FUNCTOR ANALYSIS: A METHOD OF QUANTIFYING FUNCTION WORDS FOR COMPARING AND CLASSIFYING LANGUAGES

R. David Paul Zorc

School of Australian Linguistics
Darwin Community College

It would seem quite reasonable that function words, or functors, should play an important role in subgrouping or classifying languages, because of:
(1) their obvious importance within any given speech variety, (2) their high text frequency, and (3) their tendency towards stability and a low rate of replacement. While functors can be shown to yield qualitative results in both synchronic and diachronic studies, it would also seem advantageous to have a quantitative method for dealing with them. I would like to expound briefly on these three reasons why functors are so important to subgrouping, and suggest a method that can deal with them quantitatively.

Some might wonder why it is necessary to belabor the obvious. All too often comparative studies have centered on phonological or lexical evidence, taking a "shortcut" when there isn't one in comparative linguistics [Teeter (1963:648)], and some have even ignored or dismissed counterevidence presented by grammatical structures.

1. IMPORTANCE OF FUNCTORS.

A language is more readily defined by its grammar than by its lexicon. Let us take the following two examples:

Naka-*pay* na 'akú sa *cashier*, *ásk*-a 'imáw.

I've already *bayád*-ed the *manòg-balígya'*, *kutána* her.

The grammar of the first is Aklanon (Bisayan, Philippine), while the lexicon is English; the grammar of the second is English, while the lexicon is Aklanon.

Even if 50% of the vocabulary of any given language were to go, that language would still be that language (how much of English is still *English*?!); if even 10% of the grammar (or functors) went, one would suspect he was dealing with a pidgin.

There has been much outrage expressed in the Philippines over a speech variety called *"Taglish"* (Tagalog with a heavy English overlay). While the grammar is still very much Tagalog, it is the lexicon that some find objectionable. With all the tinkering that goes on over national languages, it is rare if so much as one functor is involved (conjunctions excepted), while a massive part of the lexicon becomes "purer" or "more international" as the case may be. All this is so because the grammar and functors are "assumed"; the language is still the language--which is why the label *Taglish* was chosen instead of *Englog*. One never

doubts that it is still Tagalog being spoken--just the *kind* of Tagalog.

2. HIGH TEXT FREQUENCY OF FUNCTORS.

Hockett (1958:264-5) lists four types of grammatical forms that may be classed as functors: *substitutes, markers, inflectional affixes,* and *abstract governing derivational affixes.* While distinguishing between contentives and functors may be difficult from language to language, forms that have any of the above four functions or attributes may be usefully classed as functors, especially if they form a *paradigm.* The important thing about functors is that a few hundred of them form the grammatical core of a language, while thousands of contentives make up the lexicon. Perhaps an added advantage of working with functors is the limitation on choice. Selecting basic functors is far less tedious--and argumentative--than selecting basic vocabulary.

If one records any text and collates the transcribed results, chances are that even the most basic lexical items (e.g., *eat, sleep, eye, tongue, full,* etc.) may not occur more than a few times, but the text would be replete with pronouns, deictics, articles or markers, negatives, interrogatives, and the like. It is thus functors, rather than contentives, that form the core or basic vocabulary of any given language.

McFarland (1974:313-9) ranked 150 morphemes that occurred with the highest frequency out of six texts in each of twelve Bikol area dialects. Of these only twelve were strictly lexical: *say, arrive, person, tell, finish, name, happen, house, time, see, good, man/male*; the remaining 138 were functors.

3. STABILITY AND LOW PROBABILITY OF REPLACEMENT OF FUNCTORS.

McFarland has stated that the above two factors (high frequency of occurrence and syntactic importance) "would seem to predict high stability, that is, low probability of replacement, for the functors and other restricted-class morphemes." (1974: 121-2). Teeter has also stressed an important fact: "...[D]irect outside influence on grammars is literally impossible, since... each child constructs his own grammar by extrapolating from the utterances he hears. Words are learned, but grammatical rules are invented. (1963:646)

One of the best examples of the stability of functors is the case of Ilongot in the Philippines. Dyen (1965:32) classified it outside of all Philippine groups as an independent member of the Northwest Hesion. Walton (1977:18) classified it as the first to split from Northern Philippine languages; he discounted its higher percentages shared with Southern Cordilleran (Pangasinan and Karaw), with which it does belong, as attributable to mutual borrowing from Pangasinan.

The Ilongot lexicon shows much independent innovation, and hence the interference with its lexicostatistical scores. But if one looks at Ilongot functors, particularly the pronouns and deictics, the similarity to other Southern Cordilleran languages becomes apparent. [See Table 1.]

TABLE 1. ILONGOT PRONOUNS AND DEICTICS COMPARED WITH SOME OTHER MEMBERS OF SOUTHERN CORDILLERAN.

PRON	Kayapa	Inibaloi	Pangasinan	Ilongot	COG?	PROTO-SC
1	hi'gak	si'kaḳ	siák	si'ak	+	*si'ka-ak
1+2	hi'gata	si'kata	sikatá	sikisi	+	*si'ka-ta
2	hi'gam	si'kam	sikáˊ	sika	+ -	*si'ka-ka *si'ka-m(u)
3	hi'gatu	si'kato	sikatúˊ	siya	-	*si'ka-tu
1+1	hi'gami	si'kami	sikamíˊ	sikami	+	*si'ka-mi
1+2+	hi'gatayu	si'kito	sikatayúˊ	sikisi	+	*si'ka-tayu
2+2	hi'gayu	si'kayo	sikayúˊ	siki	+	*si'ka-yu
3+3	hi'gada	si'kara	sikaráˊ	siyay-də	-	*si'ka-da
DEC-Topic forms						
1	hi'aday	sajay	iyáˊ/sáyay	tu	-	*s()-yay
2	hi'atan	satan	itán/sátay	ta	+	*s()-tan
3	hi'amman	saman	imán/sámay	ma	+	*s()-man
DEC-Locative forms						
1	diyay	ciyay	diyáˊ	'itut	-	*di-yay
2	ditan	citan	ditán	'itat	+	*di-tan
3	diman	ciman	dimán	'imat	+	*di-man

TABLE 2. SOME DIFFERENCES IN FUNCTORS BETWEEN NORTHERN-SAMAR AND WARAY-WARAY(BISAYAN).

Northern-Samar	Waray-Waray	GLOSS
siyáˊ	hiyáˊ	*he/she*
siráˊ	hiráˊ	*they*
si	hi	nominative person marker
si(n)	hin	indefinite genitive marker
sa(n)	han	definite genitive marker
sa	ha	oblique marker
'a:k(u')	'a:kun	*mine*
'a:m(u')	'a:mun	*ours (exclusive)*
'a:t(u')	'a:tun	*ours (inclusive)*

This subgrouping of Ilongot with other Southern Cordilleran languages is further substantiated on the basis of exclusively shared innovations: the replacement of PPH *ña *his/her* by *tu, the *si'ka- nominative pronoun formative, the deictic *tan denoting position near addressee, the deictic *man denoting remote position; the assimilation of PPH *a in the penult to the vowel in the ultima (PPH *ta:[']uh *person* > SC, Ilongot tu'u) or to a final diphthong (PPH *ka:yuh *tree, wood* > PSC *kiyəw > Ilongot kiyu); lexical innovations including Ilongot də:gin, Inibaloi cadin, Pangasinan dálin *earth* (replacing PPH *ta:naq, *lu:paq), Ilongot tawən, Inibaloi tabən, Pangasinan táwən *sky* (replacing PPH *la:ŋit; note PPH *taqwən *year*).

While any innovation can be borrowed or can spread across language boundaries, functors tend to be less open to large scale borrowing or systematic replacement since functors consist of closed paradigms or restricted-class morphemes. Thus, a single pronoun or verb affix might be borrowed, but not an entire paradigm. The Ilongot functors that do not agree with those of the other SC languages turn out to be either retentions (e.g., siya *he/she* < PPH *siyá, or tu *this* < PPH *'i-tú), or independent innovations (e.g., Ilongot siyay-də < PSC *siyay *this* + *-da *they*, or the final -t on the locatives replacing the final consonant of the stem) [note a similar paradigmatic replacement by -y in the Pangasinan topic deictic alternates].

4. THE QUANTITATIVE USE OF FUNCTORS.

A number of scholars, past and present, have used functors qualitatively. Some admirable studies include those of Greenberg (1963) on African languages, and Schebeck (n.d.) on Yuulngu (Australian) languages. However, only two studies to my knowledge have developed a method for dealing with functors quantitatively [McFarland (1974) and Zorc (1977)]. Those interested in the independent evolution and rationale of these methods are referred to those studies.

Basically, McFarland's method, called *morphemic differentiae analysis*, systematically compares all paradigms of all functors between speech varieties. Each difference is scored negatively [see below], and the score reflects the total number of differences observed. Thus the *lower* the numerical score, the closer the genetic relationship posited. Zorc's method, originally called *functor classification*, selects one-hundred basic functors specific to a language family which are observed (or likely) to differ from one speech variety to another. Pairs are scored according to a strict principle of morphological identity (i.e., any difference not directly attributable to a sound change is scored negatively). Thus, the resultant scores reflect the total number of exact cognates, so that the *higher* the numerical score, the closer the genetic relationship posited.

Both methods agree in one principle: once counted, a difference is not counted again. For example [Table 2], the common-noun case-marking particles and the third person pronouns of Northern Samar and Waray-Waray (Bisayan) differ in the replace-

ment of s- in the former by h- in the latter. Further, the first person possessive pronouns in the former end in -' (glottal stop), while those of the latter end in -n. None of these are regular sound changes or correspondences; they don't occur elsewhere in the lexicon or grammar. Rather than subtract a point for *each pair* with the discrepancy, only one point is subtracted for *each discrepancy*. Thus the total negative score by both Zorc's and McFarland's methods would be -1 for the s-:h- difference, -1 for the -':-n difference, -1 for the alternate (short) forms in the N-S case markers, and -1 for the alternate forms of N-S genitive pronouns. A stricter (and less defensible) system of scoring would yield up to -9, instead of -4 for the paradigmatic differences.

The method advocated here may be devised and applied in either of two ways, which may be termed *fine tuning* and *broad-band tuning*. If one is working with closely related speech varieties, one may wish the overall scores among dialect pairs to reflect the greatest amount of difference. Hence, Zorc (1977: 186-91) selected 50 out of the 100 functors which were found to differ. Forms observed to be the same (cognate in every regard) were excluded, e.g., Pan-Bisayan [limá] *five* and [pitú] *seven*; while *one* (reflecting *'əsá, *'isá, *'isa-rá, *sayú', *'isád, *'usád), *two* (reflecting *duhá, *duwá [with unexplained loss of *-h-], or *da-rwá), *three* (reflecting *təlú or *ta-tlú) were included. This, particularly if taken with lexicostatistical comparison [see 5.2 below], resulted in a fine tuning effect.

With widely divergent languages one might like broad-band tuning, i.e., selecting functors that are basically cognate, possibly relaxing the requirement of strict morphological identity. For example, in doing fine tuning, the comparison of Aklanon sánda : Masbate sinda *they* would be scored negatively; in broad-band tuning they could be scored positively [the differences are not the product of regular sound change, but are based on an analogy: si- (singular name marker) : sa- (plural name marker) + -n- ligature + -da *they* (enclitic); they are cognate in part]. However, the principle regarding paradigmatic differences never being counted more than once is applied in both comparisons.

Table 3 is a list of 100 functors devised on the fine tuning model for the Bisayan subgroup of closely related dialects; Table 4 is devised on the broad-band tuning model for the Yuulngu group of distantly related Australian languages. The following classes are useful in drawing up similar lists:

SUBSTITUTES

pronouns (including various case forms, enclitics, alternates),
deictics (including various case forms and verbal derivatives, e.g., *go there, come here*),
locationals (*right, left, this side, otherside, above, below, downriver, upriver, downhill, uphill*, etc.),
temporals (*today/now, yesterday, tomorrow, temporarily, later on, earlier, late, early, last (night)*, etc.),
low numbers (including indicators of number, e.g., dual or plural affixes) and quantifiers (*all, many, some, few*, etc.),

TABLE 3. LIST OF 100 FUNCTORS CHOSEN FOR BISAYAN (PHILIPPINE) CLASSIFICATION.

1. Top/pro-1	26. Verb/dec-1	51. because	76. day(time)
2. 2	27. Verb/dec-3	52. what?	77. year
3. 3	28. Neg-nominal	53. who?	78. today/now
4. 1+1	29. Neg-existential	54. whose?	79. tomorrow
5. 1+2	30. Neg-past	55. when (fut)?	80. yesterday
6. 1+2+	31. Neg-future	56. when (past)?	81. later on
7. 2+2	32. Neg-prohibitive	57. where (past)?	82. earlier
8. 3+3	33. CN/topic	58. where (fut)?	83. morning
9. Gen/pro-1	34. CN/indef.gen.	59. why?	84. afternoon
10. 2	35. CN/defin.gen.	60. how many?	85. act.intr.prog.
11. 3	36. CN/locative	61. how much?	86. act.intr.fut.
12. 1+1	37. Existential	62. how (degree)?	87. act.trans.prog.
13. 1+2	38. Name/topic.sg.	63. one	88. act.trans.past
14. 1+2+	39. Name/gen.sg.	64. two	89. act.trans.fut.
15. 2+2	40. Name/obl.sg.	65. three	90. act.trans.perf.
16. 3+3	41. Name/topic.pl.	66. four	91. passive progressive
17. Obl/formative	42. Name/gen.pl.	67. six	92. passive past
18. Top/dec-1	43. Name/obl.pl.	68. ten	93. passive imperative
19. 1+2	44. now, already	69. on top of	94. passive neg. imper.
20. 2	45. still, yet	70. under	95. instrumental future
21. 3	46. first, please	71. across	96. instrumental imperative
22. Loc/dec-1	47. because (excuse)	72. left	97. instrumental potential
23. 1+2	48. don't know	73. right	98. instrumental perfect
24. 2	49. and	74. within	99. local imperative
25. 3	50. if, when(ever)	75. night	100. local neg. imperative

Composition: pronouns (1-17), deictics (18-27), negatives (28-32), common-noun case-marking particles (33-37), personal-noun case-marking particles (38-43), discourse particles (44-48), conjunctions (49-51), interrogatives (52-62), numerals (63-68), locationals (69-74), temporals (75-84), verb suffixes (85-100). Note: a number of forms that are cognate in every regard have been omitted from this list, e.g., oblique pronouns (cognate with genitive), numbers *limá five, *pitú seven, *walú eight, etc.

TABLE 4. LIST OF 100 FUNCTORS CHOSEN FOR YUULNGU (AUSTRALIAN) CLASSIFICATION.

1. Nom/pro-1
2. 2
3. 3
4. 1+1
5. 1+2
6. 2+2
7. 3+3
8. 1+1+
9. 1+2+
10. 2+2+
11. 3+3+
12. Acc/pro-2
13. 3
14. 1+2
15. Gen/pro-1
16. 2
17. 3
18. 1+1
19. 1+2
20. 2+2
21. 3+3
22. 1+1+
23. 1+2+
24. Obl/pro-2
25. -3
26. Nom/dec-1
27. 1+2
28. 2
29. 3
30. Loc/dec-1
31. 1+2
32. 2
33. 3
34. Topic suffix
35. Nominative
36. Ergative/Instr.
37. Accusative
38. Genitive/Dative
39. Originative (done by)
40. Locative (in/at)
41. Ablative (from)
42. Allative (to/towards)
43. Pergressive (through)
44. Associative (with/by)
45. Locative Increment
46. all
47. one
48. two
49. three
50. Dual suffix
51. Plural
52. that way
53. this way
54. on this side
55. up/above
56. other side
57. downward/bottom
58. down-hill/river
59. up-hill/river
60. Question particle
61. what?
62. why? for what?
63. when?
64. how? by what?
65. who?
66. who? [ergative]
67. what-you-call-it
68. which?
69. from where? [assoc]
70. which way?
71. do what? [verb]
72. none [existential]
73. not [preverbal]
74. not having [suffix]
75. having [suffix]
76. now
77. by the way
78. temporarily
79. only, merely
80. vainly
81. always
82. other [same kind]
83. other [diff.kind]
84. probably, indeed
85. there! look!
86. then, after(wards)
87. and
88. indeed, to be sure
89. because
90. later on [same day]
91. tomorrow
92. today
93. earlier [same day]
94. reflexive/reciprocal
95. causative
96. nominalisation [verb]
97. greedy-for
98. comitative [prefix]
99. past potential
100. having-many [suffix]

Composition: pronouns (1-25), deictics (26-33), case-marking suffixes (34-45), numerals and quantifiers (46-51), locationals (52-59), interrogatives (60-71), negatives (72-74), existentials (74-75), discourse particles (76-89), temporals (90-93), verb suffixes (94-100).

interrogatives (most forms and derivatives, including the verbal *do what?* and the filler *what-you-may-call-it*),

MARKERS:

case, person, number, and class markers; discourse particles (that indicate mood, attitude, belief, time, e.g., *maybe, indeed, don't know, hopefully, vainly, still, yet, only, just, so there,* etc.); negatives; affirmatives; connectives or conjunctions; existentials; pseudo-verbs or preverbs (e.g., *know how, can, want, like, may, might, should,* etc.);

INFLECTIONAL AFFIXES:

voice, tense, aspect, mode, case, number, gender, class, etc.;

DERIVATIONAL AFFIXES:

noun-, verb-, adjective-, and adverb-forming, etc.

5. USEFULNESS OF FUNCTOR ANALYSIS.

The following summarize some of the benefits derived from the use of functor analysis:

1. making explicit the relationship(s) of languages based on synchronically-derived evidence, although it can be posited that the scores must correlate with historical developments;
2. devising a hypothesis about the genetic relationship(s) of speech varieties, which can then be tested by the isolation and evaluation of exclusively shared innovations [a reasonably sound subgrouping hypothesis helps to sort out such problems as borrowing of or counterexamples to proposed innovations];
3. comparing the results of functor analysis with other methods, such as lexicostatistics and the isolation of shared innovations; agreements would serve to substantiate proposed groupings, while disagreements show the directions of influence, interference, borrowing, and the like.

5.1. Schebeck (n.d.) offers a subgrouping of Yuulngu languages of northeastern Arnhem land. He shows how one of the "native theories" correlates closely with the functors. While the subgrouping presented is reasonably sound, it can be made more explicit. Based on data gathered from Schebeck (Id.), Heath (1976), and my own research, scores have been computed for the agreement of several language pairs on the first 50 items from Table 4 [adequate data is not currently available to do the full 100-functor comparison]. These scores are presented in Table 5. They show that at least three of Schebeck's subgroups (DL, DA, DI) form a dialect chain, while three other groups (DK, DN, NN) each form a discrete subgroup equidistant from all other Yuulngu languages. This quantification allows a more refined statement of Yuulngu interrelationships, and probably of their historical development. [I would require more data in these and the other languages to make any conclusions; the present statement may be regarded as a useful hypothesis.]

5.2. Zorc (1977) offers a subgrouping of 36 Bisayan speech varieties based on the agreement of three different methods: lexicostatistics, functor analysis, and exclusively shared innovations. Generally all three agreed in delineating subgroups,

TABLE 5. FUNCTOR SCORES FOR SOME YUULNGU LANGUAGES BASED ON A 50-ITEM COMPARISON. [See Table 4.]

Liyagawumirr (DL)						
48	Djambarrpuyngu (DL)					
42	43	Gumatj (DA)				
39	40	37	Dalwangu (DI)			
26	25	23	24	Ritharrngu (DK)		
21	22	23	20	21	Rirratjingu (DN)	
16	17	18	17	20	24	Golpa (NN)

Note. The two-letter abbreviations refer to Schebeck's posited subgroups.

TABLE 6. RELATIONSHIPS OF KAMAYO TO VARIOUS LANGUAGES OF EASTERN MINDANAO AS SHOWN BY LEXICOSTATISTICS AND FUNCTOR ANALYSIS.

	Surigao	Butuan	Mansaka	Mandaya	Davao	Mamanwa	Tausug
100 lex.	80	78	76	81	78	66	62
100 func.	56	66	77	?	68	61	54
Differs:	-24	-12	+1	?	-10	-5	-8

especially at the topmost node--five branches split from Proto Bisayan (West, Banton, Central, Cebuan, and South).

However, one problem case was the position of the Gubat dialect of Southern Sorsogon (Bikol Province). Lexicostatistically, the highest scores of Gubat were with Sorsogon (83%) and with Masbate (78%); its scores with Northern Samar and Waray were somewhat lower (73%). The functor scores derived in the overall study were generally *lower* than the lexicostatistical scores: for closely related dialects from 2 to 6 points, for distant dialects from 10 to 25 points. Bearing in mind that the functor list was devised to show differences (fine tuning), the proximity and comparability of lexicostatistical and functor scores originally came as some surprise. However, the most surprising result of all was that the functor score of Gubat compared with Northern Samar (82%) was nine points *higher* than its lexicostatistical score (73%). Based on the conservative nature of functors, one is led to make historical inferences. The discovery of several exclusively shared innovations and of shared contrastive features confirmed these inferences [Zorc (1977:272-5)]. A group of Northern Samar speakers had migrated across the rough San Bernardino Strait, and subsequently lost all contact with the Waray group. They began to borrow from the more prestigious Bikol language.

While the lexicostatistical scores are thus inflated, the grammatical system (as reflected in the functors) shows the underlying genetic relationship of this community with the Warayan subgroup, further confirmed by (yet supportive of) the exclusively shared innovations.

5.3. Zorc (1977:18-9,194,287-8) put Kamayo, a language spoken in southern Surigao, into the Mansakan family on the basis of its high functor scores with Mansaka. Recently, some studies have cast doubt on this subgrouping [Walton (1977:27), Gallman (1977:29-31)] based on high lexicostatistical scores with some Bisayan languages, and the failure of Kamayo to share in some phonological innovations attributable to Mansakan languages. [Zorc found that Kamayo did not share enough lexicon (that could not be discounted as borrowings or retentions), functors, or innovations for inclusion within the Bisayan family.]

Table 6 shows the differences between the lexicostatistical and functor scores for Kamayo. While Kamayo shares 80% of vocabulary from the Swadesh 100-meaning list with Surigao (its northern Bisayan neighbor), 81% with Mandaya (its southern Mansakan neighbor), and 76% with Mansaka, its functor score with Surigao is 24 points lower than its lexicostatistical score, but 1 point higher with Mansaka. [Data is not yet available to compute the full functor score with Mandaya, but it should be above the 77% score of Kamayo:Mansaka (based on the skewing of the lexicostatistical scores and the agreement of the functors now available for comparison)]. Thus, the lexicostatistical score with Surigao can be discounted as inflated, due to borrowings from each other and mutually from Cebuano (the lingua franca in that area). But the functor score with Mansaka must be taken as a fairly close indication of the genetic relationship since it is one point higher than the lexicostatistical score. Two points need note:

First, while no Bisayan dialect has a functor score higher than 66% with Kamayo, this relatively high score is an indication that the Bisayan and Mansakan groups are closely related. They are both immediately descended from Proto Central Philippine [Zorc (1977:19,31-3,223-40)]. Kamayo is problematic then because it neighbors both Bisayan and Mansakan language communities, and it has reasonably high scores with members of each--although its functor scores are clearly skewed towards members of the Mansakan subgroup.

Second, Kamayo has some exclusively shared innovations with the Mansakan group. One is the second person plural oblique pronoun *mayú, reflected in the language name [ka-mayú] *to you*, clearly a distinguishing feature for a speech variety located in Bisayan territory where [ka-niyú] or [ka-nínyu] are used. Other innovations include functors such as Kamayo, Mansaka ya'án *he/she*, na'án *his/her*, da *now,already*, 'aw *if*, 'a-du'ún *today*, Kamayo ka-lin-tu'ú, Mansaka ka-rin-tú *right(side)*, Kamayo ki-suúm, Mansaka ki-sərəm *tomorrow*, and the syncope of the penult vowel in the second person singular oblique pronoun, Kamayo, Mansaka ka-nmu *your* from *ka-nímu (attested in Bisayan). Mansakan verb morpho-

logy has a paradigmatic (but otherwise irregular) replacement of PPH *n- by Mansakan y- to show perfective aspect, e.g., *nag- > Mansaka, Kamayo, Davaw yag- (active past), *naka- > Mansaka, Kamayo, Davaw yaka- (active potential past), *naga- > Mansaka, Kamayo, Davaw yaga- (active progressive), and *-in- > Mansaka, Kamayo -i(y)- (passive past infix). There are some exclusively-shared lexical innovations, such as the replacement of PPH *hapúy *fire* by Mansakan *'atulun (Mansaka 'aturun, Kamayo 'atuun) [note Bisayan and Bikol reflect *kalá́yu]; the complex reformation of PPH *kukúh *fingernail* as Mansakan *kulkulhun (Mamanwa kulkulhun, Mansaka kukurun, Kamayo kukuhún); the replacement of PPH *'anák or *'uŋáq *child* by Mansakan *'ísə' (Mansakan 'isə', Kamayo, Davaw 'ísu') [note Bisayan and some other Southern Philippine languages reflect *báta']. Other lexical innovations for which Kamayo has cognates include Mansakan *hambuŋ *afternoon*, *sugbu *bathe*, *pəsa' *bone*, háku' *cough*, *tigám *know (how)*, *hikəl *laugh*, *tának *lose*, *ma-da'ig *many*. There is also contrastive evidence that while some forms are not innovations limited to Mansakan, not one cognate is found in a single Bisayan dialect, e.g., *bubáy *woman*, *'utáw *person*, *sirán *they*, *yaŋ topic marker--cognates of these are found in Kamayo, Mansaka, and other speech varieties that may be subgrouped together as Mansakan.

Thus, Kamayo belongs in a subgroup with Mansakan languages, although at a higher order since it fails to share at least one qualitative phonological innovation--the assimilation of *Cl clusters to Mansakan *ll. [The exact position need not be discussed here, but is posited in Zorc (1977) and Gallman (1977).] This subgroup is substantiated initially on the basis of functor analysis, and, most importantly, exclusively shared innovations. But the indication of this subgrouping given by functor analysis is not to be disregarded or dismissed. It helps sort out the directions of borrowing (almost exclusively *from* Bisayan) and certain irregularities (e.g., the failure to share in some innovations).

6. CONCLUSIONS.

Comparison of the results of functor analysis and other methods such as lexicostatistics, brings to light secondary contacts: high functor scores (as compared with lower lexicostatistical scores) indicate a close genetic relation undone by long-term contact. Similarly, low functor scores (compared with high lexicostatistical scores) can show the directions of borrowing across linguistic boundaries, i.e., the grammatical systems of languages prove to be more conservative.

It is suggested that synchronic or historical studies would profit from the use of functor analysis in conjunction with other established methods, allowing initial working hypotheses about language interrelationships, and giving considerable weight to the classification(s) thereby obtained, where substantial lines of agreement are found.

REFERENCES

Dyen, Isidore. 1965. A Lexicostatistical Classification of the Austronesian Languages. Memoir 19--Research Center in Anthropology, Folklore, Linguistics. Baltimore: Waverly Press.

Gallman, Andrew F. 1977. "Proto South East Mindanao and its Internal Relationships." Paper presented at the Austronesian Symposium, Honolulu, August 1977. Manila: S.I.L.

Greenberg, Joseph H. 1963. The Languages of Africa. Publication 25, Indiana University Research Center in Anthropology, Folklore, and Linguistics. IJAL, Vol. 29, No. 1 (Part II). Bloomington: Indiana University.

Heath, Jeffrey. 1976. Ritharngu Grammar, Texts, and Dictionary. Manuscript: Australian Institute of Aboriginal Studies files.

Hockett, Charles F. 1958. A Course in Modern Linguistics. New York: The MacMillan Company.

McFarland, Curtis D. 1974. The Dialects of the Bikol Area. Ph.D. Dissertation: Yale University.

Reid, Lawrence A. 1971. Philippine Minor Languages: Word Lists and Phonologies. Oceanic Linguistics Special Publication #8. Hawaii: University Press.

Schebeck, Bernhard. n.d. Dialect and Social Groupings in North East Arnhem Land. Manuscript: Australian Institute of Aboriginal Studies files.

Teeter, Karl V. 1963. "Lexicostatistics and genetic relationship," Language 39.4:638-48.

Walton, Charles. 1977. "A Philippine language tree." Paper presented at the Austronesian Symposium, Honolulu, August 1977. Manila: S.I.L.

Zorc, R. David. (1976) "On the development of contrastive word accent: Pangasinan, a case in point." To appear in South-East Asian Linguistic Studies, Vol. 3; Pacific Linguistics, Series C, No. 45. Canberra: The Australian National University.

-----. 1977. The Bisayan Dialects of the Philippines: Subgrouping and Reconstruction. Pacific Linguistics, Series C, No. 44. Canberra: The Australian National University. [Originally Ph.D. Dissertation: Cornell University, June 1975.]

FUNCTOR ANALYSIS OF YOLŊU. R. David Zorc

SCORES FROM THE COMPARISON. (100 FUNCTOR LIST, see enclosure)

Madarrpa												
88	Wagilak											
88	95	Ritharrŋu										
43	44	42	Dhalwaŋu									
46	50	47	67	Liyagawumirr								
48	50	49	65	94	Gupapuyŋu							
50	52	50	65	88	88	Djambarrpuyŋu						
50	50	51	66	83	86	91	Gumatj					
43	47	46	66	79	79	85	81	Djapu				
38	40	40	45	52	52	55	52	48	Gälpu			
35	36	37	42	50	50	50	52	45	90	Rirratjiŋu		
33	32	33	37	47	47	45	43	42	85	84	Wangurri	
31	30	31	41	43	44	41	44	42	84	79	80	Warramiri
36	32	32	39	41	40	35	35	38	46	46	50	40 Golpa

PHONOLOGY OF YOLŊU LANGUAGES ("typewriter phonetics" used hereafter):

1	2	3	4	5	6	7	
p	t	t	c	T	k	'	*fortis*
b	d	d	j	D	g		*lenis*
m	n	n	ñ	N	ŋ		*nasal*
w		r	y	R			*continuant*
		l		L			*lateral*

1 *Bilabial*
2 *Lamino-dental: th, dh, nh*
3 *Apico-alveolar: rr*
4 *Lamino-alveolo-palatal: tj, dj, ny*
5 *Retroflex (apico-domal): ṯ, ḏ, ṉ, ṟ, ḷ*
6 *Velar*
7 *Glottal*
8 *Front: i, e*
9 *Central: a, ä*
10 *Back: u, o*

8	9	10
i		u
i:		u:
	a	
	a:	

100 FUNCTOR LIST FOR YOLŋULANGUAGES. [Revised from that of 8.78] rdz
Final Version as of 11.79.

001 Nom/pro-1 *I* ŋara / ŋaya
002 Nom/pro-2 *you-1* ni: / nunu
003 Nom/pro-3 *he/she* ŋayi / na:n
004 Nom/pro-1+3 *we-2-excl* ŋaliñu
005 Nom/pro-1+2 *we-2-incl* ŋali
006 Nom/pro-2+2 *you-2* numa-maNDa /-wulay
007 Nom/pro-3+3 *they-2* maNDa / dupal
008 Nom/pro-1+3+ *we-all-excl* ŋanapu(ru)
009 Nom/pro-1+2+ *we-all-incl* ŋilimuru / ŋalma
010 Nom/pro-2+2 *you-all* numa / ñi:li
011 Nom/pro-3+3+ *they-all* walala / danal
012 Acc/pro-2 *you-1 (object)* nuna
013 Acc/pro-3 *him/her (object)* ŋaña / nana
014 Acc/pro-1+2 *us-2-incl (obj)* ŋicalaŋa / ŋaliñ
015 Gen/pro-1 *mine* ŋaraku / ña:ku
016 Gen/pro-2 *yours-1* nuŋu / nuŋgu
017 Gen/pro-3 *his/hers* nanŋu / nanŋu
018 Gen/pro-1+3 *ours-2-excl* ŋiñalaŋgu / ŋaliñuŋgu
019 Gen/pro-1+2 *ours-2-incl* ŋicalaŋgu / ŋaliŋgu
020 Gen/pro-2+2 *yours-2* numalaŋgu / numaliŋgu
021 Gen/pro-3+3 *theirs-2* maNDaŋgu / dupaliŋgu
022 Gen/pro-1+3+ *ours-all-excl* ŋanapuruŋgu / ŋanapiliŋgu
023 Gen/pro-1+2+ *ours-all-incl* ŋilimuruŋgu / ŋalmaliŋgu
024 Obl/pro-3 *to(wards) him* nanu-kala / nan-guLi
025 Question particle *really?* muka / ŋaca
026 Nom/deic-1 *this (near me)* duwala / daŋu
027 Nom/deic-1+2 *this (near us)* duwali / duwan
028 Nom/deic-2 *that (near you)* ŋuni(yi) / ŋuna
029 Nom/deic-3 *that (yonder)* ŋuna / bana
030 Loc/deic-1 *here (near me)* diyala / jinal
031 Loc/deic-1+2 *here (near us)* diyali(yi) / jinalaya
032 Loc/deic-2 *there (near you)* ŋunili(yi) / ŋunala(ya)
033 Loc/deic-3 *there (yonder)* ŋunala / banalaya

100 FUNCTOR LIST FOR YOLŊU LANGUAGES (11.79)

034 Topic suffix *"the"* -(ñ)ja / -ma
035 Ergative/Instrumental *'with (X)'* -du (-yu)
036 Genitive/Dative *"to/for"* -gu (-wu)
037 Originative *"produced by, from"* -guŋu / -wuŋ
038 Locative-inanimate *"at"* -ŋuRa / -ŋa
039 Ablative-inanimate *"from"* -ŋuRu
040 Allative-inanimate *"to(wards)"* -lili / -Li
041 Pergressive-inanimate *"through"* -kuru / -muru
042 Associative-inanimate *"about"* -buy (-wuy)
043 Associative-animate *"about/concerning"* -gala-ŋu-wuy / -wuRu-wuy
044 all bukmak / warpam' / guLku
045 one waŋgañ
046 two ma:rma' / bulal'
047 three Lurkun'
048 Dual-suffix -maNDa / -wulay
049 Plural-suffix -mala / -wuru / -wara
050 not-having (suffix) -miRiw / -nara

---

051 having (suffix) -miri / -mi
052 that-way, to there bala
053 this-way, to here lili / Ra:li
054 on this side balakuru / duwalayaku / dawican
055 otherside Laypa / Raypa
056 up/above garwaR / garamat
057 downward/bottom ŋu:y-ŋuRa / ŋu:y-ŋa
058 down-hill/river yarup-turu / yarup-tuwa
059 up-hill/river DuwaT-turu / DuwaT-tuwa
060 what? na:
061 why?, what for? na:ku / na:wu
062 when? na:ta
063 how?, by what? na:yu / na:liy
064 who? yu:l
065 who (ergative)? yu:ltu
066 what-you-may-call-it na:wuy / na:pa
067 which? wana(ka) / ŋala
068 from where? [associative] wanaŋuwuy / ŋalaŋuwuy
069 which way?, where to? wanamala / wanakuru / ŋalawican
070 do(ing) what? nalcan / nalpiyan

100 FUNCTOR LIST FOR YOLŊU LANGUAGES (11.79) rdz

071 none (existential) ba:yŋu(na) / da:wul(na)
072 not (preverbal) yaka
073 today ga:tuRa / jinaŋbala
074 tomorrow gu:Dar' / baRktu
075 yesterday baRpuRu / yawungu
076 later-on (today) yalala / yalŋuwa
077 earlier (today) ŋa:tili / ŋa:cil
078 other (same); again bulu / biyapuL
079 other (different) wiRipu / waRipu
080 then, so; because bili
081 afterwards; and then bi:ŋuRu(ñ)ja / bi:waLi
082 and ga
083 causative (suffix) -maRa- / -ma-
084 inchoative (suffix) -ti-ri / -yi-
085 progressive (preverb) yukura / ga' / yaka / ma
086 definite-future (preverb) yuru / du / ŋaru
087 tomorrow-future (preverb) buŋguŋu / buŋgama / baRktu
088 habitual/repetitive (preverb) ŋuli / bayiŋ
089 go; walk marci / ŋaruŋ
090 stand da:ra / da:ya
091 sit ni:na / ñi:na
092 lie-down ŋu:ra / yukura / ŋu:ya
093 slowly bulna
094 quickly bu:ndi / ganjaryu
095 carefully gurum'
096 maybe maku / wilak
097 still, yet baDak
098 emphatic/reflexive (pronoun) -pi / (-bay)
099 greedy-for [prefix/suffix] buku- / Da:mbu- / gayawak-
100 (it is) finished Liñguna / bilin

FORMS OMITTED FROM PREVIOUS LISTS:

Oblique-pronoun-2 '(to)wards you-1' [covered #012, 016, 024]
Nominative [Pan-Yolŋu * -Ø suffix]
Accusative [Pan-Yolŋu *-na suffix, realised as -ñ in vowel-dropping lgs]
Rightful-owner [Pan-Yolŋu *-waTa-ŋu; not felt to be a functor per se]
Animate locative, allative, ablative [formation as for Animate-Assoc.#043]
Reflexive/Reciprocal Verb suffix [same as #051 'having']

EVIDENCE FOR SUBGROUPING WITHIN YOLŊU:

1. Golpa - independent - low scores throughout; 50% with Warramiri.
2. Gälpu-Rirratjiŋu-Wangurri-Warramiri - cluster - mutually high scores.
3. Djapu-Gumatj-Djambarrpuyŋu-Gupapuyŋu-Liyagawumirr-Dhalwaŋu - cluster - mutually high scores; with Dhalwaŋu as "well-marked" dialect, possibly forming independent subgroup.
4. Madarrpa-Wagilak-Ritharrŋu - cluster - mutually high scores.
5. Djinaŋ - independent [Insufficient data to date to complete full list, but enough cognates to insure macro-group membership.

Further evidence for subgrouping?

1 + 2 "Northern Yolŋu"

3 + 4 "Southern Yolŋu"

5 "Inland Yolŋu"

THE "POWER" OF THE FIRST 50 FUNCTORS (pronouns, deictics, case-marking suffixes, numerals).

	1	2	3	4	5	6	7	8	9	10	11	12	13
Madarrpa													
Wagilak	46												
Ritharrŋu	48	47											
Dhalwaŋu	24	23	22										
Liyagawumirr	23	24	23	37									
Gupapuyŋu	25	23	24	35	47								
Djambarrpuyŋu	24	23	22	36	44	42							
Gumatj	26	23	25	35	43	45	46						
Djapu	20	22	21	35	40	39	43	39					
Gälpu	17	15	16	18	18	17	20	18	16				
Rirratjiŋu	17	15	16	17	19	18	19	20	16	49			
Wangurri	15	13	14	15	18	19	19	18	16	46	45		
Warramiri	15	13	14	16	16	17	16	16	16	44	41	42	
Golpa	18	15	16	19	18	18	18	18	16	23	23	23	21

SOUTHERN YOLŊU

NORTHERN YOLŊU

EVIDENCE FOR SUBGROUPING WITHIN YOLŊU.

1a. "NORTHERN YOLŊU" [(1) Gol; (2) Gal, Rir, Wan, War]

1a.1. 002 *nunu 'you-1' : 1 nu:nu, 2 nunu; (2) War ñunu;(5 ñone).

2. 007+048 'they-2; dual' : 1 -balay; 2 - wulay; (5 biliŋi).

3. 010 *nuruli 'you-3' : 1 nuruli; 2 ñi:li (< *nuyuli); (5 ŋiliji).

4. 026 *daŋu 'this-1' : 1 naŋu (assimilation); 2 daŋu; War jaŋu.

5. 038 *-ŋa Locative : 1, 2 -ŋa.

6. 041 *-muru Pergressive : 1, 2 -muru.

7. 043 *wuRu- Animate Increment : 1, 2 -wuRu- [-wuRu-wuy / -wuRu-y].

8. 050 *-nara Privative : 1 -nara-ŋu; 2 -nara.

1b. "SOUTHERN YOLŊU" [(3) Dal, Liy, Gup, Jam, Gum, Jap; (4) Mad, Wag, Rit]

1b.1. 002 *ni: 'you-1' : 3, 4 ni:.

2. 007+048 'they-2; dual' : 3, 4 maNDa 'they-2'; 3 -maNDa 'dual suffix'.

3. 021 *maNDaŋgu 'theirs-2' : (3) Dal, Liy maNDaŋ, Jam maNDak, Gup maNDaŋgu, Gum maNDaku, [Jap maNDal]; 4 maNDaŋu.

4. 041 *-kuru Pergressive : (3) Gum, Gup -kuru, Dal, Liy, Jam, Jap -kur; (4) Mad, Rit -kuru, Wag kur.

5. 043 *Gala- Animate Increment : (3) Dal, Jam, Jap -wala-ŋu-, Gum -gala-ŋu-, Gup -gala-ŋa-, Liy -wala-ŋa-; 4 -gala-

6. 050 *-miRiw Privative : (3) Dal -miR, others -miRiw; 4 -miRiw.

7. 053 *lili 'this-way' : (3) Gum, Jam, Jap lili, Dal Li; 4 lili.