On Implicational and Distributional Universals of Word Order
Author(s): John A. Hawkins
Source: *Journal of Linguistics*, Vol. 16, No. 2 (Sep., 1980), pp. 193-235
Published by: Cambridge University Press
Stable URL: http://www.jstor.org/stable/4175546
Accessed: 06/06/2009 18:24

J. Linguistics 16 (1980) 171–338 Printed in Great Britain

On implicational and distributional universals of word order[1]

JOHN A. HAWKINS

Department of Linguistics, University of Southern California

(Received 15 January 1979)

1. Introduction

This paper argues that the theory of Universal Grammar must include both implicational universals and universals of language distribution in the description and explanation of word order.

Implicational universals are of the form: if a language has some property (or properties) P, then it will also have some property (or properties) Q. These 'if P then Q' statements require that a language must not have property P without property Q (*P & −Q). If a language does have P, therefore, it must have Q as well (P & Q), but in the absence of P, a language may still have Q (−P & Q), or else it may have neither property (−P & −Q). Three co-occurrence types are permitted, and just one (*P & −Q) is disallowed. The implicational statement thus defines universal parameters on types of language variation, allowing for some co-occurrences of linguistic phenomena, but not others.

A distributional universal is supplementary to such implicational statements. It is a principled and predictive statement defining the relative frequency across the languages of the world of the co-occurrences which are permitted by the various implicational universals.

Implicational universals of word order were first proposed in Greenberg (1966).

---

[1] I would like to thank Bernard Comrie, Paul Friedrich, Larry Hyman, Ed Keenan, Bob Stockwell, and Sandra Thompson for helpful discussion and comments in connection with the work which is reported here. I would also like to thank the students in my seminar on word order (USC, Spring 1978) who helped me check the data of Greenberg's Appendix II and who provided valuable feedback and comments: Saeed Ali (UCLA), Linda Arvanites (UCLA), Alan Kim, Wei-Lin Lei, Galust Mardirussian (UCLA), Charles Randriamasimanana, Jack Roberson, Maner Thorpe, Yukiko Uchida, and Emily Yarnall.

The following abbreviations are used in this paper: VSO = Verb–Subject–Object; SVO = Subject–Verb–Object; SOV = Subject–Object–Verb; SOVr = the rigid subtype of SOV; SOVnr = the non-rigid subtype of SOV; V–initial = Verb–initial order in the sentence; Prep (or Pr) = Preposition (before NP); Postp (or Po) = Postposition (after NP); Adp = Adposition; NA = Noun–Adjective; AN = Adjective–Noun; NG = Noun–Genitive; GN = Genitive–Noun; NDem = Noun–Demonstrative Adjective; DemN = Demonstrative Adjective–Noun; NRel = Noun–Relative Clause; RelN = Relative Clause–Noun; N– = Noun–initial order within NP; –N– = Noun–medial order within NP; –N = Noun–final order within NP; S = Subject; DO (or O) = Direct Object; IO = Indirect Object; OBL = Oblique NP; ADV = Adverbial Phrase.

0022-2267/80/0016-0023$02.00

He derived his implicational generalizations both from a detailed 30-language sample, and (for basic word order data) from a large sample of languages presented in his Appendix II. More recently, Vennemann (see references) has provided a reanalysis of Greenberg's implicational universals. He divides all of Greenberg's 'meaningful elements' (verb and direct object, noun and adjective, etc.) into 'operand' and 'operator' categories on semantic and syntactic grounds. He proposes that languages serialize all these elements in a consistent order: either operator before operand, or operand before operator.

The first goal of the present paper is to offer revised implicational statements for Greenberg's data, and to discuss the general properties of implicational statements within Universal Grammar. It is my contention that Greenberg's implicational statements are not as precise as they might have been, relative to his data, and that this lack of precision has resulted in universal word order generalizations being missed. I shall argue also that his statistical implicational universals (i.e. those admitting of exceptions) are of questionable legitimacy, and that the facts which he was attempting to capture by means of them are just a few of the many facts which a distributional universal will predict with greater adequacy. With regard to Vennemann's reanalysis, I shall argue that this, in effect, redefines Greenberg's unilateral implications (if P then Q, but not vice versa) as bilateral implications (if P then Q, and if Q then P), and that the result, when considered in relation to Greenberg's data, is a description with less than 50 per cent accuracy.

The data which we find most revealing for our purpose is Greenberg's Appendix II, which is still the biggest single language sample available giving data on the co-occurrence of the following four word orders: verb position, adposition order (i.e. prepositions or postpositions), and adjective and genitive order in relation to the noun. This appendix (henceforth AII) is reproduced as Table 1. It will be readily apparent that many of the AII co-occurrence cells are totally unattested. Our implicational statements aim to predict the balance between attested and non-attested co-occurrences in a regular way.[2]

---

[2] What is important about Greenberg's Appendix II is the wide genetic and areal coverage. Many of the entries refer to whole families (e.g. Celtic, Bantu, Dravidian), and the individually named languages are genetically diverse, with the result that the majority of the world's language families are here represented. In a seminar on word order which I organized at USC in the Spring of 1978 one of the collective projects was to check as many as possible of these entries, particularly the families and groups, and to add to the sample. Our research provided strong confirmation for the implicational patterns derivable from AII. We could find NO counterexamples to the non-statistical implications (I), (II), (III'), and (IV') presented in Section 3.1. However, our findings did suggest that some of Greenberg's entries should be reclassified among the already attested types. Our expanded sample was compiled and classified too late to be included in the present paper. In any case the purpose of this paper is to reveal the word order regularities which are contained in Greenberg's original data base. The following classification errors should be pointed out, however:

Type 9: Modern Greek should be type 10;

Type 9: 'almost all' should replace 'all Bantu languages' (Tunen and Bandem are type 17, Larry Hyman, personal communication);

It will also be apparent that the sizes of the attested co-occurrence cells vary considerably. And our second goal is to propose a distributional universal, the principle of Cross-Category Harmony (henceforth CCH), which accounts for these distributional differences. In stating this principle we retain Vennemann's distinction between operator and operand categories, but our harmony principle differs in all other major respects from his serialization principle.

It must be stressed at the outset that Greenberg's AII does give only a convenience sample of the world's languages, albeit a large sample which names numerous families of languages in addition to individual languages. Whether my generalizations carry over to even larger samples remains to be seen. In this context I shall simply describe AII as it stands, and shall draw from it the very suggestive generalizations which I believe to underlie it, as a stimulus for further research in this direction, and as an antidote to many word order universals which are currently accepted as valid.

The order of presentation is as follows. We first provide, in Section 2, a summary and critique of Vennemann's word order universals, testing them against Greenberg's data. Section 3 defines implicational universals for the AII data, discusses the general properties of adequate implicational universals, and offers an explanation for these universals being the way they are. Section 4 defines the distributional universal, tests it in detail against Greenberg's data, compares it with Vennemann's and Greenberg's universals, and points to a possible explanation for it.

Type 14: Rutulian and other SVO Daghestan languages in the Caucasus should be type 15 (Maner Thorpe, personal communication);
Type 15: Ijo should be type 23 (cf. Givon, 1975);
Type 16: most Mandingo (i.e. Mande) languages should be type 24 (cf. Givon, 1975; Larry Hyman, personal communication).

In addition, for the following entries, the verb, adposition or adjective order turn out to be of undecidable basicness:

Type 10: German has both SOV and SVO orders;
Type 15: Chinese also has both SOV and SVO (cf. Li & Thompson, 1975);
Type 16: Nupe is both pre- and postpositional (Larry Hyman, personal communication).

Our expanded sample would further permit us to delete the word 'probably' before:

Type 1: 'other Austronesian languages';
Type 2 'other Philippine Austronesian languages';
Type 15: 'Algonquian'.

Finally, Papago (Postp & VSO) has been entered into AII as reproduced here. This language is mentioned by Greenberg in a note on p. 107, although not included in the Appendix as such. Ed Keenan has confirmed for me (personal communication) on the basis of his work on verb–initial languages that Postp & VSO is indeed an existing, though limited, type. Its occurrence is in fact predicted by the existence of languages of the opposite type: Prep & SOV (types 17 and 19). The limited distribution of both Postp & VSO and Prep & SOV languages, and the slightly larger distribution of the latter, are predicted by our principle of Cross-Category Harmony.

Distribution of Basic Order Types:	*Number of languages*
1. *VSO/Pr/NG/NA.* Celtic languages; Hebrew, Aramaic, Arabic, Ancient Egyptian, Berber; Nandi, Masai, Lotuko, Turkana, Didinga; Polynesian languages and probably other Austronesian languages; Chinook, Tsimshian; Zapotec, Chinantec, Mixtec, and probably other Ote-Mangue languages.	19
2. *VSO/Pr/NG/AN.* Tagabili and probably other Philippine Austronesian languages; Kwakiutl, Quileute, Xinca.	5
3. *VSO/Pr/GN/AN.* Milpa Alta Nahuatl.	1
4. *VSO/Pr/GN/NA.* No examples.	0
5. *VSO/Po/NG/NA.* No examples.	0
6. *VSO/Po/NG/AN.* No examples.	0
7. *VSO/Po/GN/AN.* Papago.	1
8. *VSO/Po/GN/NA.* No examples.	0
9. *SVO/Pr/NG/NA.* Romance languages, Albanian, Modern Greek; West Atlantic languages, Yoruba, Edo group, most languages of Benue–Congo group including all Bantu languages; Shilluk, Acholi, Bari, most languages of Chad group of Hamito-Semitic but not Hausa; Neo-Syriac, Khasi, Nicobarese, Khmer, Vietnamese, all Thai languages except Khamti; many Austronesian languages including Malay; Subtiaba.	21
10. *SVO/Pr/NG/AN.* German, Dutch, Icelandic, Slavonic, Efik, Kredj, Maya, Papiamento.	8
11. *SVO/Pr/GN/AN.* Norwegian, Swedish, Danish.	3
12. *SVO/Pr/GN/NA.* Arapesh (New Guinea).	1
13. *SVO/Po/NG/NA.* No examples.	0
14. *SVO/Po/NG/AN.* Rutulian and other Daghestan languages in the Caucasus.	2
15. *SVO/Po/GN/AN.* Finnish, Estonian, Ijo, Chinese, Algonquian (probably), Zoque.	6
16. *SVO/Po/GN/NA.* Most Mandingo and Voltaic languages, Kru, Twi, Gã, Guang, Ewe, Nupe, Songhai, Tonkawa, Guarani.	11
17. *SOV/Pr/NG/NA.* Persian, Iraqw (Cushitic), Khamti (Thai), Akkadian.	4
18. *SOV/Pr/NG/AN.* No examples.	0
19. *SOV/Pr/GN/AN.* Amharic.	1
20. *SOV/Pr/GN/NA.* No examples.	0
21. *SOV/Po/NG/NA.* Sumerian, Elamite, Galla, Kanuri, Teda, Kamilaroi and other southeastern Australian languages.	7
22. *SOV/Po/NG/AN.* No examples.	0
23. *SOV/Po/GN/AN.* Hindi, Bengali and other Aryan languages of India; Modern Armenian, Finno-Ugric except Finnish group; Altaic, Yukaghir, Paleo-Siberian, Korean, Ainu, Japanese, Gafat, Harari, Sidamo, Chamir, Bedauye, Nama Hottentot; Khinalug, Abkhaz and other Caucasian languages; Burushaski, Dravidian; Newari and other Sino-Tibetan languages; Marind-Anim, Navaho, Maidu, Quechua.	28
24. *SOV/Po/GN/NA.* Basque, Hurrian, Urartian, Nubian, Kunama, Fur, Sandawe, Burmese, Lushei, Classical Tibetan, Makasai, Bunak (Timor), Kate (New Guinea), most Australian languages, Haida, Tlingit, Zuni, Chitimacha, Tunica, Lenca, Matagalpa, Cuna, Chibcha, Warrau.	24
Total number of languages =	142

No. with VSO = 26 No. with Pr = 63 No. with N– = 51
SVO = 52 Po = 79 –N– = 51
SOV = 64 –N = 40

Table 1

Greenberg's Appendix II

In calculating the language quantities for each co-occurrence cell I have counted the language families and groups as individual languages only, since Greenberg gives no information on how many members of each group he has actually checked. To assign to these groups any number of languages larger than 1 would give rise to problems in justifying why 2, rather than 3, rather than 4, etc., were assigned.

2. VENNEMANN'S THEORY

Vennemann makes an important innovation relative to Greenberg's earlier word order universals. Whereas Greenberg had operated with a three-way typology of languages, VSO, SVO, and SOV, Vennemann proposes just a two-way typology: VO (or VX) versus OV (or XV). VSO and SVO are, therefore, collapsed into one type, VO.

Languages with OV have co-occurring word order patterns in the order operator before operand, while VO languages have just the reverse, operand before operator. The following is a sample of Vennemann's operator and operand categories:

		Operator	*Operand*
I.	(a)	object	verb
	(b)	adverbial	verb
	(c)	main verb	auxiliary
	(d)	main verb	modal
II.	(a)	adjective	noun
	(b)	relative clause	noun
	(c)	genitive	noun
	(d)	determiner	noun
III.	(a)	adverbial	adjective
IV.	(a)	noun phrase	adposition (prep or postp)

The criteria for allocating some category to operator or operand status are that

> semantically the application of an operator results in a specification of the operand predicate, and, syntactically, that the application of an operator to an operand results in a constituent of the same general category as that of the operand. . . .

For example,

> Adverbials are operators on verbs because the result of the operation is a complex verb rather than a complex adverbial. Main verbs which are infinitives are operators on finite modals, which are finite verbs, because the result of the operation is a finite verb rather than an infinitive. And noun phrases are operators on prepositions, which are transitive adverbials, because the result of the operation is a prepositional phrase, i.e. an adverbial, rather than a noun phrase (Vennemann, 1972).

Having classified these word order categories as either operators or operands, Vennemann then formulates his 'Natural Serialization Principle'. According to this principle, languages will serialize all their operator–operand pairs either operator before operand, or operand before operator. It is stated as follows:

$$\{\text{Operator } \{\text{Operand}\}\} \Rightarrow \begin{cases} [\text{Operator } [\text{Operand}]] \text{ in OV languages} \\ [[\text{Operand}] \text{ Operator}] \text{ in VO languages} \end{cases}$$

Vennemann thus defines his two language types abstractly. He notes that not all languages are fully consistent with his schema, and hypothesizes that this is for historical reasons. The inconsistent languages are those which are moving from one type to the other. He offers a phonological and morphological explanation for why a language should shift V relative to O, and he then proposes the Natural Serialization Principle as the source of an analogical pull which brings about the subsequent acquisition of the other operator-operand orderings which are compatible with VO or OV, respectively. The inconsistent languages are, therefore, moving towards a target type which is as yet only partially attained.

2.1. *Critique of Vennemann's Natural Serialization Principle*

The data of Greenberg's AII enable us to test the predictions of the Natural Serialization Principle (henceforth NSP) with regard to four operator–operand pairs: object and verb, adjective and noun, genitive and noun, and noun phrase and adposition.

2.1.1 *Counterexamples to NSP:* Taken literally, NSP permits just two co-occurrences of the four operator–operand pairs listed in AII:

operator before operand: OV & AN & GN & NP+Po
operand before operator: VO & NA & NG & Pr+NP

Only three of the twenty-four logically possible co-occurrence types in Greenberg's Appendix satisfy these two orderings:

type 1:	VSO & NA & NG & Pr+NP	19 lgs
type 9:	SVO & NA & NG & Pr+NP	21 lgs
type 23:	SOV & AN & GN & NP+Po	28 lgs

These three types cover 68 of the 142 languages in the sample (47.89 per cent). Yet AII contains not three but sixteen attested co-occurrence cells, the other thirteen accounting for the remaining 74 languages (or 52.11 per cent). Although types 1, 9, and 23 are among the four most frequent types (along with type 24), therefore, they amount to less than half the languages of the sample.

These figures present an insuperable argument against NSP as it stands. To have over 50 per cent counterexamples is unpalatable for any theory, even one which proposes a historical explanation for synchronic exceptions. The value of this historical explanation is also undermined by the empirical and theoretical arguments in Hawkins (1979) against the whole form of historical logic to which Vennemann appeals.

2.1.2 *Patterns within the counterexamples:* Another problem for NSP comes from the eight co-occurrence cells in AII which are unattested. Given the size of the sample together with its wide genetic and areal coverage, the total absence of eight out of the twenty-four logically possible co-occurrences suggests that there are universal constraints on co-occurring word orders which NSP is failing to capture. And, indeed, there are interesting patterns within the exceptions to NSP. Consider OV languages (types 17 through 24). The predicted noun modifier co-occurrences are OV & AN & GN (types 23 and 19). The exceptions take two forms. Types 17 and 21 have both noun modifiers after the noun, OV & NA & NG; while the only other attested OV type has OV & NA & GN (type 24), i.e. with the adjective alone an exception. There are no OV languages in which the genitive alone is an exception to NSP: OV & AN & NG (potentially types 18 and 22). Consider next VO languages of the VSO subtype (types 1 through 8). The predicted noun modifier co-occurrences are: VO & NA & NG (type 1). The exceptions again take two forms. Either both noun modifiers precede, VO & AN & GN (types 3 and 7); or only the adjective precedes, VO & AN & NG (type 2). There are no VSO languages with the genitive alone as an exception, VO & NA & GN (potentially types 4 and 8). In both cases, therefore, if there is going to be just one noun ordering exception it is always the adjective, and never the genitive. The adjective is evidently a more unstable operator relative to its operand than is the genitive, a pattern which generalizes across the two ideal operator–operand types.

2.1.3 *The status of SVO languages:* A further difficulty for Vennemann's NSP concerns the status of SVO languages. In terms of co-occurring word orders, these languages are typologically ambivalent. Although they share patterns with VSO, they also share with SOV. Thus, most VSO languages in Greenberg's sample have prepositions (96.15 per cent), while most SOV languages have postpositions (92.19 per cent). SVO languages have significant numbers of both: 36.54 per cent are postpositional, 63.46 per cent prepositional. VSO languages permit three of the four logically possible co-occurrences of adjective and genitive (all but NA & GN), while SOV languages also permit three possibilities (all but AN & NG). But SVO languages have the three noun modifier co-occurrences of VSO languages, and the three of SOV languages, and hence they have all four of the logically possible co-occurrences (compare the unattested types 4 and 8 with the attested SVO types 12 and 16, and the unattested 18 and 22 with the attested SVO types 10 and 14). Also, Greenberg (1966) gives thirteen implicational universals involving verb position (nos. 3, 4, 5, 6, 7, 10, 12, 13, 15, 16, 17, 25, 41). In all thirteen, the implicationally antecedent property is either VSO or SOV, never SVO. Hence, SVO does not correlate with any other word order properties in a unique and principled way. It merely combines properties of VSO and SOV languages, and has none of its own. SVO is, therefore, not a type indicator, and so SVO cannot be collapsed with VSO into VO.

2.1.4 *Language distributions:* Finally, the distribution of languages across Greenberg's sixteen attested types varies considerably. Some co-occurrence types are evidently more preferred than others. Although the three permitted types defined by NSP are among the four most frequent, the second most frequent type, 24, is an exception. And NSP, as currently formulated, makes no predictions concerning the relative sizes of the other attested co-occurrence types.

3. Implicational universals derivable from Appendix II

The adjective and genitive data in AII point to the following implicational generalization within the exceptions to Vennemann's NSP. If the genitive departs from the operator–operand serialization of verb and object, then so does the adjective, but the adjective may depart from this serialization without the genitive doing so. More generally, it seems that we need to set up implicational co-occurrence statements defined on specific operator–operand constructions, in order, in effect, to predict which will depart first from Vennemann's ideal NSP ordering. The result is universal statements more akin to Greenberg's earlier formulations than to Vennemann's reformulation.

An implicational 'if P then Q' statement, where P and Q are individual operator–operand constructions (e.g. NA and NG, respectively) requires that languages must not have property P without property Q (*P & −Q). If a language does have P, therefore, it must have Q (P & Q), but in the absence of P, a language may still have Q (−P & Q), or else may have neither property (−P & −Q). Three co-occurrence types are permitted, and just one is disallowed. By contrast, Vennemann's NSP consists in effect of a set of bilateral implications of the form 'if P then Q, and if Q then P', the result of which is to permit only two co-occurrences (and hence two language types): P & Q and −P & −Q (e.g. NA & NG and AN & GN).

The data of AII permit the formulation of the following three-way typologies defined by unilateral implicational statements mentioning specific operator–operand constructions.

3.1.1 *Implicational universal (I):* Consider SOV languages again. The adjective and genitive co-occurrences permit the following implication:

If the adjective precedes the noun, then the genitive precedes the noun,
i.e. AN ⊃ GN

SOV languages co-occur with:

SOV lgs:l	(P & Q)	AN & GN:	types 19 and 23
	(−P & Q)	NA & GN:	type 24
	(−P & −Q)	NA & NG:	types 17 and 21
	(*P & −Q)	*AN & NG:	no examples (potentially types 18 and 22)

Our first implicational universal is, therefore:

(I) If a language has SOV word order, and if the adjective precedes the noun, then the genitive precedes the noun, i.e. SOV ⊃ (AN ⊃ GN)

3.1.2 *Implicational universal (II)*: VSO languages exhibit the exact mirror-image pattern:

If the adjective follows the noun, then the genitive follows the noun, i.e. NA ⊃ NG

VSO languages co-occur with:

VSO lgs:	(P & Q)	NA & NG:	type 1
	(−P & Q)	AN & NG:	type 2
	(−P & −Q)	AN & GN:	types 3 and 7
	(*P & −Q)	*NA & GN:	no examples (potentially types 4 and 8)

Our second implicational universal is, therefore:[3]

(II) If a language has VSO word order, and if the adjective follows the noun, then the genitive follows the noun, i.e. VSO ⊃ (NA ⊃ NG)

3.1.3 *Implicational universal (III)*: The property Prep seems to have, throughout AII, the same conditioning effect on the co-occurrence of adjective and genitive orders as does VSO. Languages with Prep co-occur with:

Prep lgs:	(P & Q)	NA & NG:	types 1 and 9 and 17
	(−P & Q)	AN & NG:	types 2 and 10
	(−P & −Q)	AN & GN:	types 3 and 11 and 19

The co-occurrence *NA & GN (*P & −Q) does not occur together with Prep in both VSO and SOV languages:

*NA & GN: no examples of types 4 and 20

However, in SVO languages there is just one isolated example of *NA & GN

[3] No corresponding universal with SVO is possible here (SVO ⊃ (NA ⊃ NG)) since the co-occurrence SVO & NA & GN is attested in types 12 and 16 (recall Section 2.1.3). As a result VSO and SVO cannot be collapsed into a common antecedent property VO. However, notice that in addition to languages with basic VSO, there are also languages with basic VOS, e.g. Malagasy and Gilbertese (Charles Randriamasimanana, personal communication) and also languages which are best classified as verb–initial on account of the frequency of both VSO and VOS, e.g. Fijian, Samoan, and Tongan. The five languages just illustrated all have Prep & NA & NG, and so satisfy one of the co-occurrence possibilities for VSO languages as defined by universal (II). We can, therefore, convert (II) into the more general statement (II′) by changing the antecedent property VSO into V–initial:

(II′) If a language has Verb–initial order, and if the adjective follows the noun, then the genitive follows the noun; i.e. V–initial ⊃ (NA ⊃ NG).

co-occurring with Prep (Arapesh, type 12). Bearing in mind this single counter-example, we have the following (statistical) universal:[4]

(III) If a language has Prep word order, and if the adjective follows the noun, then the genitive follows the noun, i.e. Prep ⊃ (NA ⊃ NG)

Implicational universal (III) generates interesting predictions in conjunction with implications (I) and (II). Implication (I) SOV ⊃ (AN ⊃ GN) permits an SOV language to have AN & GN, NA & GN or NA & NG. The first and third of these co-occurrences, but not the second, overlap with the permitted co-occurrences of the mirror-image implication NA ⊃ NG within universal (III):

implication (I)	implication (III)
AN & GN	AN & GN
NA & GN	*NA & GN
NA & NG	NA & NG
*AN & NG	AN & NG

It follows that if a language has both SOV & Prep, then it can have neither of the starred co-occurrences. This is what we find. The only two attested SOV & Prep co-occurrences are types 17 and 19, which have NA & NG and AN & GN, respectively.

Both universals (II) and (III) define the same noun modifier co-occurrences (NA ⊃ NG), and so do not further restrict the possibilities for a language with both VSO & Prep.

3.1.4 *Implicational universal (IV):* The property Postp has a similar conditioning effect on the co-occurrence of adjective and genitive orders throughout AII, as does SOV. Languages with Postp co-occur with:

Postp lgs:	(P & Q)	AN & GN:	types 7 and 15 and 23
	(−P & Q)	NA & GN:	types 16 and 24
	(−P & −Q)	NA & NG:	type 21

The co-occurrence *AN & NG (*P & −Q for universal (I) SOV ⊃ (AN ⊃ GN)) does not occur with Postp in both VSO and SOV languages:

*AN & NG: no examples of types 6 and 22

But again, in SVO languages there are a limited number of examples with *AN & NG co-occurring with Postp (Rutulian and other Daghestan languages, type 14). We therefore have the following (statistical) universal:[5]

[4] A non-statistical version of (III) would be (III′):

(III′) If a language has Prep word order and either VSO or SOV, then if the adjective follows the noun, the genitive follows the noun also, i.e. Prep & (VSO v SOV) ⊃ (NA ⊃ NG).

[5] A non-statistical version of (IV) would be (IV′):

(IV) If a language has Postp word order, and if the adjective precedes the noun, then the genitive precedes the noun, i.e. Postp ⊃ (AN ⊃ GN)

Universal (IV) defines the same noun modifier co-occurrences as universal (I), and so does not further restrict the range of possibilities in a language with both SOV & Postp. The collective predictions of universals (II) VSO ⊃ (NA ⊃ NG) and (IV) are, however:

implication (II)	implication (IV)
NA & NG	NA & NG
AN & NG	*AN & NG
AN & GN	AN & GN
*NA & GN	NA & GN

If a language has both VSO & Postp, therefore, it cannot have either of the starred co-occurrences. It is significant that the one attested VSO & Postp type (type 7) has one of the two permitted orders, AN & GN.

3.1.5 *Summary of the implicational universals:* Our four implicational universals are (with the excluded co-occurrence types in parentheses):

(I) SOV ⊃ (AN ⊃ GN) (18, 22)
(II) VSO ⊃ (NA ⊃ NG) (4, 8)
(III) Prep ⊃ (NA ⊃ NG) (4, 12, 20)
(IV) Postp ⊃ (AN ⊃ GN) (6, 14, 22)

Notice that implications (I) and (IV) are mirror-images of (II) and (III), respectively. The exceptionless versions of (III) and (IV) (presented in footnotes [4] and [5]) are:

(III′) Prep & (VSO v SOV) ⊃ (NA ⊃ NG) (4, 20)
(IV′) Postp & (VSO v SOV) ⊃ (AN ⊃ GN) (6, 22)

We have just seen that these universals define attested versus non-attested word order co-occurrences in AII. Operating in conjunction with one another they then predict whether whole co-occurrence types in AII will be attested or not. A co-occurrence type will not be attested if any one of its constituent word orders violates any one of these implications. There are eight unattested co-occurrence types in AII. Six of these eight (types 4, 6, 8, 18, 20, 22) are ruled out by the

(IV′) If a language has Postp word order and either VSO or SOV, then if the adjective precedes the noun, the genitive precedes the noun also, i.e. Postp & (VSO v SOV) ⊃ (AN ⊃ GN).

However, according to information given me by Maner Thorpe, the SVO Daghestan languages of the Caucasus should be classified as type 15 rather than type 14 (cf. note 2), which then removes the exceptions to (IV). Hence, (IV) is already non-statistical, and is to be preferred over (IV′) since it is more general.

exceptionless implications (I), (II), (III′), (IV′). The two remaining unattested types (type 5 VSO & Po & NG & NA, and type 13 SVO & Po & NG & NA) are actually logically possible (being consistent with the independently motivated implications (II) and (IV′)), but will be predicted by our distributional principle of the next section to be rare or non-occurring in a sample of this size. The sixteen co-occurrences that are attested are all permitted by the exceptionless universals. In fact, therefore, implications (I), (II), (III′), and (IV′) allow for eighteen co-occurrences to exist, the sixteen that do in fact co-occur, plus types 5 and 13, whose possible absence is accounted for by the distributional principle. Formal proof of the predictions made by these implications, using truth tables to illustrate exactly which values of P and Q, etc., are permitted to co-occur, is given in Hawkins (1978).

3.2 *General properties of adequate implicational universals*

We have seen that a small number of either exceptionless or practically exceptionless implicational statements can account for a large body of data. As a result, many of Greenberg's original statistical implications are unnecessarily imprecise:

> With overwhelmingly more than chance frequency, languages with dominant word order VSO have the adjective after the noun (Greenberg, 1966: 85).

Co-occurrence types 2, 3, and 7 provide exceptions in the form of VSO & AN co-occurrences. Similarly,

> In languages with prepositions, the genitive almost always follows the governing noun, while in languages with postpositions it almost always precedes (Greenberg, 1966: 78).

Co-occurrence types 3, 11, 12, and 19 provide exceptional Prep & GN co-occurrences, while types 14 and 21 provide exceptions with Postp & NG. In our formulations greater accuracy is achieved by using multi-valued implicational statements. Instead of relating just two properties to one another, as Greenberg generally does (if a language has some single word order P, then it also has some single word order Q), my statements involve at least three properties (if a language has some word order P, and if it has word order Q, then it will also have word order R). By, in effect, increasing the conditioning property P from one word order specification to at least two, it becomes possible to predict the co-occurrence of a third word order in a straightforward way, and to reveal significant generalizations about co-occurring word orders.

There are also three general theoretical reasons why the overuse of statistical universals should be avoided. First, to the extent that the available data reveal that certain co-occurrence types are totally unattested, we must attempt to distinguish as restrictively as possible between the attested and the unattested co-occurrences. But this requires using stronger, exceptionless universals, rather than tendency-based statements. These latter necessarily allow the offending *P & −Q language types to

exist, albeit as a minority type, even though no such languages have yet been attested. Of course, if such languages are subsequently found to exist in a different or larger sample, then our universals must be reformulated accordingly, or abandoned. But at each stage in the research process we must attempt to predict the balance between attested and non-attested co-occurrences as accurately, and simply, as possible. And this requires the regular use of stronger, exceptionless statements, rather than weaker statistical ones.

Notice that exactly this point is assumed also by Keenan and Comrie (1977) in their formulation of implicational universals of relativization. Their hierarchy constraints are NOT statistical approximations of the form 'if a primary relativization strategy operates on oblique NPs, then with greater than chance frequency it operates on indirect and direct objects as well', or 'identical relativization strategies in a language will in general operate on a continuous segment of the Accessibility Hierarchy'. They are non-statistical claims which account for all and only the co-occurring relativization possibilities in the languages of their sample. Exceptions are clearly recognizable and would lead to a reformulation of their claims (and have in fact already done so: compare the 1972 and 1977 versions of their paper). The data of this present paper suggest that there is no reason in principle why the rigour of Keenan and Comrie's relativization implications cannot be extended to suitably reformulated universals defining word order co-occurrences. In both cases, currently available data samples indicate that some co-occurrences of linguistic phenomena are attested, while others are not.[6]

Secondly, Greenberg uses statistical implicational universals in order to make what are in essence distributional statements, distinguishing, e.g. very frequent

[6] It might be objected, however, that 'basic' word orders are less easily recognizable, cross-linguistically, than the relativization strategies and grammatical relations of the Accessibility Hierarchy, and that this jeopardizes the validity of precise, non-statistical implications. But this objection misses the point. First, both statistical and non-statistical universals of word order assume equally the recognizability of basic VSO, NA, etc., and any difficulties in this regard are potentially damaging to both. Second, I can see no difference in principle between the problems raised in defining 'relativization on subject', 'on direct object', 'on indirect object', and 'separate relativization strategy' etc., on the one hand, and 'basic' word order on the other. Third, the question of the logical form and status of the descriptive statements of Universal Grammar is independent of questions of definition for the major cross-language categories. To the extent that we even suspect that certain logically possible co-occurrences of linguistic phenomena are unattested, we have justified the need for exceptionless implicational statements defining all and only the attested facts. Fourth, there seems to be an implicit view held by some linguists that if some language has, say, both VSO and SVO, and it is difficult to decide which is basic, then this is an exception to the implicational universals. But this is not necessarily so. It may be an exception to dogmatic assertions that all languages will have either basic VSO, SVO, or SOV word order. But the implications state that 'if a language has basic word order P, then it also has basic word order Q'. In the event that some languages do not clearly have basic word order P, then no predictions are made for these languages simply, since the antecedent property is not satisfied.

Without wishing, therefore, to minimize the need for tighter and more adequate definitions of cross-language categories, I see good reason for elevating the universals of word order to the same degree of precision and falsifiability proposed by Keenan and Comrie for relativization. And I see nothing but loose theorizing and misguided generalizations resulting from a failure to do so (cf. Hawkins (1979) for a discussion of such loose theorizing within a historical context).

co-occurrences such as VSO & NA, from less frequent ones, VSO & AN. But we shall argue in section 4.7 that implicational statements of this kind are not the appropriate device with which to capture regularities of language distribution. Hence, if distributional generalizations are more adequately captured in another way, the motivation for statistical implicational statements disappears.

Thirdly, there is an important reason why it is NOT unrealistic to expect that exceptionless universals of word order will be both possible and productive relative to data samples larger than those presented in Greenberg (1966). It is a familiar finding in word order studies that the attested co-occurrences are regularly significantly fewer than the mathematically possible co-occurrences. To give a simple example, Steele (1975) found, using a genetically broad sample of languages, that the only positions which a modal element could occupy in SVO, VSO, and SOV languages were the following:

SModVO		
ModVSO	VModSO	
SOVMod	SModOV	ModSOV

There are four mathematically possible positions which a modal could occupy in each of these three language types (e.g. ModSVO, SModVO, SVModO, SVOMod in SVO languages) but only six of the twelve possibilities are actually attested.

And more generally, the mathematically possible co-occurrences of all operators and operands throughout a language are quite enormous. In any language the verb may potentially precede or follow the other major constituents of the sentence (subject, direct object, indirect object, oblique NPs, etc), and the noun may potentially precede or follow the determiner, adjective, genitive, relative clause, and so on, within the NP. If one calculates the mathematically possible combinations of all verb positions with all noun positions (e.g. VSO may combine with either noun–initial, noun–second, noun–third order, etc., within the NP, and so may SVO, non-rigid SOV, rigid SOV), and if one then adds in the possible co-occurrences of the adposition within the adposition phrase, of the adjective within the adjective phrase, and of the auxiliary elements relative to the verb, etc., the mathematical combinatorial possibilities that languages might potentially exhibit are indeed very large. Yet the attested word order variation patterns that we find languages actually exhibiting are not nearly so large. Time and time again, certain word orders are found to co-occur, while others are not. And both my own research, and my reading of the research of others, have convinced me that many more of these mathematical possibilities are UNattested than are attested. Such considerations provide a further motivation for avoiding statistical in favour of non-statistical implicational universals.

Moreover, one major reason for this discrepancy between the mathematical possibilities and what is actually attested is that many cross-language word order co-occurrence possibilities, just like relativization possibilities, appear to be

arranged in hierarchies. For example, Hawkins (1979) argues that the antecedent property Prep in (III) Prep $\supset$ (NA $\supset$ NG) is actually the antecedent for a whole chain of overlapping implications, as in (V):

(V) Prep $\supset$ ((NDem $\supset$ NA) & (NA $\supset$ NG) & (NG $\supset$ NRel))
i.e. P $\supset$ ((Q $\supset$ R) & (R $\supset$ S) & (S $\supset$ T))

the effect of which is to define a hierarchy of noun modifier co-occurrence possibilities for prepositional languages. Specifically, prepositional languages are claimed to exhibit the following five word order co-occurrences only:

(i)	Prep & NDem & NA & NG & NRel	(i.e. P & Q & R & S & T)
(ii)	Prep & DemN & NA & NG & NRel	P & −Q & R & S & T
(iii)	Prep & DemN & AN & NG & NRel	P & −Q & −R & S & T
(iv)	Prep & DemN & AN & GN & NRel	P & −Q & −R & −S & T
(v)	Prep & DemN & AN & GN & RelN	P & −Q & −R & −S & −T

But there are $2^4 = 16$ mathematically possible co-occurrences of Q, R, S, and T properties in type P (prepositional) languages, 11 of which are claimed to be impossible. If we were to add just one more implication to this hierarchy, for example:

$$P \supset ((Q \supset R) \,\&\, (R \supset S) \,\&\, (S \supset T) \,\&\, (T \supset V))$$

the result would be only one more permissible co-occurrence (making six in all), and $2^5 = 32$ mathematical possibilities. Because of the overlapping nature of these implications, therefore, the co-occurrence possibilities available to a type P language are very limited, and the impossible co-occurrences very many. And as more such implicational dependencies are found between linguistic properties, so the discrepancy between the attested co-occurrences and the mathematical possibilities becomes quite enormous.

Notice, finally, that we saw ample evidence in Section 3.1 for the need for unilateral implications (if P then Q), rather than bilateral implications (if P then Q, and if Q then P, as in Vennemann's theory). Unilateral implications define three-way typologies, whereas bilateral implications define two-way typologies. The permitted co-occurrences of, for example, AN $\supset$ GN (P $\supset$ Q) in SOV languages are precisely the three which correspond to the values P & Q, −P & Q and −P & −Q (with *P & −Q ruled out). But in Vennemann's operator–operand schema (i.e. NSP), the logical status of all Greenberg's unilateral implications has been changed into bilateral, reversible implications permitting just two co-occurrences, P & Q and −P & −Q. In effect, P & Q languages serialize their operator–operand structures in one order, while −P & −Q languages serialize them in the reverse order. There is no place in this system for the third co-occurrence type, −P & Q, whose existence we have seen to be both valid and productive. Hence, there is a systematic failure in Vennemann's work to distinguish −P & Q from *P & −Q co-occurrences, and this

accounts for most of the exceptions to his principle. The data of Section 3.1 reveal many three-way rather than two-way typologies.

Summarizing, implicational universals of word order should be preferably non-statistical rather than statistical, multi-valued rather than bi-valued, and unilateral rather than bilateral, if they are to distinguish adequately between attested and non-attested word order co-occurrences.

3.3 *Explaining the implicational universals*

Universals (I), (II), (III), and (IV) all have the logical form $P \supset (Q \supset R)$. In each case, P, Q, and R have consistent operator–operand orderings. For example, where P = SOV, Q = AN, and R = GN, there is a consistent operator before operand serialization. And Q and R are constructions with one and the same operand (noun). These formulations permit languages with P to co-occur with the following values for Q and R:

P & Q & R
P & −Q & R
P & −Q & −R

*P & Q & −R

The property P does not, therefore, impose its own serialization pattern on Q and R, necessarily. It defines possible and impossible co-occurrences. Both the operators upon the noun may be consistent with the serialization of P, or they may both be inconsistent with it. But if just one of these operators departs from P's serialization, it is always Q (the adjective) and not R (the genitive).

The following is a plausible explanation, which is part grammatical and part historical. The adjective is syntactically and semantically less complex than the genitive. The latter consists of a whole NP modifying another NP, in conjunction often with a genitive marker. Thus, the genitive may consist of a nominal referring predicate plus all the syntactic–semantic material which may occur within an NP, adjectives, determiners, etc. But the adjective is a single descriptive predicate, and is both less complex and morphologically and syntactically shorter than a genitive. If, in the course of its history, a language is going to change the order of its noun modifiers, therefore, away from the serialization pattern defined by property P, it can reorder its adjective with greater ease and with less drastic restructuring of the NP, than by moving the whole genitive expression. Given the gradualness of language change, the prior shifting of adjectives away from the serialization pattern of P would seem to be a natural consequence of the relative semantic–syntactic complexity distinguishing the two. This historical order of events (which is supported by the historical data of Hawkins (1979)) is then reflected in the co-occurrences of languages currently available for synchronic inspection. In many, the adjective alone is at variance with the serialization of P, while the genitive is at variance with P only if the adjective is as well.

More generally, I would suggest that a similar relative complexity explanation underlies the implicational chains or hierarchies exemplified by (V) above. With each implication, NDem ⊃ NA, NA ⊃ NG, and NG ⊃ NRel, the antecedent property is arguably less complex and morphologically and syntactically shorter than the consequent property, and again the antecedent is comparatively and historically less stable than the consequent property, and departs first from the serialization pattern defined by property P (prepositions).

Recall now that universals (III) and (IV) have a small number of exceptions in SVO languages. What is significant about universals (I), (II), (III), (IV), and also (II′), (III′) and (IV′) (cf. notes 3–5) is that the ultimate antecedent P properties, SOV, VSO, V-initial, Prep, and Postp are all what we might call 'operand-peripheral'. The verb occurs either leftmost or rightmost in the sentence, and the adposition is either leftmost or rightmost within the adposition phrase. These operand-peripheral orders seem to exert strong co-occurrence requirements on their noun modifiers. By contrast, we have seen that SVO, in which the operand is not peripheral in the sentence, does not exert any word order co-occurrence requirements at all. There are no implicational universals of word order (in either Greenberg's or my own formulations) which have SVO as the ultimate antecedent (P) property. And it is possibly this non-predictive nature of SVO which underlies also the limited number of exceptions to (III) and (IV), precisely in SVO languages.

Because of this typological ambivalence of SVO languages, verb position becomes a bad indicator of language type. In general we have seen that the notion of a language type no longer means uniform serialization for all operator–operand categories. It means instead that the languages in question conform to a restricted set of co-occurrence possibilities defined by our implicational universals. If we generalize from these implications and say, as Greenberg did, that verb position is the major type indicator, we define into existence three language types, VSO, SVO, and SOV, one of which (SVO) has no type characteristics. But the notion of a type becomes meaningless if there are no regular correlations with the selected type indicator. The ambivalence of SVO, therefore, destroys the generality of a verb-based typology of word order. Although VSO and SOV do give good correlations, SVO, the second most frequent verb position, does not. Much better type indicators are Prep and Postp. These word orders make genuine co-occurrence predictions possible across all the languages of Greenberg's sample. We, therefore, propose that there exist two major language types: prepositional and postpositional. Both exert co-occurrence requirements on adjective and genitive orders, but only distributional preferences, as we shall see, in favour of VSO, SVO, or SOV.

Finally, I attribute the fact that there are ultimately just two major or basic language types, obeying different implicational laws, to the fact that there are only two basic positions that a modifier can adopt in relation to its head: before or after. Where the number and nature of modifiers is semantically and syntactically limited (as with adposition modifiers) we get just the two possible orders: Prep and Postp.

But where the number of modifiers is much greater (e.g. with the noun) we get a whole set of three-way co-occurrences between each pair of implicationally related modifiers, reflecting the relative semantic–syntactic stability of the members of each pair.

4. THE DISTRIBUTIONAL GENERALIZATION

Having set up implicational universals distinguishing attested from non-attested word order co-occurrence types in Greenberg's AII, we can now show that the rather uneven distribution of languages across the attested types is principled. Underlying the relative sizes of the word order co-occurrence types there emerges a distributional generalization which I shall call the principle of 'Cross-Category Harmony' (CCH).

4.1 *The basic insight*

As a preliminary illustration notice that there is a regular decline in the numbers of attested languages as the adjective and genitive operators depart from the serialization pattern of verb and adposition. Universals (I) SOV ⊃ (AN ⊃ GN) and (IV) Postp ⊃ (AN ⊃ GN) permit the following co-occurrence possibilities:[7]

23.	SOV & Po & AN & GN	28 lgs	decline
24.	SOV & Po & NA & GN	24 lgs	
21.	SOV & Po & NA & NG	7 lgs	

Universals (II) VSO ⊃ (NA ⊃ NG) and (III) Prep ⊃ (NA ⊃ NG) permit the following:

1.	VSO & Pr & NA & NG	19 lgs	decline
2.	VSO & Pr & AN & NG	5 lgs	
3.	VSO & Pr & AN & GN	1 lg	

Prepositional languages with SVO pattern just like VSO & Pr languages:

9.	SVO & Pr & NA & NG	21 lgs	decline
10.	SVO & Pr & AN & NG	8 lgs	
11.	SVO & Pr & AN & GN	3 lgs	

If the operators on the verb and on the adposition are all preposed (SOV & Po), then the most favoured languages are those in which the operators on the noun are all

[7] Type 23 is actually genetically and areally more diverse relative to type 24 than these figures indicate. The decision to count language groups as individual languages has given to type 24 a disproportionately high figure. The figures for my expanded sample (cf. note 2) are approximately:

Type 23:	80 lgs	decline
Type 24:	50 lgs	
Type 21:	11 lgs	

which gives a decline somewhat more in line with the VSO & Prep and SVO & Prep figures.

preposed as well (AN & GN), the next favoured languages have only one noun operator postposed (NA & GN), and the least favoured have two postposed (NA & NG). And if the operators on the verb and adposition are postposed (VSO/SVO & Pr), then the most favoured languages are those with both operators on the noun postposed (NA & NG), the next favoured have only one noun operator preposed (AN & NG), and the least favoured have both preposed (AN & GN).

Evidently, there are quantifiable preferences for the position of the noun in relation to its operators to mirror the position of the verb and the adposition in relation to their operators. As the order of these operands across their respective categories comes into increasing conflict, the number of exemplifying languages decreases. We might, therefore, hypothesize that: the more similar the position of operands relative to their operators across the different operand categories, the greater will be the number of exemplifying languages, and the more dissimilar is this ordering, the fewer will be the exemplifying languages.

How can we test such a hypothesis and make it more precise? Notice first that the quantitative data considered so far relate the position of the noun to both verb and adposition orders. But this kind of quantification gives us no means of checking whether it is crucially the verb position alone, the adposition order alone, or some combination of both which results in fewer exemplifying languages as a consequence of the conflict with noun position. We must, therefore, first compare noun position with verb position separately, and noun position with adposition order separately. More generally, given that we have three distinct operand categories represented in AII, noun, verb, and adposition, we must compare noun with verb, noun with adposition, and also verb with adposition order before considering all three together (cf. Section 4.5).

Consider, for example, verb and adposition orders in relation to one another. The following figures show the percentages of VSO, SVO, and SOV languages which are prepositional and postpositional respectively:

Pr/VSO = 25/26 = 96.15%	decline	Po/VSO = 1/26 = 3.85%	decline
Pr/SVO = 33/51 = 63.46%		Po/SVO = 18/51 = 36.54%	
Pr/SOV = 5/64 = 7.81%		Po/SOV = 59/64 = 92.19%	

As the verb moves to the right in the sentence (VSO > SVO > SOV) it becomes increasingly disharmonic with prepositions, since the increased preposing of operators on the verb conflicts with the postposing of operators on the adposition (i.e. prepositions). This increasing disharmony is reflected in the decreasing percentage numbers of the languages concerned. Correspondingly, as the verb moves leftwards (SOV > SVO > VSO), the postposing of operators on the verb becomes increasingly disharmonic with the preposing of operators on the adposition (i.e. postpositions), which results again in decreasing numbers of the relevant languages.

The use of percentages rather than actual language numbers requires comment. In general we are not primarily interested in the actual number of, say, VSO as

opposed to SVO or SOV languages in the sample. Our interest lies in the proportions of each of these groups which co-occur with the different positions of noun and adposition in relation to their operators, for we suspect that these proportions reflect the operator–operand ordering balance across the three operand categories. Thus, the fact that the actual number of VSO & Pr languages (= 25) is less than the actual number of SVO & Pr languages (= 33) should not be taken to indicate that the orders SVO and Pr exhibit a stronger attraction towards one another than VSO and Pr. Quite the reverse. What is important is that the 25 VSO & Pr languages are 25 out of only 26 possible VSO languages, whereas the 33 SVO & Pr languages are 33 out of a total of 52 SVO languages. Whereas VSO co-occurs almost exclusively with prepositions alone, therefore, SVO co-occurs with both prepositions and postpositions (in the ratio 2 to 1). By using percentages rather than actual language totals we quantify more accurately the mutual attraction between two word orders by taking into account both the number of languages actually exhibiting some word order co-occurrence, and the potential total that might have had this co-occurrence within the sample.

Notice also that the distributional evidence of this section suggests, contra Vennemann, that subjects should also be regarded as operators on verbs as operands. Prepositions do not occur as frequently in SVO languages as they do in VSO languages. If subjects are operators, then their preposing before the verb produces an operator–operand ordering, harmonic with postpositions. SVO languages would, therefore, have one solid operator before operand ordering, even though the other operators on the verb are all generally postposed. The resulting predominance of prepositions in SVO languages would reflect the majority serialization relative to the verb: operand before operator. But the operator before operand ordering of the subject would explain why roughly one third of SVO languages are nonetheless postpositional, while hardly any VSO languages are.

Theoretical support for regarding the subject as an operator just like a direct object, is not lacking. Semantically, a subject NP in transitive clauses is one argument of the predicate among others, these others being in Vennemann's system, operators. Considerations of semantic parallelism would, therefore, argue for the operator status of subjects. And syntactically, Vennemann's claim that subjects cannot be analysed as operators since the result of the operation is a sentence, S, rather than a complex V or VP, is undermined by the $\bar{X}$ Theory, in which the category S is replaced by $\bar{\bar{V}}$. In this theory, a subject NP ($\bar{\bar{N}}$) takes a $\bar{\bar{V}}$ (or VP) and makes a $\bar{\bar{V}}$, i.e. a more complex category of the same type as that of the operand, $\bar{\bar{V}}$.

This reanalysis of subjects, for which the distributional data provide the initial motivation, now results in an interesting generalization. Greenberg (1966) proposed two types of SOV languages: rigid and non-rigid. In the former, exemplified by Japanese, the verb is rigidly final in the clause, while in the latter, exemplified by Basque, at least one operator on the verb typically follows the verb as the basic

order. The VSO/SVO distinction can now be seen to mirror the rigid/non-rigid SOV dichotomy. VSO languages have basic verb–initial order, while SVO languages have one operator on the verb before the verb.

In fact, there seems to be a continuum from verb–initial order at one end to rigid verb–final order at the other. Languages with the following basic verb positions are attested:

(i)	V S DO IO OBL ADV]	VSO
(ii)	S V DO IO OBL ADV]	SVO
(iii)	S DO V IO OBL ADV]	
(iv)	S IO DO V OBL ADV]	SOVnon-rigid
(v)	S OBL IO DO V ADV]	
(vi)	S ADV OBL IO DO V]	SOVrigid

(i) and (ii) are VSO and SVO languages, respectively. Some languages of type (iii) are discussed by, for example, Givon (1975). Types (iii), (iv), and (v) represent different degrees of non-rigid verb-finality, while type (vi) languages are rigidly verb–final.

4.2 *Further cross-categorial word order co-occurrence quantities*

In Table 2 we list co-occurrence frequencies for all pairs of word orders in AII. For each pair, e.g. Postp & SOV, we give the percentage of SOV languages which are postpositional, and also the percentage of postpositional languages which are SOV, followed by a total of these two percentages (maximum 200). By totalling these two percentages we quantify the attraction which each word order exerts towards the co-occurrence of the other. Thus, if all SOV languages had postpositions, and all postpositional languages were SOV, the total would be 200. To the extent that any total is less than 200, we quantify the degree to which the mutual attraction of two word orders is less than perfect. Where one percentage total exceeds another, we know that the word order combination in question occupies a quantitatively larger slice of the potentially available languages than does a word order combination with a lower total. For the higher these totalled percentages, the greater are the numbers of languages with the co-occurrence in question, relative to the total number having each word order.

The fact that the percentage total for Postp & SOV exceeds that of Postp & SVO, therefore, and also that of Postp & VSO, Prep & SVO and Prep & SOV, means that the total frequency with which Postp and SOV select one another exceeds the frequency with which these other verb and adposition orders select one another. Hence, the mutual attraction of Postp & SOV is greater than for Postp & SVO, etc.

Proceeding from top to bottom in Table 2, the percentage totals decline from Postp & SOV to Postp & SVO to Postp & VSO, and from Prep & VSO to Prep & SVO to Prep & SOV. Postp & SOV and Prep & VSO represent the ideal states: the operators on both verb and adposition all precede their operands in the first case,

				Total
Postp & SOV:	SOV/Po	= 59/79 = 74.68%	Po/SOV = 59/64 = 92.19%	166.87
Postp & SVO:	SVO/Po	= 18/79 = 24.05%	Po/SVO = 18/52 = 36.54%	60.59
Postp & VSO:	VSO/Po	= 1/79 = 1.27%	Po/VSO = 1/26 = 3.85%	5.12
Prep & VSO:	VSO/Pr	= 25/63 = 39.68%	Pr/VSO = 25/26 = 96.15%	135.83
Prep & SVO:	SVO/Pr	= 33/63 = 52.38%	Pr/SVO = 33/52 = 63.46%	115.84
Prep & SOV:	SOV/Pr	= 5/63 = 7.94%	Pr/SOV = 5/64 = 7.81%	15.75
Prep & N–:	N–/Pr	= 44/63 = 69.84%	Pr/N– = 44/51 = 86.27%	156.11
Prep & –N–:	–N–/Pr	= 14/63 = 22.22%	Pr/–N– = 14/51 = 27.45%	49.67
Prep & –N:	–N/Pr	= 5/63 = 7.94%	Pr/–N = 5/40 = 12.5%	20.44
Postp & –N:	–N/Po	= 35/79 = 44.3%	Po/–N = 35/40 = 87.5%	131.8
Postp & –N–:	–N–/Po	= 37/79 = 46.84%	Po/–N– = 37/51 = 72.55%	119.39
Postp & N–:	N–/Po	= 7/79 = 8.86%	Po/N– = 7/51 = 13.73%	22.59
VSO & N–:	VSO/N–	= 19/55 = 37.25%	N–/VSO = 19/26 = 73.08%	110.33
VSO & –N–:	VSO/–N–	= 5/51 = 9.8%	–N–/VSO = 5/26 = 19.23%	29.03
VSO & –N:	VSO/–N	= 2/39 = 5%	–N/VSO = 2/26 = 7.69%	12.69
SVO & N–:	SVO/N–	= 21/51 = 41.18%	N–/SVO = 21/52 = 40.38%	81.56
SVO & –N–:	SVO/–N–	= 22/51 = 43.14%	–N–/SVO = 22/52 = 42.31%	85.45
SVO & –N:	SVO/–N	= 9/40 = 22.5%	–N/SVO = 9/52 = 17.31%	39.81
SOV & –N:	SOV/–N	= 29/40 = 72.5%	–N/SOV = 29/64 = 45.31%	117.81
SOV & –N–:	SOV/–N–	= 24/51 = 47.06%	–N–/SOV = 24/64 = 37.5%	84.56
SOV & N–:	SOV/N–	= 11/51 = 21.57%	N–/SOV = 11/64 = 17.19%	38.76

Table 2

Percentage co-occurrence quantities for word order pairs[8]

N– is an abbreviation for noun–initial, i.e. noun before both adjective and genitive; –N stands for noun–final, i.e. noun after both; while –N– represents noun–medial, i.e. one operator precedes and one follows. Which operator precedes and which follows will depend on the accompanying verb and adposition order. Each language type, prepositional or postpositional, permits at most one noun–medial possibility.

[8] Although the co-occurrence totals shown here exhibit the predicted relative differences in size, the extent of these relative differences is not always the same from one set of co-occurrences to another. I believe there is a reason for this. Greenberg's AII gives only a small sample of the total number of operators which can modify their respective operands. When the predictably co-occurring operator orders omitted from this sample are taken into consideration, the discrepancies in relative size become

and all follow in the second. But as the imbalance of operator preposing to postposing increases across these two operand categories, the percentage totals then decline.

The decline from Prep & N– to Prep & –N– to Prep & –N, and from Postp & –N to Postp & –N– to Postp & N– is explainable in a similar way. In Prep & N– languages the operators on both the adposition and the noun are postposed. The

principled. Thus, the decline from the *Postp & SOV* co-occurrence to the *Postp & SVO* co-occurrence is from 166.87 to 60.59, whereas the decline from *Prep & VSO* to *Prep & SVO* is from 135.83 to 115.84. Evidently, the mutual attraction of *Prep & SVO* is closer to that of *Prep & VSO*, than is *Postp & SVO* to *Postp & SOV*. In Section 4.1 we saw that languages with SVO have all operators on the verb postposed, except for the subject. It follows that SVO languages are still very harmonic with prepositional phrases, in which all the operators on the adposition are postposed. Conversely, although the varying degrees of verb finality are harmonic with postpositions, SVO is not. SVO is, therefore, the tip of an iceberg full of predictably co-occurring postposed operators on the verb, the effect of which is to make an SVO language only marginally less harmonic with prepositions than VSO. And the co-occurrence quantities thus appear to reflect the operator orderings which are absent from the sample and which are independently known to co-occur in languages having the sample properties selected.

There is a similar discrepancy between the decline from *Prep & N–* to *Prep & –N–* (156.11 to 49.67), and from *Postp & –N* to *Postp & –N–* (131.8 to only 119.39). Thus, the mutual attraction between *Postp & –N–* is much closer to that between *Postp & –N*, than is *Prep & –N–* to *Prep & N–*. In Hawkins (1979) I show that the noun–medial co-occurrence which is permitted in prepositional languages (AN & NG by implication (III)) implies the simultaneous preposing of numerous other operators on the noun (determiners, numerals, possessive adjectives). Hence, a Prep & AN & NG language necessarily has many noun operators preceding the noun, whereas all the operators on the adposition follow the adposition. By contrast, the logically permitted noun–medial co-occurrence in postpositional languages (NA & GN by implication (IV)) does not actually require any other operators to be postposed after the noun, although it does not rule out this possibility. Thus, a postpositional language like Burmese (SOV) has NA as its unique postposed nominal modifier. Hence, the noun–medial co-occurrence NA & GN is more harmonic with postpositions, than is AN & NG with prepositions. Postp & NA & GN is indicative of a language state in which the noun may still be predominantly final within the NP, while the adposition is final within the adposition phrase. But Prep & AN & NG is indicative of much preposing of operators on the noun, which contrasts with the postposing of operators on the adposition.

The discrepancy between *VSO & N–* to *VSO & –N–* (110.33 to 29.03) and *SOV & –N* to *SOV & –N–* (117.81 to only 84.56) mirrors that between *Prep & N–* to *Prep & –N–* and *Postp & –N* to *Postp & –N–*. The permitted noun–medial co-occurrence in VSO languages (AN & NG by implication (II)) again implies the co-occurrence of other preposed operators on the noun, which creates disharmony with verb–initial position. But SOV & NA & GN does not necessarily imply the co-occurrence of other postposed operators on the noun, and hence such a language state can be relatively harmonic with SOV (particularly since many of the SOV languages in AII are non-rigid).

The SVO co-occurrence quantities with noun position are complicated both by the fact that SVO languages may be regularly prepositional or postpositional, and by the absence of unique noun modifier co-occurrences with SVO. As a result, –N– may refer to both AN & NG and to NA & GN, which rules out possible inferences about co-occurring noun operator orders in languages which are *SVO & –N–*. Nonetheless, comparison of VSO and SVO with noun position co-occurrences reveals regular patterns. *SVO & N–* is less frequent (81.56) than *VSO & N–* (110.33), on account of the preposing of the subject in SVO languages. Conversely, the extreme infrequency of *VSO & –N–* (29.03) is relieved in *SVO & –N–* languages (85.45), although the fact that –N– in prepositional SVO languages (AN & NG) implies the co-occurrence of other preposed operators on the noun creates disharmony with the single operator before the verb (the subject) and prevents the overall SVO & –N– figure from being very high. Finally, the disharmony between *VSO & –N* (12.82) is predictably less extreme in *SVO & –N* languages (39.81), again on account of the preposing of the subject.

preposing of first one operator on the noun (Prep & –N–), and then of two (Prep & –N), produces operator–operand orderings in the NP progressively at variance with those in the adposition phrase. As this imbalance increases, the percentage totals decrease. Similarly, in Postp & –N languages the preposing of operators on the adposition matches the preposing of operators on the noun. As one operator on the noun is postposed (Postp & –N–), and then two (Postp & N–), the operator ordering imbalance increases across the categories, and as it does so the percentage correlation totals decrease.

The postposing of operators on the verb in verb–initial languages is most harmonic with the postposing of operators on the noun (i.e. N–), less harmonic with the preposing of one noun operator (–N–), and least harmonic with the preposing of two (–N). The declining percentage totals reflect this. Conversely, the preposing of operators on the verb in verb–final languages is most harmonic with the preposing of noun operators (SOV & –N), less harmonic with the postposing of one noun operator (SOV & –N–), and least harmonic with the postposing of two (SOV & N–). The declining percentage totals again reflect this. The co-occurrence quantities for SVO and noun position are more difficult to interpret on account of the fact that SVO languages regularly have many more operators on the verb after the verb in addition to the direct object (indirect objects, etc), and only the subject precedes the verb. Hence, although the verb is medial in relation to S and O, which are the only operators on the verb considered in this data sample, the verb still stands in a predominantly leftward position in the sentence as a whole in SVO languages (cf. note 8). However, for the purposes of the data at hand we must, for consistency's sake, interpret SVO & –N– as being most balanced, having one operator before both verb and noun, and one operator after. The postposing of both operators on the noun after the noun (SVO & N–) distorts the ideal verb–medial/noun–medial match-up, as does the preposing of both operators on the noun before the noun (SVO & –N). The predicted decline in percentage totals from SVO & –N– to SVO & –N (85.45 to 39.81) is convincing, but the decline from SVO & –N– to SVO & N– (85.45 to 81.56) is not, although it is not a counterexample. We attribute this marginal result to the kinds of considerations discussed in note 8.

Table 2 does not present all the predictions for relative percentage totals that can be made by our emerging principle of CCH. Before we present the full predictions (in Tables 3, 4, and 5) let us formulate the principle itself.

4.3. *The principle of Cross-Category Harmony (CCH)*

The principle which we have been developing so far can be summarized informally as follows:

> The more similar the position of operands relative to their operators across different operand categories considered pairwise (verb in relation to adposition order, noun in relation to adposition order, verb in relation to noun order), the greater are the percentage numbers of exemplifying languages.

More precisely, we have been comparing throughout two word order co-occurrence pairs, W and W′, where W comprises word orders A & B, and W′ comprises A′ & B′. (For example, W = Postp & SOV (A & B), and W′ = Postp & SVO (A′ & B′), i.e. in this example A = A′ and B ≠ B′). All four word orders, A, A′, B, and B′, can be regarded as ordered sets of grammatical categories satisfying the following conditions. Sets A and A′ have the same categories as members: one operand, a, and at least one operator upon a (designated by lower case letters late in the alphabet, x, y, z, etc). And sets B and B′ have the same categories as members: one operand, b, where b ≠ a, and at least one operator upon b (similarly designated by late lower case letters). The ordering of operand to operators differs either between A and A′, or between B and B′, or, as we shall see below, between both.

The illustrative example above can be set out as follows:

Co-occurrence pair W=Postp & SOV (A & B)
Co-occurrence pair W′=Postp & SVO (A′ & B′)

A:	⟨x, a⟩	i.e. NP+Adposition (x = Adp)
A′:	⟨x, a⟩	i.e. NP+Adposition (x = NP, a = Adp)
B:	⟨y, z, b⟩	i.e. S+O+V (y = S, z = O, b = V)
B′:	⟨y, b, z⟩	i.e. S+V+O (y = S, b = V, z = O)

Between sets A ⟨x, a⟩ and B ⟨y, z, b⟩ there is greater cross-category harmony than between A′ ⟨x, a⟩ and B′ ⟨y, b, z⟩, since operands a and b are both rightmost in A & B, whereas in A′ & B′ the rightmost position of a is not matched by b. ⟨x, a⟩ and ⟨y, b, z⟩, therefore, deviate by one position, and ⟨x, a⟩ and ⟨b, y, z⟩ (= VSO) would deviate by two.

Similarly, consider the following paired sets:[9]

A:	⟨w, x, a⟩	i.e. –N (w = G, x = A, a = N)
A′:	⟨w, a, x⟩	i.e. –N– (w = G, a = N, x = A)
B:	⟨y, z, b⟩	i.e. SOV (y = S, z = O, b = V)
B′:	⟨y, z, b⟩	i.e. SOV (y = S, z = O, b = V)

Sets A ⟨w, x, a⟩ and B ⟨y, z, b⟩ are maximally harmonic, since their respective

[9] We are not strictly interested in this context in the relative order of the operators to one another, but only in the order of each single operator relative to its operand. Thus, it makes no difference for our distributional calculations whether –N stands for ⟨w, x, a⟩ or ⟨x, w, a⟩, as long as both w and x precede a. Our predictions relate to the number of operators which precede and the number which follow the operand, irrespective of the sequencing of the operators. We should, therefore, more accurately represent –N as a set of ordered subsets: {⟨w, a⟩, ⟨x, a⟩}. Similarly, we make no distinction between VSO ⟨b, y, z⟩ and VOS ⟨b, z, y⟩ (for the purposes of CCH), and should more accurately collapse these as: {⟨b, y⟩, ⟨b, z⟩}. However, it will greatly improve readability if we do NOT complicate the formalism in this way. And we shall retain ⟨w, x, a⟩ as a shorthand for {⟨w, a⟩, ⟨x, a⟩}, understanding thereby that the relative ordering defined holds only for each operator in relation to its operand, and not necessarily for each operator in relation to each other operator. I shall, nonetheless, attempt in these formulae to order the operators relative to one another in the way that most languages do in fact order them (e.g. VSO as opposed to VOS).

operands are both rightmost. Sets A′ ⟨w, a, x⟩ and B′ ⟨y, z, b⟩ differ by one operand position. ⟨a, w, x⟩ (N–) and ⟨y, z, b⟩ (SOV) would differ by two. Similarly, ⟨w, a, x⟩ (–N–) is most harmonic with ⟨y, b, z⟩ (SVO). ⟨w, x, a⟩ (–N) would differ from ⟨y, b, z⟩ (SVO) by one operand ordering, as would ⟨a, x, w⟩ (N–). ⟨b, y, z⟩ (VSO) would also differ from ⟨w, a, x⟩ (–N–) by one operand ordering, as would ⟨y, z, b⟩ (SOV).

Notice now that the number of operators is not always the same for each operand in our calculations. In the present sample of properties, we are considering two operators on the verb, two on the noun, but only one on the adposition. Thus, even though ⟨x, a⟩ (Postp) is similar to ⟨y, b, z⟩ (e.g. SVO) in that one operator precedes the operand in both cases, yet there is also one postposed operator in ⟨y, b, z⟩ but not in ⟨x, a⟩, and hence operand b is not rightmost relative to the total number of operators in its set. We therefore consider ⟨x, a⟩ to be most harmonic with ⟨y, z, b⟩ (since all operators, regardless of their different number, precede their respective operands), less harmonic with ⟨y, b, z⟩ and least harmonic with ⟨b, y, z⟩. In so doing we are measuring the comparative position of the operand relative to however many operators there are in each member of the word order co-occurrence pair.

Where the number of operators is identical across a co-occurrence pair, the calculation of CCH is most straightforward (for any number of operators). Comparison between a one-operator set and a two-operator set (or more generally between a one-operator and an n-operator set) is also straightforward. Some interpretation of similarity must be imposed, however, when comparing sets of unequal numbers of operators greater than two, for example, a two-operator set with a three-operator set, or a three-operator set with a six-operator set. In section 4.4 we shall compare two-operator sets ⟨a, x, w⟩ with three-operator sets ⟨b, y, z, j⟩. We shall, as before, consider orders ⟨b, y, z, j⟩, ⟨y, b, z, j⟩, ⟨y, z, b, j⟩ and ⟨y, z, j, b⟩ progressively disharmonic with ⟨a, x, w⟩ on account of the increased preposing of operators on b. Similarly, for the comparison of any co-occurrence pair in which a or b are operand-peripheral (i.e. rightmost or leftmost). The only problem lies in the interpretation of co-occurrence pair ⟨x, a, w⟩ and ⟨y, b, z, j⟩ compared to ⟨x, a, w⟩ and ⟨y, z, b, j⟩. We shall consider these pairs equally harmonic, since there is no principled reason for considering either non-peripheral position of b to be more harmonic with the medial position of a. No other problems of similarity interpretation arise in the data of this paper.[10]

---

[10] For more complex cases of similarity interpretation we might adopt either of two approaches. In comparing a three-operator set with a six-operator set, for example, we could assign to each operator in the former case the value of two operators in the latter. Thus, the following word order pairs would all be maximally harmonic:

⟨a, x, w, k⟩ and ⟨b, y, z, j, l, m, n⟩
⟨x, a, w, k⟩ and ⟨y, z, b, j, l, m, n⟩
⟨x, w, a, k⟩ and ⟨y, z, j, l, b, m, n⟩
⟨x, w, k, a⟩ and ⟨y, z, j, l, m, n, b⟩

The general prediction which we are making is that whichever co-occurrence pair, W or W′, has more cross-category harmony (i.e. between A & B or between A′ & B′), the greater will be the percentage number of type W or W′ languages. We have seen that these percentages are measured by totalling the percentage number of languages with A which also have B, with the percentage of languages with B which also have A (for W); and by comparing this total with the percentage number of languages with A′ which have B′, added to the percentage of B′ languages with A′ (for W′). The higher these totals, the greater are the numbers of languages exemplifying the word order co-occurrence pair in question, relative to the numbers of languages having each of the constituent word orders.

4.3.1 *Defining CCH:* The Principle of Cross-Category Harmony (CCH)

GIVEN: Two word order co-occurrence pairs, W and W′, satisfying the following conditions:

(1) W consists of co-occurring word orders A & B, while W′ consists of A′ & B′;

(2) A, A′, B, and B′ are all ordered sets of grammatical categories (but cf. note 9);

(3) Sets A and A′ have the same categories as members: one operand, a, and at least one operator upon a. Sets B and B′ have the same categories as members: one operand, b, where b $\neq$ a, and at least one operator upon b;

(4) The relative ordering of operand to operator(s) differs either between A and A′, or between B and B′, or between both.

THEN: The relative cross-category harmony of W compared to W′ is defined according to the following formula: the more similar the position of operand a to operand b within A & B or A′ & B′ (and hence the more similar the ratio of operator preposing to postposing within A & B or A′ & B′) the greater will be the CCH of co-occurrence pair W (A & B) or W′ (A′ & B′).

PREDICTION: Whichever word order co-occurrence pair, W or W′, has more CCH, the greater will be the percentage numbers of type W or type W′ languages (as measured relative to the total numbers of languages having each of the constituent word orders, A and B (for W), and A′ and B′ (for W′)).

4.3.2 *Testing CCH:* Tables 3, 4, and 5 present the full predictions that CCH makes for the percentage co-occurrence quantities of Table 2. Table 3 gives the noun and adposition quantities, Table 4 the verb and adposition quantities, and Table 5 the noun and verb quantities. The percentage totals for co-occurring word orders are placed at the intersection of the relevant row and column. For example, in Table 3 Prep & N– has a percentage total of 156.11, Prep & –N– a total of 49.67, Prep & –N 20.76, and so on. In these tables, the direction of an arrow represents a decline in

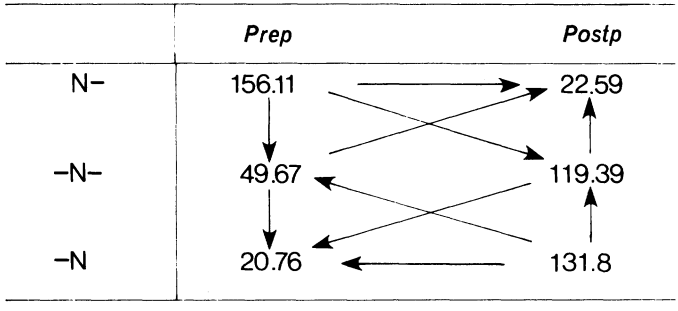
and departures from these ideal co-occurrences could be measured in the usual way. Alternatively, we might break down the six-operator set into two sets of three, for example, ⟨b, y, z, j⟩ and ⟨b, l, m, n⟩, each of which could then be compared with ⟨a, x, w, k⟩ in turn.

CCH, and hence a predicted decline in the numbers of exemplifying languages. Thus, Prep & N– has greater CCH than Prep & –N–, Prep & –N– greater CCH than Prep & –N, and the predicted decline in language quantities is fulfilled in each case.

Under each table we enumerate explicitly all the word order co-occurrence pairs for which CCH makes predictions of relative CCH. The direction of the arrow again signifies a decline in CCH. Thus, Prep & N– → Prep & –N– means that the former has more CCH than the latter. All the word order co-occurrence pairs in each table are presented in two lists. In the first list we enumerate all pairs in which the relative ordering of operator to operand differs either between A and A′, or between B and B′, but not between both, i.e. pairs such as Prep & N– and Prep & –N–, where one operator–operand ordering, Prep, is common to both. In the second list we enumerate all co-occurrence pairs in which neither A and A′, nor B and B′, have identical operator–operand orderings. For example, Prep & N– (A & B) has greater CCH than Postp & –N– (A′ & B′). Letting a stand for adposition, and b for noun, Prep & N– languages combine word order set A ⟨a, x⟩ with set B ⟨b, y, z⟩, while Postp & –N– languages combine word order set A′ ⟨x, a⟩ with set B′ ⟨y, b, z⟩. Now, A ⟨a, x⟩ (Prep) and B ⟨b, y, z⟩ (N–) are perfectly harmonic, since operands a and b are leftmost in both sets. But the medial position of b in B′ ⟨y, b, z⟩ (–N–) is one position at odds with the rightmost position of a in A′ ⟨x, a⟩ (Postp). Similarly, Prep & –N– are more harmonic than Postp & N–. The medial position of the noun ⟨y, b, z⟩ (–N–) is only one position removed from the leftmost position of the adposition ⟨a, x⟩ (Prep), whereas the initial position of the noun ⟨b, y, z⟩ (N–) is two positions removed from the rightmost position of the adposition ⟨x, a⟩ (Postp). But a co-occurrence such as Postp & –N (⟨x, a⟩ and ⟨y, z, b⟩) is not more harmonic than Prep & N– (⟨a, x⟩ and ⟨b, z, y⟩), since the harmonic rightmost position of the operands in the former case is matched by the equally harmonic leftmost position of the operands in the latter. The measurement of CCH thus generalizes readily to cases where all four sets, A, B, A′, and B′, differ, and none have identical operator–operand orderings.

Notice that the relation 'greater CCH than' is transitive. If co-occurrence pair W has more CCH than W′, and W′ has more CCH than W″, then W has more CCH than W″. For this reason the number of word order co-occurrence pairs for which CCH makes predictions, listed at the foot of each table, may be greater than the number of arrows drawn in the tables themselves. For example, in Table 5, N– & VSO has greater CCH than –N & SVO. and –N & SVO has greater CCH than –N & VSO. Hence, N– & VSO has greater CCH than –N & VSO.

CCH makes relative harmony predictions for 12 sets of word order co-occurrence pairs in each of Tables 3 and 4, and for 22 pairs in Table 5, making in all 46 predictions of relative CCH. All 46 predictions are confirmed by the quantitative data.



W → W′ signifies co-occurrence pair W has greater CCH than co-occurrence pair W′. Hence, co-occurrence pair W is predicted to have a greater percentage number of exemplifying languages than W′.

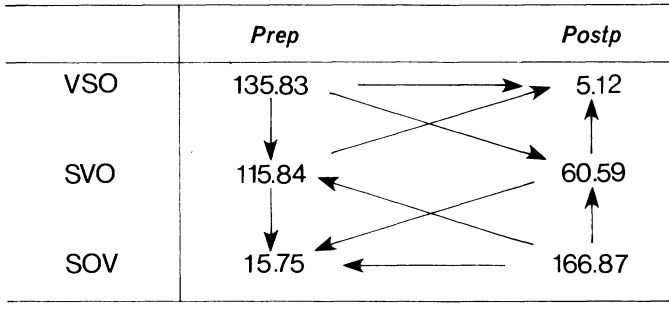
Total relative harmony preferences defined by CCH

Either A and A′, or B and B′ differ
Prep & N– → Prep & –N–
Prep & –N– → Prep & –N
Prep & N– → Prep & –N
Prep & N– → Postp & N–
Postp & –N → Postp & –N–
Postp & –N– → Postp & N–
Postp & –N → Postp & N–
Postp & –N → Prep & –N

Both differ
Prep & N– → Postp & –N–
Prep & –N– → Postp & N–
Postp & –N → Prep & –N–
Postp & –N– → Prep & –N

Results
CCH defines relative harmony preferences between 12 sets of word order co-occurrence pairs. All 12 predictions for the greater percentage numbers of the relevant languages are correctly matched by the AII data.

Table 3
Noun and adposition predictions



W → W′ signifies co-occurrence pair W has greater CCH than co-occurrence pair W′. Hence, co-occurrence pair W is predicted to have a greater percentage number of exemplifying languages than W′.

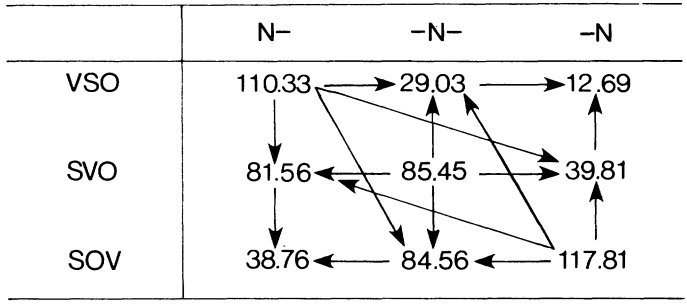
Total relative harmony preferences defined by CCH

Either A and A′, or B and B′ differ	*Both differ*
Prep & VSO → Prep & SVO	Prep & VSO → Postp & SVO
Prep & SVO → Prep & SOV	Prep & SVO → Postp & SOV
Prep & VSO → Prep & SOV	Postp & SOV → Prep & SVO
Prep & VSO → Postp & VSO	Postp & SVO → Prep & SOV
Postp & SOV → Postp & SVO	
Postp & SVO → Postp & VSO	
Postp & SOV → Postp & VSO	
Postp & SOV → Prep & SOV	

Results

CCH defines relative harmony preferences between 12 sets of word order co-occurrence pairs. All 12 predictions for the greater percentage numbers of the relevant languages are correctly matched by the AII data.

Table 4
Verb and adposition predictions



W → W′ signifies that co-occurrence pair W has greater CCH than co-occurrence pair W′. Hence, co-occurrence pair W is predicted to have a greater percentage number of exemplifying languages than W′.

Total relative harmony preferences defined by CCH

Either A and A′, or B and B′ differ

N– & VSO → N– & SVO
N– & SVO → N– & SOV
N– & VSO → N– & SOV
N– & VSO → –N– & VSO
–N– & VSO → –N & VSO
N– & VSO → –N & VSO
–N & SOV → –N & SVO
–N & SVO → –N & VSO
–N & SOV → –N & VSO
–N & SOV → –N– & SOV
–N– & SOV → N– & SOV
–N & SOV → N– & SOV
–N– & SVO → –N– & VSO
–N– & SVO → N– & SVO
–N– & SVO → –N– & SOV
–N– & SVO → –N & SVO

Both differ

N– & VSO → –N & SVO
N– & VSO → –N– & SOV
–N & SOV → –N– & VSO
–N & SOV → N– & SVO
–N– & SVO → N– & SOV
–N– & SVO → –N & VSO

Results

CCH defines relative harmony preferences between 22 sets of word order co-occurrence pairs. All 22 predictions for the greater percentage numbers of the relevant languages are correctly matched by the AII data.

Table 5
Noun and verb predictions

4.4 *Further evidence for CCH: noun and verb position*

In Section 4.1 we argued for the existence of a continuum in verb position: from verb–initial order, through SVO, through varying degrees of non-rigid SOV to rigid verb–final position. According to our principle of CCH, the position which the verb occupies along this continuum should be reflected in the position which the noun occupies in relation to its operators within the NP. The data of Greenberg's 30-language sample, which distinguishes between VSO, SVO, non-rigid SOV, and rigid SOV, enables us to test CCH in greater detail with regard to verb and noun order. We predict, for example, that rigid SOV languages will prefer to co-occur with AN & GN, rather than with NA & GN or NA & NG, since they are most harmonic with the rightmost position of the noun in the NP. But non-rigid SOV languages will be more harmonic with noun modifier co-occurrences in which at least some operators on the noun follow the noun, just as some operators on the verb follow the verb. There are 11 SOV languages in Greenberg's 30-language sample. Six are rigid and 5 non-rigid. It transpires that 5 of the 6 rigid SOV languages do indeed have AN & GN (Burushaski, Hindi, Kannada, Japanese, Turkish), while the sixth, Burmese, has a unique postposed operator: the adjective. By contrast, 4 of the 5 non-rigid SOV languages have NA & GN (Basque, Chibcha, Loritja, Nubian) while the fifth, Quechua, has AN & GN. Clearly, rigid verb finality matches noun finality, while non-rigid verb finality matches non-rigid noun finality.

With VSO and SVO we expect, correspondingly, that the more leftward position of the verb in VSO languages will be matched by a more leftward position of the noun. There are 6 VSO and 13 SVO languages in Greenberg's 30-language sample. All 6 VSO languages have NA & NG (Berber, Hebrew, Maori, Masai, Welsh, Zapotec). Of the 13 SVO languages there are only 6 with NA & NG (Fulani, Italian, Malay, Swahili, Thai, Yoruba), 5 with –N– (Greek, Guarani, Maya, Serbian, Songhai), and 2 with AN & GN (Finnish and Norwegian). Thus, whereas 100 per cent of the VSO languages have NA & NG, only 6 out of 13 SVO languages (46.15 per cent) have this co-occurrence, the remaining 7 having at least one noun operator preposed. The preposing of the subject operator in SVO languages is clearly reflected in significantly less noun–initial ordering within the NP than in VSO languages.

We can now test the predictions of CCH for the data of the 30-language sample more precisely. The relevant facts are summarized in Table 6. Table 7 sets out the precise percentage co-occurrence quantities for noun and verb position, in the manner of Table 2 above. The predictions of CCH are then tabulated in Table 8, in the manner of Tables 3, 4 and 5. A total of 38 predictions are made for the relative numbers of languages exhibiting specific co-occurrence pairs. Of these 38, 36 hold, and just two (both asterisked) are marginal counterexamples: –N & SOVnr → –N & SVO, and –N– & SVO → N– & SVO.

Notice that we are regarding –N– & SVO and –N– & SOVnr as equally

6 SOVr lgs	*5 SOVnr lgs*	*13 SVO lgs*	*6 VSO lgs*	
SOVr & –N 5	SOVnr & –N 1	SVO & –N 2	VSO & –N 0	8 –N lgs
SOVr & –N– 1	SOVnr & –N– 4	SVO & –N– 5	VSO & –N– 0	10 –N– lgs
SOVr & N– 0	SOVnr & N– 0	SVO & N– 6	VSO & N– 6	12 N– lgs

SOVr = rigid SOV
SOVnr = non-rigid SOV

Table 6
Greenberg's 30-language sample

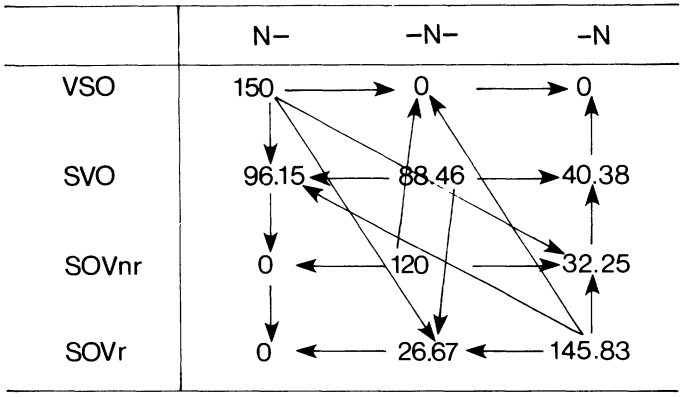
harmonic, for reasons discussed in section 4.3.0 (where –N– & SVO corresponds to ⟨x, a, w⟩ and ⟨y, b, z, j⟩ and –N– & SOVnr to ⟨x, a, w⟩ and ⟨y, z, b, j⟩).

4.5 *The predictability of the relative sizes of AII co-occurrence types*

We have been dealing so far exclusively with pairs of co-occurring operator–operand constructions, making predictions for the relative numbers of languages exhibiting one word order pair rather than another. But every language consists of many such pairs. For example, the co-occurrence type SOV & Postp & –N can be analysed into three overlapping word order pairs: SOV & Postp, Postp & –N, and SOV & –N. We might expect, therefore, that the relative frequency with which languages will select these three pairs in combination will reflect the overall CCH of all three together. In this instance all three word order pairs are maximally harmonic. But consider the co-occurrence type SOV & Prep & N–, which is analysable into SOV & Prep, Prep & N–, and SOV & N–. Prep & N– are maximally

				Total
SOVr & –N:	SOVr/–N	= 5/8 = 62.5%	–N/SOVr = 5/6 = 83.33%	145.83
SOVr & N–:	SOVr/–N–	= 1/10 = 10%	–N–/SOVr = 1/6 = 16.67%	26.67
SOVr & N–:	SOVr/N–	= 0/12 = 0%	N–/SOVr = 0/6 = 0%	0
SOVnr & –N:	SOVnr/–N	= 1/8 = 12.25%	–N/SOVnr = 1/5 = 20%	32.25
SOVnr & –N–:	SOVnr/–N–	= 4/10 = 40%	–N–/SOVnr = 4/5 = 80%	120
SOVnr & N–:	SOVnr/N–	= 0/12 = 0%	N–/SOVnr = 0/5 = 0%	0
SVO & –N:	SVO/–N	= 2/8 = 25%	–N/SVO = 2/13 = 15.38%	40.38
SVO & –N–:	SVO/–N–	= 5/10 = 50%	–N–/SVO = 5/13 = 38.46%	88.46
SVO & N–:	SVO/N–	= 6/12 = 50%	N–/SVO = 6/13 = 46.15%	96.15
VSO & –N:	VSO/–N	= 0/8 = 0%	–N/VSO = 0/6 = 0%	0
VSO & –N–:	VSO/–N–	= 0/10 = 0%	–N–/VSO = 0/6 = 0%	0
VSO & N–:	VSO/N–	= 6/12 = 50%	N–/VSO = 6/6 = 100%	150

Table 7
Percentage co-occurrence quantities for word order pairs in 30-language sample



W → W′ signifies that co-occurrence pair W has greater CCH than co-occurrence pair W′. Hence, co-occurrence pair W is predicted to have a greater percentage number of exemplifying languages than W′. (In fact, because of the total absence of some co-occurrence types in this small sample, we will allow in this instance that greater CCH be matched by a greater *or equal* percentage number of exemplifying languages.)

Total relative harmony preferences defined by CCH

Either A and A′, or B and B′ differ				*Both differ*	
N– & VSO	→ N– & SVO	–N & SOVr	→ –N & VSO	N– & VSO	→ –N & SOVnr
N– & SVO	→ N– & SOVnr	–N & SOVnr	→ –N & VSO	N– & VSO	→ –N & SVO
N– & SOVnr	→ N– & SOVr	–N & SOVr	→ –N– & SOVr	N– & VSO	→ –N– & SOVr
N– & VSO	→ N– & SOVnr	–N– & SOVr	→ N– & SOVr	–N & SOVr	→ –N– & VSO
N– & VSO	→ N– & SOVr	–N & SOVr	→ N– & SOVr	–N & SOVr	→ N– & SVO
N– & SVO	→ N– & SOVr	–N– & SVO	→ –N– & VSO	–N & SOVr	→ N– & SOVnr
N– & VSO	→ –N– & VSO	*–N– & SVO	→ N– & SVO	–N– & SVO	→ N– & SOVnr
–N– & VSO	→ –N & VSO	–N– & SVO	→ –N– & SOVr	–N– & SVO	→ N– & SOVr
N– & VSO	→ –N & VSO	–N– & SVO	→ –N & SVO	–N– & SVO	→ –N & VSO
–N & SOVr	→ –N & SOVnr	–N– & SOVnr	→ –N– & VSO	–N– & SOVnr	→ N– & SOVr
*–N & SOVnr	→ –N & SVO	–N– & SOVnr	→ N– & SOVnr	–N– & SOVnr	→ –N & SVO
–N & SVO	→ –N & VSO	–N– & SOVnr	→ –N– & SOVr	–N– & SOVnr	→ –N & VSO
–N & SOVr	→ –N & SVO	–N– & SOVnr	→ –N & SOVnr		

Results

CCH defines relative harmony preferences between 38 sets of word order co-occurrence pairs. Of the 38 predictions for the greater percentage numbers of the relevant languages, there are just 2 exceptions (both asterisked) in the 30-language sample data.

Table 8
Noun and verb predictions for the 30-language sample

harmonic, but SOV & Prep and SOV & N– are maximally disharmonic. Similarly, the co-occurrence SVO & Prep & –N is analysable into the pairs SVO & Prep, Prep & –N and SVO & –N. In SVO & Prep the verb is one operator removed from the ideal leftmost position (VSO), in Prep & –N the noun is two operator positions removed from the ideal N–, and in SVO & –N there is one operator deviation from the ideal verb–medial/noun–medial or verb–final/noun–final match-up.

We would, therefore, expect that the actual number of SOV & Postp & –N languages will exceed that of both SOV & Prep & N– languages and SVO & Prep & –N languages. More generally, we expect that the actual number of languages in each co-occurrence cell in AII should be predictable from the overall CCH of all the word order pairs which it comprises.

In order to quantify the overall CCH for each co-occurrence type we first divide each type into its three constituent word order pairs. For each pair we calculate the number of deviations (if any) made by the verb or the noun from the most harmonic ordering, and then add up the total number of deviations for all three pairs. Thus, in the case of SVO & Prep & –N, SVO & Pr deviate by one operator position, Pr & –N by two, and SVO & –N by one, making four deviations in all.[11] By contrast, SOV & Po & –N exhibits no deviations, and a type such as SOV & Po & –N– would exhibit two: Po & –N– is one operator position at variance with the ideal Po & –N, and SOV & –N– is also one position at variance with the ideal SOV & –N or SVO & –N–.

The results of this method are set out in Table 9. Because of the limited number of operators in AII (two for the verb, two for the noun, and one for the adposition), we are, unfortunately, unable to make very fine predictions. The only permissible numbers of deviations are 0, 2, and 4. Nonetheless, all the types with 4 deviations (except for type 21) are correctly predicted to have fewer exemplifying languages than all the types with 2 or 0 deviations. And the 2-deviation types include the medium-sized groups (types 16, 10, and 15), while their highest attestation (type 24 with 24 languages) has fewer languages than type 23 with 0 deviations. We would expect, however, that type 1, also with 0 deviations, would have between 28 and 24

[11] The number of deviations within each co-occurrence pair has to be quantified on the basis of noun and verb position, rather than adposition order, since both noun and verb have more than one operator in our sample. The adposition can occur in only two positions relative to its single (NP) operator, before or after. When comparing VSO & Pr, SVO & Pr, and SOV & Pr we can say that the verb is 0, 1, and 2 operator positions respectively removed from the ideal VSO & Pr match-up. Similarly, the noun is 0, 1, and 2 operator positions removed from the ideal in the progression N– & Pr, –N– & Pr and –N & Pr. Seen from the side of the adposition, however, Pr & VSO would be just one adposition operator position removed from Po & VSO, and Po & SOV just one position removed from Pr & SOV. But how would we then measure Pr & SVO, Po & SVO, and also Pr & –N– and Po & –N–? Though Pr & VSO and Pr & N– are most harmonic, and Po & VSO and Po & –N least harmonic, there is no way of assessing the intermediate cross-category harmony of the adposition with these verb–medial and noun–medial orders, on account of the unique operator on the adposition. Only on the basis of the verb and noun operators can we quantify the intermediate status of Pr & SVO (between Pr & VSO and Pr & SOV), and of Pr & –N– (between Pr & N– and Pr & –N). Hence, only verb and noun position permit a meaningful quantification of the cross-category harmony of whole co-occurrence types within this sample.

	Language Type	*No. of Deviations*	*Co-occurrence Total*	*Number of lgs. in AII*
23.	SOV & Po & –N	0	416.48	28
1.	VSO & Pr & N–	0	402.27	19
24.	SOV & Po & –N–	2	370.82	24
9.	SVO & Pr & N–	2	353.51	21
16.	SVO & Po & –N–	2	265.43	11
10.	SVO & Pr & –N–	2	250.96	8
15.	SVO & Po & –N	2	232.2	6
2.	VSO & Pr & –N–	2	214.53	5
21.	SOV & Po & N–	4	228.22	7
17.	SOV & Pr & N–	4	210.62	4
11.	SVO & Pr & –N	4	176.09	3
3.	VSO & Pr & –N	4	168.96	1
13.	SVO & Po & N–	4	164.74	0
19.	SOV & Pr & –N	4	154	1
7.	VSO & Po & –N	4	149.61	1
5.	VSO & Po & N–	4	138.04	0

Table 9
Predictions for relative quantities of AII co-occurrence types[12]

attestations rather than 19. But the approximate ranking at least of these language types is clearly predicted by this method. In a sample which contained many more operators, finer predictions would be possible since the number of deviations could be much greater.

There is another way, however, in the present context for making finer predictions within these three deviation classes. We can use the percentage co-occurrence figures for the individual word order pairs, the relative differences between which were shown in Tables 3, 4, and 5 to be predictable by CCH. By adding up the percentage totals for the three pairs in each co-occurrence type (maximum 600: 3×200), we would predict that: the higher the overall percentage total, the

[12] Language types 12 (SVO & Pr & GN & NA, with one exemplification) and 14 (SVO & Po & NG & AN, with two exemplifications) have been omitted from this table. These languages were exceptions to our otherwise very general implications (III) Prep $\supset$ (NA $\supset$ NG) and (IV) Postp $\supset$ (AN $\supset$ GN) (though cf. note 5). The existence of a limited number of exceptions makes (III) and (IV) statistical rather than non-statistical universals. Since our implicational universals feed the distributional universal (CCH) with possible language types, it follows from the definition of 'statistical implication' that such statements will permit a small number of exceptional languages to exist, regardless of any (perhaps contrary) predictions that the distributional universal might make. On this occasion types 12 (SVO & Pr & –N–) and 14 (SVO & Po & –N–) have an identical co-occurrence array for distributional purposes to types 10 and 16, respectively (with eight and eleven attestations, respectively). The fact that type 12 has only one attestation and type 14 only two can, therefore, be accounted for as a result of their exceptional status *vis à vis* universals (III) and (IV). Such exceptions will be independently predicted to have a very low frequency of occurrence, and so they have been excluded from Table 9.

greater will be the actual numbers of languages in each type. For since the percentage figures reflect the relative frequency with which individual word order pairs co-occur, they should be able to predict (when added together) the relative frequency with which these pairs will occur in combination to constitute a whole language type. We have in any case suggested (note 8) that these percentage figures not only obey CCH, but also reflect the presence of other preposed or postposed operators in the relevant languages, whose co-occurrence can be inferred from the chosen sample word order properties in conjunction with implicational laws. As a result these figures enable us to anticipate a more detailed study which tests more operator–operand deviations than we can here.

Table 9 gives the percentage totals in descending order within each deviation type. The actual numbers of languages do follow extremely closely the descending percentage figures. Hence, the percentage totals do enable us to predict the frequency with which word order pairs will combine with one another to form a whole co-occurrence type. As a result they make predictions for the relative numbers of languages in AII having each co-occurrence type.[13]

---

[13] It might be objected that there is a certain circularity in the percentage co-occurrence predictions of Table 9, since the numbers of languages whose relative size we are attempting to predict (e.g. 21 languages of type 9, 11 languages of type 16, and so on) are themselves some of the languages which figure in the word order quantities on the basis of which the predictions are being made. It might, therefore, appear as if type 9 was being predicted to have 21 exemplifying languages largely because 21 such languages figured in the word order quantities to begin with. But this impression is misleading. The paired word order quantities of Table 2 are founded on data drawn from ALL the languages in the sample. They measure the total number of languages having each word order pair (SVO & Pr, etc.), relative to the total number of languages with each individual word order (SVO, Pr, etc.). Thus, the prediction that type 9 languages exceed type 16 is made on the basis of all the languages in the sample having either and both of the word orders concerned, and not just on the basis of the actual type 9 and type 16 languages themselves. This prediction, therefore, reflects the frequency with which Prep & SVO combine throughout the whole sample, relative to the other combinatorial possibilities for these same operators and operands (Prep & VSO, Postp & SVO, etc.).

Confirmation of this interpretation comes from the following test. We have recalculated the predicted relative size of each co-occurrence type by systematically removing the actual number of languages to be predicted from the percentage quantities for the three constituent word order pairs in each type. Thus, for type 16 SVO & Postp & –N– (11 languages) we have removed 11 languages from the calculation of each of the SVO & Po, Po & –N–, and SVO & –N– pairs. Compare the following percentage co-occurrence quantities with the corresponding pairs in Table 2:

			Total
Postp & SVO:	SVO/Po = 7/68 = 10.29%	Po/SVO = 7/41 = 17.07%	27.36
Postp & –N–:	–N–/Po = 26/68 = 38.24%	Po/–N– = 26/40 = 65%	103.24
SVO & –N–:	SVO/–N– = 11/40 = 27.5%	–N–/SVO = 11/41 = 26.83%	54.33

The new co-occurrence total for SVO & Po & –N–, therefore, becomes 27.36+103.24+54.33 = 184.93. What this figure now represents is a prediction for the relative number of type 16 languages, derived from all the languages in the sample apart from those of type 16 itself.

By pursuing this method for all the co-occurrence types of Table 9 we obtained the following ranking of the co-occurrence totals:

Type 1: 234.29 (19 lgs), type 23: 230.06 (28 lgs), type 24: 222.93 (24 lgs), type 9: 198.71 (21 lgs), type 10: 191.51 (8 lgs), type 16: 184.93 (11 lgs), type 15: 182.89 (6 lgs), type 17: 182.83 (4 lgs), type 21:

Notice that language type 5, which we predicted in Section 3.1.5 to be possible on the basis of our implicational generalizations, even though unattested, can now be predicted on distributional grounds to be possibly absent in a sample of this size. Given that type 5 has a lower percentage total than the two co-occurrence types immediately above it, each with only one entry, it can be predicted to have either one or no attestations. Similarly, although we expect type 13 to have exactly one attestation, its total absence is only a marginal counterexample, and indeed its very low overall percentage total provides a reason for why this co-occurrence should be possibly absent, whereas others, with higher totals, are not.

4.6 *Comparison of CCH with Vennemann's NSP*

It is important to appreciate how our principle of CCH differs in significant ways from Vennemann's NSP, even though both principles are founded upon a similar division of categories into operators and operands:

(1) In contrast to NSP, CCH does not claim that languages tend to serialize all their operators on a consistent side of their respective operands (all on the left or all on the right). Far too many (in fact most) languages regularly have some operators on the noun to the left and some to the right, some operators on the verb to the left and some to the right, and so on. Instead, CCH asserts the importance of a balance in the position of the operand relative to its operators across the different operand categories. Whatever position the operand of one category occupies should preferably be matched by the position of the operand in each of the other categories. Now, for languages with an operand-peripheral verb, for example, a uniform leftward or rightward serialization of operators within the other categories will, of course, be most balanced, and in these cases Vennemann's predictions overlap with those of CCH. But for all categories which are not operand-peripheral (e.g. non-rigid SOV, –N–, etc.) the most preferred accompanying word orders will NOT involve uniform leftward or rightward serialization but instead a cross-categorial balance.

(2) CCH, but not NSP, makes a precise quantitative prediction to the effect that the more a language type departs from the 'ideal' harmonic ordering, the fewer exemplifying languages there will be.

(3) NSP consists of a set of (bilateral) implicational statements defining permitted versus non-permitted word order co-occurrences. CCH does not consist of a set of such implicational statements. It defines quantitative prefer-

---

179.56 (7 lgs), type 2: 164.81 (5 lgs), type 13: 164.74 (0 lgs), type 3: 158.05 (1 lg), type 11: 148.42 (3 lgs), type 19: 145.74 (1 lg), type 5: 138.04 (0 lgs), type 7: 137.4 (1 lg).

Although there are some differences relative to Table 9, it will be apparent that the overall ranking is fundamentally similar, and the predictions for the relative numbers of languages of each co-occurrence type in AII still very good. When the actual languages whose relative size is to be predicted are then added in to these calculations (as in Table 9), the results are almost perfect.

ences among word order co-occurrence types whose wellformedness is already predicted by independently needed implicational statements, and without which incorrect frequency predictions would be made. For example, as far as CCH is concerned, the co-occurrence SOV & –N– should be as frequent with AN & NG standing for –N– as with NA & GN. In fact, the former is totally unattested, being ruled out by implication (I) SOV ⊃ (AN ⊃ GN). CCH is therefore a distributional universal which is supplementary to, and which operates upon the output of, a set of implicational universals, but which does not of itself define possible language types.

(4) The word order harmony principle which CCH defines applies equally to prepositional and to postpositional languages, i.e. to the two major language types whose existence we argued for in 3.3. Quantitative predictions reflecting the degree of cross-categorial balance therefore hold irrespective of the major type affiliation of the languages concerned. By contrast, NSP actually divides the languages of the world into two types (VX and XV), and comprises separate statements defining the properties of each.

4.7 *Comparison of CCH with Greenberg's statistical implications*

We must also consider CCH in relation to Greenberg's statistical implications. Some of the statistical universals which he proposes to cover the data of AII are:

> In languages with prepositions, the genitive almost always follows the governing noun.
>
> In languages with postpositions, the genitive almost always precedes the governing noun.
>
> With overwhelmingly more than chance frequency, languages with dominant order VSO have the adjective after the noun.

Exceptional *P & –Q languages exist for all these statements, in the form of *Prep & GN languages, *Postp & NG languages, and *VSO & AN languages, respectively. Now although these co-occurrences are all attested, they are extremely disharmonic according to CCH. *Prep & GN languages co-occur only with AN and not with NA (recall implication (III) Prep ⊃ (NA ⊃ NG), which is logically equivalent to Prep ⊃ (GN ⊃ AN)). Prep & GN & AN is symbolized as Prep & –N in our notation. Similarly, *Postp & NG languages co-occur only with NA and not with AN (recall implication (IV) Postp ⊃ (AN ⊃ GN), which is logically equivalent to Postp ⊃ (NG ⊃ NA)). Postp & NG & NA is symbolized as Postp & N–. But both Prep & –N and Postp & N– are maximally disharmonic. The co-occurrence *VSO & AN is compatible with both VSO & AN & GN (i.e. VSO & –N) or with VSO & AN & NG (i.e. VSO & –N–) by implication (II) VSO ⊃ (NA ⊃ NG).

Notice, therefore, how Greenberg's statistical implications differ in an important respect from his non-statistical implications. The latter are defined in terms of two

co-occurring word orders, P & Q, and they rule out all co-occurrences of *P & −Q. But his statistical statements are not defining attested versus non-attested co-occurrences. They are, in effect, distinguishing between harmonic (and frequent) P & Q co-occurrences (e.g. Prep & N–), and disharmonic (and infrequent) *P & −Q co-occurrences (e.g. Prep & –N). As a result they constitute an alternative method for defining language distributions.

In Section 3.2 we observed that Greenberg's overuse of statistical statements was unfortunate since it resulted in a failure to define precisely the attested versus completely unattested co-occurrences in his data. Notice now that statistical implications cannot adequately capture distributional regularities either. For if we were to attempt to capture all distributional facts using them, we would need to set up a separate implicational statement for each (frequent) P & Q versus (infrequent) *P & −Q pair, i.e. we would need a separate implicational statement for each of the output predictions of CCH, thereby missing the overall regularity.

Worse yet, statistical implications could only be set up for values of P & Q which had a very high frequency of occurrence, where *P & −Q had a very low frequency of occurrence. This is because the number of *P & −Q exceptions which any implicational statement can tolerate is limited. One cannot claim that the presence of one linguistic property, P, guarantees the co-occurrence of another (which is what an 'if P then Q' statement asserts in this context) if there are many languages with P which do not in fact have Q. Hence the only distributional facts which statistical statements could potentially define are those involving very high versus very low frequencies. But the data of AII reveal many fine distinctions between large, medium- and small-sized language groups, distinctions which CCH can show to be principled.

By reserving implicational statements for the task of distinguishing attested from non-attested co-occurrences, we can, therefore, formulate just one supplementary distributional regularity which avoids many unnecessary statistical implications, and which captures generalizations which these latter are intrinsically unable to state.

4.8 *Explaining the distributional universal*

We now see that, underlying the varying numbers of languages in AII, there is a quantifiable preference for semantically similar elements, operators and operands, to be given a similar surface word order in relation to one another. In offering an explanation for CCH I would propose the following three interrelated factors.

CCH points firstly to the validity of the semantic–syntactic parallelism between the verb and its modifiers, the noun and its modifiers, and the adposition and its modifiers. Within the three categories, S, NP, and Adposition Phrase (AdpP), therefore, the constituents V, N, and Adp can be regarded as semantic head or operand constituents, with the other constituents functioning semantically as modifiers or operators, essentially as in Vennemann's semantic analysis, except that

the subject should be regarded as a verb-modifier just like all modifiers of the verb. I see evidence for the reality of this semantic–syntactic schema in the very fact that our distributional generalization refers to an organizational balance across all three categories.

Secondly, I see in CCH also evidence for the rôle of analogy. The operator preposing and postposing balance within one category generalizes to another as a result both of the operator–operand generalization linking the two categories, and of a natural tendency observable throughout language for like elements to be treated in a like manner. To the extent that the preposing/postposing balance of operators to operand within one category moves away from that of the other categories, the language becomes progressively non-preferred and infrequent. And historically the language then comes under increasing pressure to reintroduce a balance, as is argued in Hawkins (1979) (cf. Section 7 and note 25). Thus, the rôle of analogy as defined by CCH differs from its putative rôle in Vennemann's theory. Analogy does not necessarily strive to create a uniform left to right or right to left serialization of operators, but instead to achieve a cross-categorial operator preposing/postposing balance.

However, I believe that there are grammatical reasons for why languages exhibit the quantitative preferences which they do, beyond the analogical preference for like semantic elements – operators and operands – to be treated in a like manner. And the third factor for which I would argue in explaining CCH is grammatical complexity. The languages with steadily decreasing frequencies of occurrence are those with increasingly complex grammars. This point is difficult to justify in the present state of syntactic theory, since detailed syntactic analyses of the appropriate languages are still in their infancy, and since so much of our general syntactic theory still derives from English. Nonetheless, in a forthcoming note in this Journal I discuss briefly one very promising direction for defining syntactic complexity in this context, using insights from the $\bar{X}$ Theory of generative grammar (cf. Jackendoff, 1977). I suggest that languages with a greater balance of operators to operands across the different categories (i.e. with more CCH) permit the formulation of more, or more general, cross-categorial rules of grammar than languages with disharmonic orderings. Disharmonic word orders (with low CCH) preclude the collapsing of individual rules into more general cross-categorial rules, and as a result decreasing CCH is matched by increasing grammatical complexity. Since CCH predicts the quantitative decline of languages in AII, it follows that the quantitatively preferred languages have relatively simpler, and the non-preferred languages relatively more complex, grammars.

Conclusions

I shall summarize the major conclusions of this paper very briefly. We argued in Section 2 against Vennemann's Natural Serialization Principle on account of the

excessive number of counterexamples within Greenberg's data. Many of these occur in SVO languages, and most are attributable to Vennemann's reformulation of Greenberg's unilateral implications as bilateral statements.

In Section 3 we presented our own reformulated implications for Greenberg's data. These used operand-peripheral orders of verb and adposition as the ultimate antecedent property, and defined the possible and impossible noun operator deviations from the serialization of the antecedent verb or adposition order. The first noun operator to deviate is the adjective rather than the genitive, and we explained this on the grounds that the adjective is syntactically and semantically less complex, and historically less stable, than the genitive. We argued more generally that implicational universals of word order will need to be multi-valued rather than bi-valued, unilateral rather than bilateral, and preferably non-statistical rather than statistical, if they are to succeed in distinguishing attested from non-attested word order co-occurrences.

In Section 4 we proposed our second type of universal, the distributional universal, CCH. This universal predicts the relative frequencies of pairs of co-occurring word orders. The more similar the position of the operand across each pair, the greater the percentage number of the relevant languages. These percentage figures were then used, together with the number of operator–operand deviations across the categories, to predict the relative numbers of whole language types in Greenberg's data. CCH differs from Vennemann's NSP in the nature of the 'ideal' operator–operand ordering which it defines. Languages strive not for a uniform leftward or rightward serialization of operators relative to their operand, but for a balance in the position of the operand across the categories. It differs also in that it is an explicitly quantitative principle which operates upon the co-occurrences permitted by implicational statements, and does not itself define permitted versus non-permitted co-occurrences. CCH is also an improvement over Greenberg's statistical implications in that it avoids a multiplicity of such statements, each defining one of the output predictions of CCH (in terms of very frequent P & Q co-occurrences versus very infrequent *P & −Q co-occurrences). It also captures distributional differences between large and medium-sized, and medium-sized and small groups, which cannot be captured at all by statistical implications. In explaining CCH we saw evidence for Vennemann's operator–operand schema, for the rôle of analogy, and also for increasing grammatical complexity matching decreasing language frequencies.

Finally, we must stress again, as we did in the introduction, that Greenberg's data is only a convenience sample of languages. And though it is a large sample, with a very wide genetic coverage, our conclusions must be interpreted relative to this sample. However, my current work, aimed at broadening Greenberg's data base, has given me confidence that the generalizations of the present paper do have more general validity. Before publishing this extended work, I have decided to present what are, I believe, the most adequate and precise implicational and distributional

universals compatible with Greenberg's original data sample, since it was this sample which led to the original insights about cross-language word order generalizations. My major point is that the universals of word order which have been extracted from this important sample, both by Greenberg himself and by subsequent researchers, are not as general or as adequate as they might have been.

REFERENCES

Givon, T. (1975). Serial verbs and syntactic change: Niger–Congo. In Li, C. N. (ed.), 49–112.

Greenberg, J. H. (1966). Some universals of grammar with particular reference to the order of meaningful elements. In Greenberg, J. H. (ed.), *Universals of language*. Cambridge, Mass: MIT Press. 73–113.

Hawkins, J. A. (1978). Theoretically significant word order patterns in Greenberg's Appendix II. In *University of Essex Department of Language and Linguistics Occasional Papers No. 20: Universal Grammar*, 87–150.

Hawkins, J. A. (1979). Implicational universals as predictors of word order change. *Lg* 55, 618–648.

Jackendoff, R. (1977). $\bar{X}$ syntax: a study of phrase structure. (Linguistic Inquiry Monograph No. 2.) Cambridge, Mass: MIT Press.

Keenan, E. L. & Comrie, B. (1972). Noun phrase accessibility and universal grammar. Paper presented at the Winter meeting of the Linguistic Society of America.

Keenan, E. L. & Comrie, B. (1977). Noun phrase accessibility and universal grammar. *LIn* 8. 63–99.

Li, C. N. (ed.) (1975). *Word order and word order change*. Austin, Texas: University of Texas Press.

Li, C. N. & Thompson, S. (1975). The semantic function of word order: a case study in Mandarin. In Li, C. N. (ed.), 163–195.

Steele, S. (1975). On some factors that affect and effect word order. In Li, C. N. (ed.), 197–268.

Vennemann, T. (1972). Analogy in generative grammar: the origin of word order. Paper read at the 11th International Congress of Linguists, Bologna.

Vennemann, T. (1974a). Topics, subjects and word order: from SXV to SVX via TVX. In Anderson, J. M. & Jones, C. (eds) *Historical linguistics*. Amsterdam: North Holland. 339–376.

Vennemann, T. (1974b). Theoretical word order studies: results and problems. *Papiere zur Linguistik* 7. 5–25.

Vennemann, T. (1975). An explanation of drift. In Li, C. N. (ed.), 269–305.