast, over half the SV tokens (54%) are in paragraph initial clauses. Again, the PRO category is shown to be the most continuous, the SV category the least continuous, and the VS category ranking between the two.

Almost all of the referential-indefinite NP's (93%) occur in paragraph non-initial clauses. This correlates with the results obtained for the persistence measurement: as highly unimportant referents, they are unlikely to initiate equi-topic chains; rather, they are satellite topics which are most often introduced in the middle or toward the end of the thematic unit, decaying almost immediately in the discourse.

Topic Continuity Properties of Objects

Referential Distance: Table 9 presents the average values for referential distance for the definite object categories:

	N	RD
PRONOUN	44	1.70
VO order	45	6.04
OV order	30	9.97

TABLE 9: Average Referential Distance (in number of clauses) for Objects

As with the subject categories, these are ranked as expected on a continuum form the most continuous, the PRO category, to the least continuous, the OV category, with the VO category intermediate between the other two. The distribution within categories is given in Table 10:

no. of clauses	PRO N	PRO %	VO N	VO %	OV N	OV %
1	33	75	15	33	10	33
2	5	11	11	24	3	10
3	1	02	1	02	--	
4	2	05	2	05	1	03
5	1	02	3	07	1	03
6	--		1	02	--	
7	--		--		--	
8	2	05	--		--	
9			--		2	07
10			3	07	--	
11			--		--	
12			--		--	
13			1	02	--	
14			--		--	
15			--		--	
16			--		1	03
17			--		--	
18			--		--	
19			--		--	
20+			8	19	12	40
TOTALS	44		45		30	

TABLE 10: Percent Distribution of Referential Distance within Object Categories

75% of all PRO occurrences are within the 1 clause range. Seventy-one percent of the VO occurrences are between 1-5 clauses, with 19% at the 20+ clause range. The OV category, like SV, exhibits clustering at the two extreme ends, with 43% in the 1-2 clause range, and a relatively high 40% at the 20+ clause range, indicating that objects as well as subjects are most likely to be introduced or reintroduced in pre-verbal position, as will be demonstrated further below.

Potential Referential Interference: Tables 11 and 12 present the average values for potential interference and the distribution within the definite object categories:

	N	PIR
PRO	44	0.25
VO	45	0.22
OV	30	0.70

TABLE 11: Average Potential Referential Interference for Objects

PIR	PRO		VO		OV	
	N	%	N	%	N	%
0	34	77	35	78	14	47
1	9	20	10	22	11	37
2	1	02			5	17
3+						

TABLE 12: Percent Distribution of Tokens with 1 or 2 Potentially Interfering Referents in (PIR) Preceding Three Clauses

Both the pronouns and the VO ordered NP's show no interfering referents in the preceding 3 clauses. The average value for the OV category is higher, with 54% of the tokens having 1 or 2 interfering referents in the preceding 3 clauses. The PRO and VO categories, then, tend to code more continuous referents, while the OV category, the most discontinuous, is the preferred order in potentially ambiguous predications.

Persistence: Table 13 presents the average values of persistence for the object categories:

	N	Persistence
VO REF/INDEF	15	0.27
OV REF/INDEF	27	1.22
PRONOUN	44	1.36
OV DEFINITE	45	1.60
VO DEFINITE	30	2.80

TABLE 13: Average Persistence (in Number of Clauses) across Object Categories

The VO-ordered definite NP's show the highest average value; this category thus tends to code both highly continuous and important referents, whereas the OV-ordered definite NP's code discontinuous referents which decay more quickly in the discourse register. The PRO category has a relatively low persistence, as this category is not used to code new topics, and is most often found in paragraph non-initial clauses.

The majority of referential-indefinite objects occur in the OV order, and this category has a lower persistence than the definite categories, indicating the relative unimportance of the topics encoded by them. Lowest of all, however, is the VO ordered referential-indefinite category, which nearly always constitutes unique mention in the discourse. These are extremely unimportant referents, the action of the verb itself usually being the most salient in the predication, thus tending to precede the object, as in ex. (22):

(22) coy honk ?at not'o:t'le:Gi,
and now throw round obj. across,

waGe:nha cacga:lam
maybe pine cone

'And then [he] threw [something] across,
maybe a pine cone.' (1:130)

The distribution within categories is recorded in Table 14:

no. of clauses	VO-R/I N	VO-R/I %	OV-R/I N	OV-R/I %	PRO N	PRO %	OV-DEF N	OV-DEF %	VO-DEF N	VO-DEF %
0	12	80	13	48	20	45	13	43	18	40
1	2	13	4	15	12	27	6	20	5	11
2	1	07	4	15	3	07	4	13	3	07
3			3	11	3	07	1	03	3	07
4			3	11	3	07	3	10	2	04
5					--		2	07	3	07
6					1	02	--		3	07
7					1	02	--		3	07
8					1	02	--		3	07
9							1	03	--	
10+									2	04
Totals	15		27		44		30		45	

TABLE 14: Percent Distribution of Persistence within Object Categories

Seventy-two percent of the PRO category shows a low persistence in the 0-1 clause range; the majority of the OV-DEF category (76%) occurs in the 0-2 clause range; and the VO-DEF category again shows the greatest persistence, with 76% distributed in the 0-5 clause range and the only tokens having the maximum persistence of 10+ clauses. Eighty percent of the VO-REF/INDEF category decays immediately at 0 clauses, and 78% of the OV-REF/INDEF NP's occur in the 0-2 clause range. The REF/INDEF categories also have the least dispersed distribution, with no tokens occurring beyond 2 clauses and 4 clauses, respectively.

Topic Continuity Properties of Locatives

Referential Distance: The results for this measurement are recorded in Tables 15 and 16:

	N	RD
V-LOC NP	30	10.93
LOC NP-V	9	17.55

TABLE 15: Average Referential Distance (in number of clauses) for Locative NP's

no. of clauses	V-LOC NP N	%	LOC NP-V N	%
1	5	17	--	
2	2	07	--	
3	1	03	--	
4	3	10	--	
5	--		--	
6	1	03	--	
7	--		--	
8	2	07	--	
9	2	07	2	22
10	1	03	--	
11	--		--	
12	--		--	
13	--		--	
14	1	03	--	
15	--		--	
16	--		--	
17	--		--	
18	--		--	
19	--		--	
20+	12	40	7	78
TOTALS	30		9	

TABLE 16: Percent Distribution of Referential Distance within Locative NP Categories

The majority of locative NP's are post-verbal, showing lower referential distance. A smaller number of locatives occur in pre-verbally; most of these are highly discontinuous, first-mention referents, as can be seen from their distribution in Table 16. The VO occurrences show a more scattered distribution throughout the 1-20 clause range, indicating the somewhat greater continuity of this category, although the locative NP category as a whole is low in topicality as compared to the subject and object categories.

Persistence: Tables 17 and 18 present the average values for persistence and their distribution within the locative NP categories:

	N	Persistence
V-LOC NP	30	0.67
LOC NP-V	9	0.22

TABLE 17: Average Persistence for Locative NP's

no. of clauses	V-LOC N	V-LOC %	LOC-V N	LOC-V %
0	18	60	7	78
1	9	30	2	22
2	1	03		
3	--			
4	1	03		
5	1	03		
6				
7				
8				
9				
10+				
TOTALS	30		9	

TABLE 18: Percent Distribution of Persistence within Locative NP Categories

The majority of occurrences for both the V-LOC and LOC-V orders are shown to have no persistence; the V-LOC, however, contains 30% which persist for one clause, and a few tokens which persist for 2-5 clauses. Therefore, although the locative NP category in general tends to code relatively unimportant referents, the more important of these occur post-verbally, in the more continuous order.

Introduction of New Referents into the Discourse: Table 19 presents the number of first-mention referents occurring in the post-verbal and pre-verbal order for each NP category:

	V NP N	V NP %	NP V N	NP V %	TOTAL N	TOTAL %
SUBJECT						
DEFINITE NP	1	03	28	97	29	100
REF/INDEF NP	0		11	100	11	100
TOTAL SUBJECT	1	02	39	98	40	100
OBJECT						
DEFINITE NP	7	35	13	65	20	100
REF/INDEF NP	12	32	25	68	37	100
TOTAL OBJECT	19	33	38	67	57	100
LOCATIVE	10	59	7	41	17	100

TABLE 19: Introduction of New Referents into Discourse

Ninety-eight percent of the new subject NP's, both definite and referential-indefinite, are introduced in preverbal order, this being another strong indication of the discontinuity of the referents expressed in that order. For the two object categories, 67% are also introduced in pre-verbal order; again a measure of the discontinuity of that order, though to a lesser extent for objects. For locative NP's, numerically more referents are introduced in post-verbal position (59%); however, as the great majority of all locative NP's occur in this order, this amounts to only 33% of all the post-verbal NP's, whereas the number of first-mention referents amounts to 78% of all pre-verbal NP's. Pre-verbal order, then, is the preferred mode for introducing new referents of all major case roles into the discourse.

Word Order Distribution: The distribution of the NP categories in the pre-verbal and post-verbal orders is summarized in Table 20:

	V NP		NP V		TOTAL	
	N	%	N	%	N	%
SUBJECT						
DEFINITE NP	36	29	89	71	125	100
REF/INDEF NP	0		15	100	15	100
TOTAL SUBJECT	36	26	104	74	140	100
OBJECT						
DEFINITE NP	45	60	30	40	75	100
REF/INDEF NP	15	36	27	64	42	100
TOTAL OBJECT	19	33	38	67	57	100
LOCATIVE	30	77	9	23	39	100

TABLE 20: Distribution of Word Order

Here the distribution across the major cases is surprising with respect to the general tendencies exhibited by the major case-roles in the results above, as will be discussed further in the last section of this paper. The pre-verbal order, demonstrated to be the least continuous both within

and across categories in terms of all 6 measurements, has the highest percentage of subject NP's, the most continuous case-role on the topicality hierarchy, and the lowest percentage of locative NP's, the least continuous case-role on the topicality hierarchy, with the object NP's ranking midway between the two. The reverse is true for the post-verbal order, demonstrated above as the more continuous order. For the subject categories, this skewed distribution results in part from the greater functional load of the SV order: it is used for introducing and re-introducing referents into the texts, in potentially ambiguous predications, following DS clauses, and at the start of thematic units. The VS order, on the other hand, is most often used when none of the above conditions are present, i.e. when the referent is predictable and expected. Furthermore, use of the VS order is complemented by the use of zero-anaphora and unstressed pronouns in coding more continuous referents, thus decreasing its overall frequency of occurrence. There are more discourse environments in which the pre-verbal, SV, order is the preferred order for coding referents, and this accounts for the greater number of occurrences of SV order.

For locative NP's the distribution is explained by the relative continuity within the category. As we have seen, the V-LOC order codes more highly continuous locative NP's than the LOC-V order. In fact, 60% of the occurrences had been previously mentioned, and 40% were seen to persist for at least 1 clause, as compared to only 22% of previously mentioned referents in the LOC-V category, and 22% which showed an persistence whatever. The greater number of post-verbal locative NP's, then, is due to the greater number of continuous referents within that category.

Table 21 shows the distribution of major word order patterns in transitive clauses containing two NP arguments.

WORD ORDER	no. of occurrences
SVO	7
SOV	5
VSO	2
OVS	1
OSV	0
VOS	0

TABLE 21: Distribution of Word Order Patterns

As can easily be seen, this type of clause is extremely rare in Klamath, as NP's are the least frequent coding device for subjects. The small number of occurrences in any of the orders does not warrant establishing for Klamath a basic word order in the sense of Greenberg (1966), as is the case for many other free word order languages (cf. Mithun to appear).

Animacy: Klamath shows the expected correlation between case role and the percentage of animate vs. inanimate topics, as presented in Table 22:

	ANIMATE		INANIMATE		TOTAL	
	N	%	N	%	N	%
SUBJECT	135	96	5	04	140	100
OBJECT	61	52	56	48	117	100
LOCATIVE	1	06	16	94	17	100

TABLE 22: Percent of Animate vs. Inanimate NP's in the Major Case-role Categories

Subjects are overwhelmingly animate, reflecting the topical importance of the referents encoded as subject; objects are roughly equally divided between animate and inanimate, and locatives are nearly always inanimate, indicating the relative unimportance of the topics expressed as locative NP's. As will be discussed below, the results obtained for animacy correlate closely with those obtained for both the Persistence and PRI measurements.

DISCUSSION

RD and topic continuity within and across case-roles: The average values of referential distance within the major case roles are repeated below for convenience:

	SUBJ	OBJ	LOC
PRO	1.30	1.70	---
V-NP	3.92	6.04	10.93
NP-V	7.82	9.97	17.55

Two parallel tendencies in the coding of the topicality of referents are evident in these results: within case roles, the grammatical devices can be ranked from the most to the least continuous as follows:

(23) PRO > V-NP > NP-V

and across case-roles, the following continuum is evident, following the same gradient of topic continuity:

(24) SUBJECT > OBJECT > LOCATIVE

This gradation in terms of topic continuity of both case roles and grammatical categories confirms the predictions resulting from the studies in Givón 1983. Within the case roles, pronouns clearly code the most continuous referents, post-verbal NP's code referents of intermediate continuity, and pre-verbal NP's clearly code the least continuous referents, therefore the most inaccessible to the speaker and hearer. Across case roles, subjects overall code the most continuous topics in terms of referential distance, objects code topics of intermediate continuity, and locatives referents of low continuity in the preceding discourse.

In Tables 23 and 24, the degree of topic continuity encoded by the grammatical devices and the case roles, respectively, is expressed as the percentage of tokens in each category having a referential distance of 1-2 clauses, the range showing maximal continuity of referents. In Table 23, 89% of the PRO category is shown to have a low RD of 1-2 clauses; 53% for the entire post-verbal category, and 44% for the entire pre-verbal category:

	TOTAL SAMPLE		RD of 2.0 or less	
	N	%	N	%
PRO	176	100	157	89
V-NP	111	100	58	52
NP-V	128	100	56	44

TABLE 23: Percentage of Tokens of each Type with RD of 2.0 or Lower

In Table 24, the major case-roles are similarly ranked, with 73% of all subjects having a low RD of 1-2 clauses, 65% of all objects, and only 18% of locatives:

	TOTAL SAMPLE		RD of 2.0 or less	
	N	%	N	%
SUBJ	257	100	187	73
OBJ	119	100	77	65
LOC	39	100	7	18

TABLE 24: Percentage of Tokens in each Case-Role with RD of 2.0 or Lower

PRI, Animacy, and Topic Continuity: The percentage of tokens within the subject- and object-marking categories as a whole having no potential for ambiguity is presented in Table 25:

	TOTAL SAMPLE		PRI of 0.00	
	N	%	N	%
PRO	176	100	114	65
V-NP	81	100	45	56
NP-V	119	100	32	27

TABLE 25: Percentage of Subject or Object Tokens with PRI Value of 0

This distribution reflects the hierarchy in (23) above: the PRO category, having the highest percentage of non-interference, is also the most continuous category with respect to the previous discourse.

The post-verbal category shows a moderate potential for ambiguity, and thus an intermediate degree of continuity, with nearly half of its tokens showing a PRI value of greater than 0, while the pre-verbal category clearly shows the lowest degree of non-interference from competing referents. As the potential for ambiguity within the clause increases, the more likely it is that the referent will be decoded in the pre-verbal order: to increase the ease of recoverability of the referent, it is fronted in the clause, making it more salient in the discourse.

Across the subject and object case roles, the relative values of PRI do not follow the hierarchy in (24) as to RD value: subjects consistently have a higher PRI than do objects. Although seemingly unexpected in terms of the lower RD of the subject categories, this is a direct result of the distribution of animates in the major case roles, a measure of the topical importance of the referents (Givón 1983). As seen in Table 22, 96% of all subject NP's are animate; the potential for ambiguity in the predication is then greater, as ambiguity results from shared semantic compatibility with the verb, a major feature of which is animacy. The object NP's, on the other hand, contain a large portion of inanimates (48%), and these do not compete directly with the majority of the verbal predications. Thus the PRI for objects is naturally lower than for subjects.

Topical importance, animacy, and persistence of referents: As mentioned above, the distribution of animate vs. inanimate referents in the major case roles is a more or less direct measurement of the topical importance encoded by the case roles. From Table 22, the continuum presented in (24), repeated below, for topic continuity can be reproduced with reference to discourse importance, as measured by the percentage of animates within the case roles, with 96% animates for subjects, 52% for objects, and 6% for locatives:

(24) SUBJECT > OBJECT > LOCATIVE

Discourse importance is also measured in terms of the persistence of the referent in the subsequent discourse. Although the results obtained for the subject categories cannot be directly compared with those of the object and locative categories, due to the differences in the methodology, the values obtained for these two cases are shown to follow (24), with object outranking locative

in topical importance. The object categories consistently show higher values of persistence than do the locative categories (Table 13), thus coding more topically important referents.

The two ordering patterns, V-NP and NP-V, discourse importance is scaled as for topic continuity, with V-NP outranking NP-V. The post-verbal order codes the more continuous and important topics than the pre-verbal order.

Of lowest topical importance are the referential-indefinite categories for both subjects and objects. For subjects, these are all in pre-verbal position; within the object categories, the post-verbal order codes extremely unimportant referential-indefinites which die out immediately in the register, while the pre-verbal order codes the relatively more salient referential-indefinite objects.

The subject-marking categories were measured for persistence as subject only. The results obtained correlate to some extent with the positioning of the referent in the thematic unit, though not exactly as predicted in Givón 1983. Pronouns show the highest persistence, though they occur overwhelmingly in paragraph non-initial clauses (93%). As discussed above, this results from their being a major component in equi-subject chains. The post-verbal category shows the lowest persistence, as expected for paragraph non-initial occurrences. The persistence of the pre-verbal category is close to that of the PRO category, a relatively high average value as expected for theme-initial occurrences, the initiators of equi-subject chains.

SS vs. DS, Thematic Continuity, and Continuity of Subjects: The hierarchy of continuity for the grammatical devices expressed in (23) is supported within the subject-marking categories by the additional measures of SS vs. DS and contiguity to thematic junctures. The largest percentage of both SS and thematic non-initial occurrences are in the most continuous category, the unstressed pronouns. The VS-ordered NP's again rank as intermediary, and the least continuous category, the SV-ordered NP's, contain the smallest percentage of both SS and thematic non-initial occurrences; thus occurring more often in the highly discontinuous environments of theme-initial clauses and following DS clauses.

Referential-indefinites, the maximally discontinuous category by definition, has the bulk of its

tokens occurring in paragraph non-initial clauses. As discussed above, this category codes extremely unimportant referents, and for this reason is highly unlikely to occur paragraph-initially at the start of equi-subject chains.

A Direct Measure of Topic Discontinuity: The Entry of Referents into the Register: The pre-verbal order is the highly preferred order for introducing referents across all case roles. The entry of a referent into the register is the most discontinuous function in terms of topic continuity, and this strongly confirms the pre-verbal order as the most discontinuous of all grammatical devices. As can be seen by the total number of new topics for each case role (Table 19), referents are most often introduced into the register as objects (N=57) and least often as locatives (N=17). First mention referents in the subject role (N=40) are intermediate.

Pragmatic Functions of Word Order in Klamath: Word order in Klamath is demonstrated to be pragmatically controlled, with the post-verbal order coding more highly continuous referents, and the pre-verbal order coding more highly discontinuous referents across the major case roles, thus confirming the predictions outlined in Givón 1983. Several functions determine the relative continuity of the two orders: for subjects, the pre-verbal order is used most often for entry and re-entry of referents into the discourse; for potentially ambiguous predications, to aid in the recoverability of the referent by increasing its saliency in the discourse, in switching reference from the preceding clause, and at the sites of thematic junctures. The post-verbal order occurs in conjunction with the pronouns in the absence of the above conditions. For objects and locatives, the pre-verbal order is also used more often for introducing and re-introducing referents, and, for objects, in the presence of interfering referents in the preceding discourse. The post-verbal position for the two non-subject cases is used to code the most continuous referents, as well as the more topically important referents which persist in the subsequent discourse. Referential-indefinites for both subjects and objects tend to precede the verb, unless they are extremely unimportant in the discourse, in which case, for objects, they occur in post-verbal position, typically as an afterthought construction. Word order in Klamath is syntactically free; but as a general

rule, unpredictable, potentially ambiguous information is fronted in the clause, thereby increasing its accessibility in the discourse.

NOTES

1) There are other problems of Klamath word order which will not be dealt with in this paper, most notably the fact that noun modifiers occasionally occur on the opposite side of the verb from their head, as in the following example from Barker 1964:315:

nanqa ?a de:Wi wokas p'aLa-tdat
some INDIC leave wokas tray-LOC
'[He] left some wokas on a tray.'

2) The texts are: 1, Old Bear and Antelope, 2, Coyote and Badger, 3, Little Porcupine and Coyote (all by Mrs. Pansey Ohles), 8, Hehji:k'a (Mr. Grover Pompey), 10, The Crater Lake Myth, 26, Killing Fish (both by Mr. Robert David), and 27, Some Historical Incidents (Mrs. Aggie Butler).

3) Citations are to text number and line in Barker 1963.

4) The honk of coy honk occurring here and in many other examples, although morphologically identical to the object form of the demonstrative, apparently functions as a discourse connective of some sort and will not be glossed.

5) One VS existential-presentative construction was omitted from the counts.

6) See Barker 1964:240. Exceptions to the *pattern of marking all and only animate objects* do occur, but none were found in the texts studied.

7) N in the tables represents the total number of tokens of the type in the data base.

8) Zero anaphora is the other major component.

REFERENCES

Barker, M.A.R. 1963. Klamath texts. University of California Publications in Linguistics 30.

_____. 1964. Klamath grammar. UCPIL 32.

Givón, T. 1979. On understanding grammar. NY: Academic Press.

_____, ed. 1983. Topic continuity in discourse: Quantitative cross-language studies. (Typological Studies in Language 3). Amsterdam: Benjamins.

Greenberg, Joseph. 1966. Some universals of grammar with particular reference to the order of meaningful elements. in J. Greenberg, ed., Universals of language. Cambridge: MIT Press.

Mithun, Marianne. to appear. Is basic word order universal? in R. Tomlin, ed., Grounding and coherence in discourse. (Typological Studies in Language 11). Amsterdam: Benjamins.