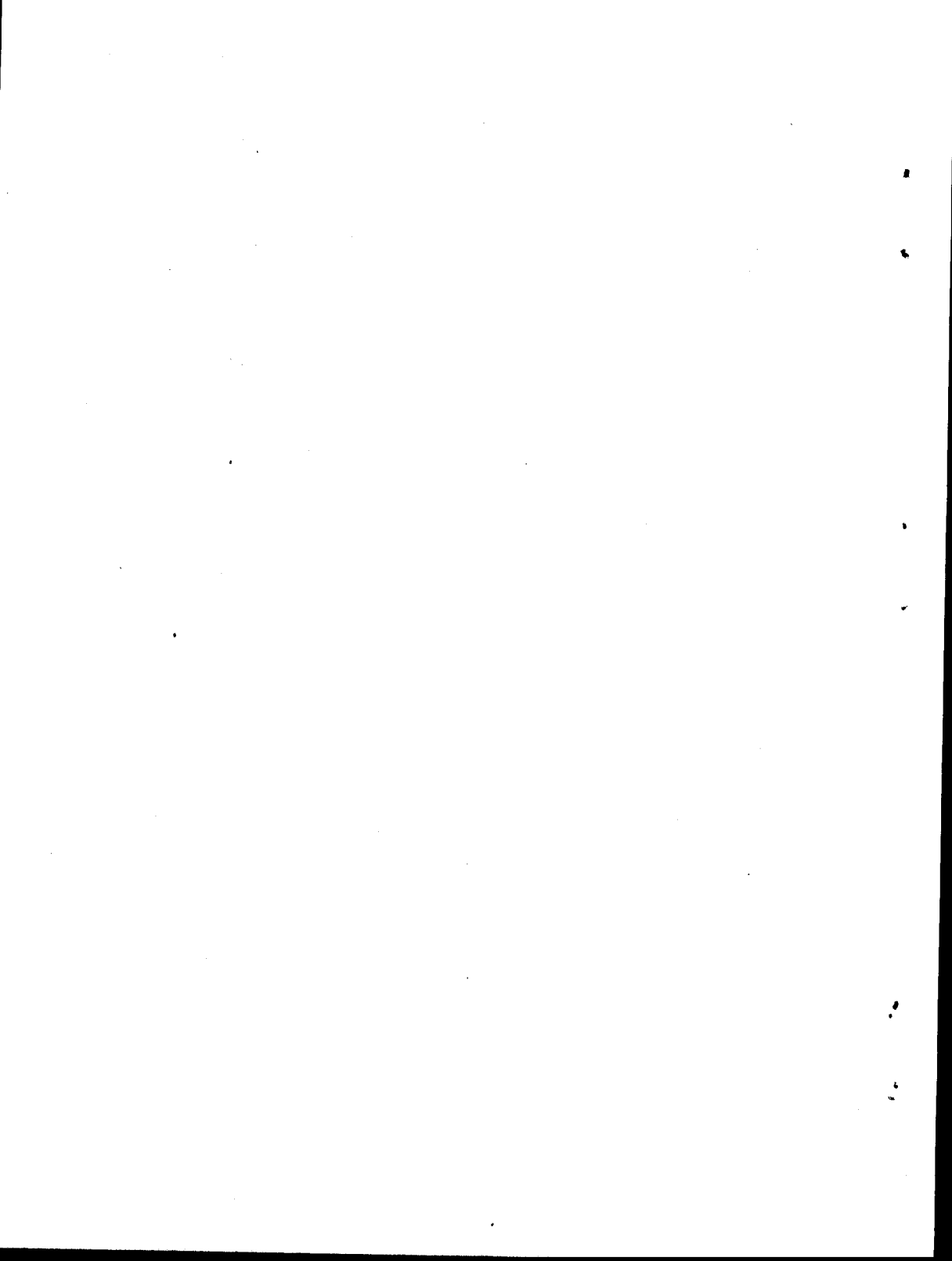


**CALGARY WORKING PAPERS
IN LINGUISTICS**

Number 15 Fall 1992

**THE
UNIVERSITY OF
CALGARY**

Department of Linguistics
The University of Calgary
2500 University Drive N.W.
Calgary, Alberta, CANADA
T2N 1N4



Foreword

The editors of this issue, Siv M. Rogmans, Margrit Prevost, and Lorna V. Rowsell, are pleased to present the fifteenth issue of the *Calgary Working Papers in Linguistics* published by the Department of Linguistics at the University of Calgary. The papers published here represent works in progress and as such should not be considered in any way final or definitive.

All the contributors to this issue are or have been associated with the Department of Linguistics at the University of Calgary. The collection of essays is quite diverse, drawing from several language families and different areas of linguistic study. Dr. John Archibald, in his article *The Structure of the Colour Lexicon*, suggests a dual-tetrahedral model of colour lexicon space which could help to explain the differences between basic and non-basic colours. Cynthia Caswell, in her article *Constraints on Reinterpretation Processes as evidenced in Particle Reinterpretation Data from Homeric Greek.*, presents an analysis of particles in Ancient Greek which shows that, contrary to Lightfoot (1980), such a syntactic change is subject to semantic constraints. Naomi Cull examines *Proto-Romance Syllable Structure* and concludes that Proto-Romance, the ancestor of the Romance languages, began with a VC\$CV syllable structure rather than the commonly accepted V\$CCV structure and that this structure was inherited from Proto-Indo-European. Michael Durrant presents a paper on a *Government and Binding Approach of Bella Coola* in which he investigates the adequacy of GB theory for the Bella Coola language and classifies Bella Coola as a VSO language. Finally, Ken Hanna and Andrea Wilhelm in their paper *On the Acquisition of Wh-Questions*, report on some interesting findings from a recent study into the acquisition of subject and object Wh-questions by children in the age range of 3;4 to 4;7 years.

The next edition of CWPL, which will appear a little earlier than usual in September 1993, will feature a collection of papers compiled from a panel discussion which took place at the annual meeting of the Alberta Conference on Language in Banff, Alberta in November 1992. Papers on all topics in Linguistics are still gladly accepted for further editions of the CWPL.

The editors wish to thank all the contributors, and especially Vi Lake for devoting so much time, skill and patience to the preparation of these papers for printing. We would also like to thank all those who helped with proof-reading and other tasks.

© 1992. Calgary Working Papers in Linguistics. Printed in Canada by University of Calgary Printing Services. ISSN 0823-0579.

CALL FOR PAPERS

Calgary Working Papers in Linguistics is an annual journal which includes papers by faculty and students in Linguistics and related disciplines, both at the University of Calgary and elsewhere.

The editors would like to encourage all readers to submit papers for future publication. The deadline for submission of papers is August, 30 in order to meet a Autumn publication date. The editors would prefer contributions on 3 1/2 inch Micro Floppy Disks (preferably Macintosh Apple and Microsoft Word format). We furthermore request that the submissions follow the API style sheet, use 10x10 point *Times* font, and leave a 2.056" top- and bottom-margin, as well as a 1.75" margin on both the left and the right hand side. In those rare circumstances in which the contributor does not have access to a computer, the editors will accept two typed copies of the paper. Those wishing to submit papers should send manuscripts to the address stated below. Postage costs should be included if the disk containing the manuscript or the paper copy are to be returned. Appearance of papers in this volume does not preclude their publication in another form elsewhere.

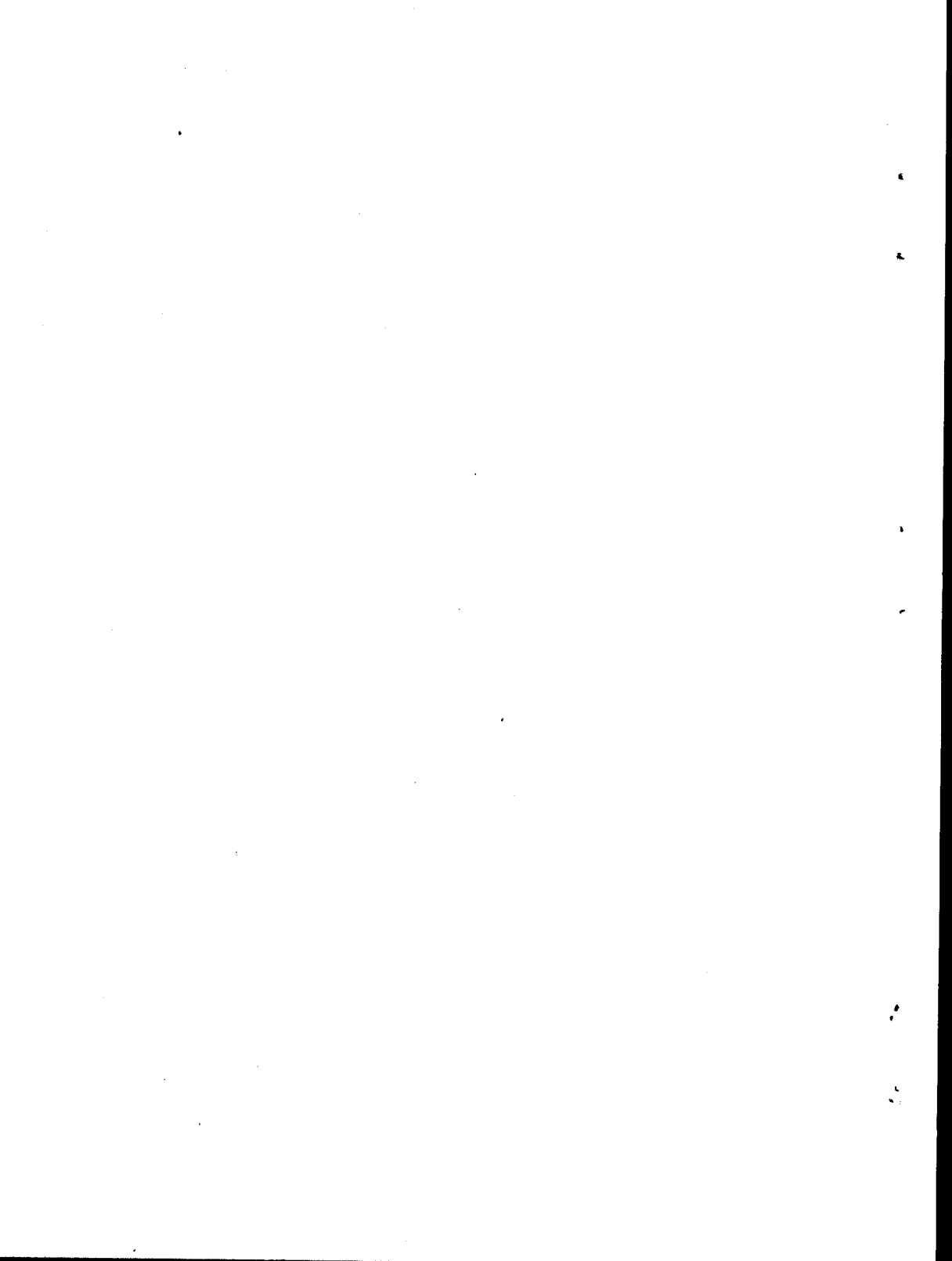
CWPL is available from:

CWPL
Department of Linguistics
The University of Calgary
2500 University Drive N.W.
Calgary, Alberta, Canada
T2N 1N4

The journal is also available on a reciprocal exchange basis. If you publish a journal or newsletter which you would like to send us, we will send you our journal gratis. Yearly subscription rate in Canada is \$ 8.- including G.S.T. For U.S. subscribers the rate is \$ 9.- (Cdn), and for overseas subscribers, \$ 10 - (Cdn).

Table of Contents

| | | |
|--|-----------------------------------|----|
| Foreword..... | i | |
| Call for Papers | ii | |
| Table of Contents | iii | |
| | | |
| <i>The Structure of the Colour Lexicon</i> | John Archibald | 1 |
| | | |
| <i>Constraints on Reinterpretation Processes as evidenced in Particle Reinterpretation Data from Homeric Greek</i> | Cynthia Caswell | 13 |
| | | |
| <i>Proto-Romance Syllable Structure</i> | Naomi Cull | 35 |
| | | |
| <i>A Government and Binding Approach to Bella Coola</i> | Michael Durrant | 75 |
| | | |
| <i>On the Acquisition of WH-Questions</i> | Andrea Wilhelm Ken Hanna | 89 |



The Structure of the Colour Lexicon

John Archibald
University of Calgary

In Archibald (1989) I proposed a model of colour space based on the colour lexicon. The research design involved having subjects sort a list of colour terms into a variety of categories. Most of the previous research on the linguistic structure of colour terms (Berlin & Kay, 1969; Mervis & Roth, 1981), had been done by having subjects perform various sorting tasks on colour chips. That is to say, most of the information came from the manipulation of visual stimuli based on the perception of colour. One of the standard models of describing colour space is shown in Figure 1. This is the model shown in Kay & McDaniel (1978).

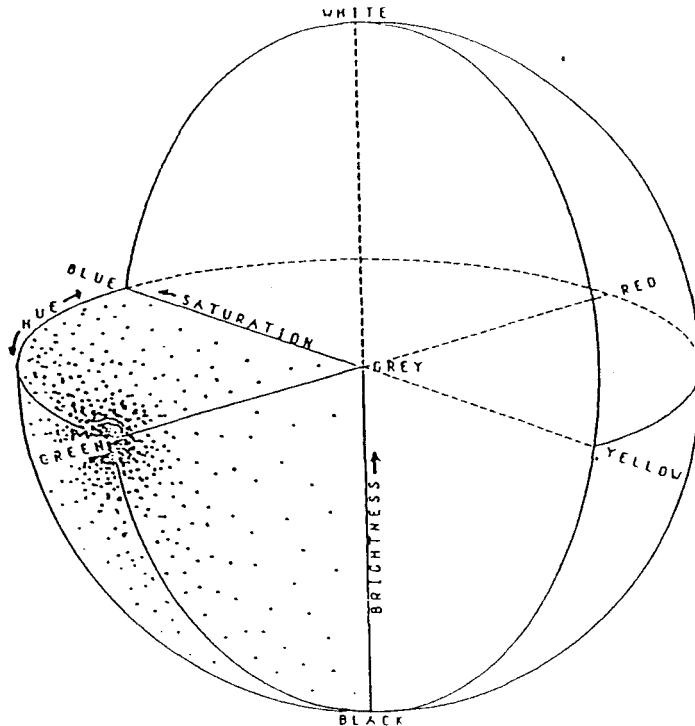


FIGURE 1: Kay & McDaniel's (1978) Model of Colour Space.

The spherical model, with its four points of equal status along the equator (*red, yellow, green, blue*), is justified physiologically (De Valois et al. 1966).

Let us briefly review Berlin & Kay's (1968) criteria for establishing whether a particular item was a *basic* term or not:

- 1) it must be monolexemic (ruling out *lime green*)
- 2) its signification must not be included in another colour term (ruling out *scarlet* which is a kind of *red*)
- 3) its usage must not be restricted to a narrow range of objects (ruling out *blond*)
- 4) it must be psychologically salient (ruling out *taupe*)

In Archibald (1989) I pointed out that the spherical model has great difficulty accounting for the distribution of basic colour terms on the colour solid. Consider the mid-points between the polar extremes and their lexical status (basic or non-basic). Between *blue* and *red* there is *purple*, a basic term. Between *red* and *yellow* there is *orange*, another basic term. But between *blue* and *green* there is *turquoise*, a non-basic term. And between *green* and *yellow*, there is *chartreuse*, another non-basic term. Just by looking at the equator of this colour model, we can see that it has no predictive power as to the lexical status of a colour term. The rest of the colour solid is likewise problematic. Similarly, the model cannot explain why terms at certain points on the model are basic while terms at other points are non-basic. Ultimately, we should like our model to be able to account for these lexical facts.

Further examination of this problem quickly reveals that it is the status of the colour terms centred on green (*chartreuse* and *turquoise*) which are problematic. Therefore, contrary to the predictions of the spherical model, I suggested that green has a different lexical status from *red, blue, and yellow*.

Essentially, I was investigating whether the two-level structure (basic versus non-basic) is enough. Do we need a third level of structure? To examine this question I divided the notion of basic colour term into two levels: primary basic and secondary basic. The primary colours are those which cannot be derived from combinations of other colours: *red, yellow, blue, black, and white*. The secondary colours are perceived as being derived from any two of the primaries². For example:

red + blue = purple
red + yellow = orange
blue + yellow = green
red + white = pink
yellow + black = brown

¹ See Appendix B for another method of classification

² This is not an uncontroversial distinction. Theatrical lighting designers, in particular, were outraged..

The task then is to determine whether green, in sorting tasks, patterns as a primary colour (red, blue, and yellow) or like a secondary colour (purple and orange). We will see that by assigning green a different status from red, blue, and yellow, our model gains considerable power.

The Testing

I will not go into details of the sorting tasks here (see Archibald, 1989 for further discussion). Suffice it to say that the following list of lexical items was given to a number of subjects who had to place the terms under given category headings. The terms were:

| | |
|----------|-----------|
| scarlet | turquoise |
| yellow | orange |
| burgundy | lavender |
| beige | lime |
| pink | navy |
| crimson | olive |
| gold | red |
| cream | blue |
| purple | peach |
| green | brown |
| mauve | emerald |

This was the random order in which the terms were presented to the subjects. The data were then analyzed by means of a statistical procedure known as dual scaling.

Dual scaling is a technique used to reveal the structure of categorical data. The dual scaling technique looks at the data matrix and determines what, if any, patterns underly the organization of these data. The major pattern that it pulls out is referred to as the first solution or first dimension. The programme will then extract other patterns (referred to as the second and third dimensions, and so on). The technique is dual in that it determines patterns in both rows and columns. That is to say, it takes into account how subject one classified all the colours, and how all the subjects responded to, for example, pink.

All in all, the subjects performed fifteen sorting tasks with the following category headings:

- 1) Four unlabeled categories
- 2) Red-Green-Yellow-Blue
- 3) Red-Purple-Yellow-Blue
- 4) Red-Orange-Yellow-Blue
- 5) Three unlabeled categories
- 6) Red-Yellow-Blue
- 7) Red-Green-Blue (no Yellow)
- 8) Red-Yellow-Green (no Blue)
- 9) Green-Yellow-Blue (no Red)
- 10) Orange-Yellow-Blue
- 11) Red-Orange-Blue

- 12) Red-Yellow-Orange
- 13) Purple-Yellow-Blue
- 14) Red-Yellow-Purple
- 15) Red-Purple-Blue

All of these tests revealed two organizing principles of the primary and secondary terms:

- 1) The primary terms remain maximally distant.
- 2) The secondary terms exhibit an elastic behaviour between their source primaries.

I concluded that although both primary and secondary colours should be considered basic, they are perceived differently. This forced me to reject the standard spherical model of colour space. I proposed the model shown below to graphically depict the linguistic structure of the colour lexicon³. This model has the explanatory power that the standard model lacked. There is now a way of classifying a colour as basic or non-basic by looking at its position on the model. The apices of the dual-tetrahedron have the potential to be encoded as basic terms, as do the midpoints of each of the sides. This is not to say that every language will encode all of these positions as basic; we know this is not the case.

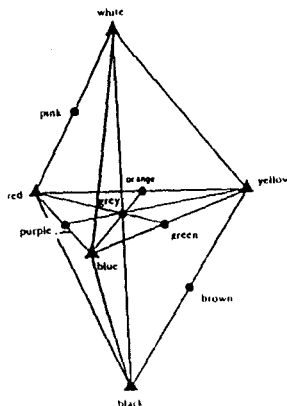


FIGURE 2: A Model of the Colour Lexicon

³ Thanks to John Giesbrecht for drawing Figures 2 and 3.

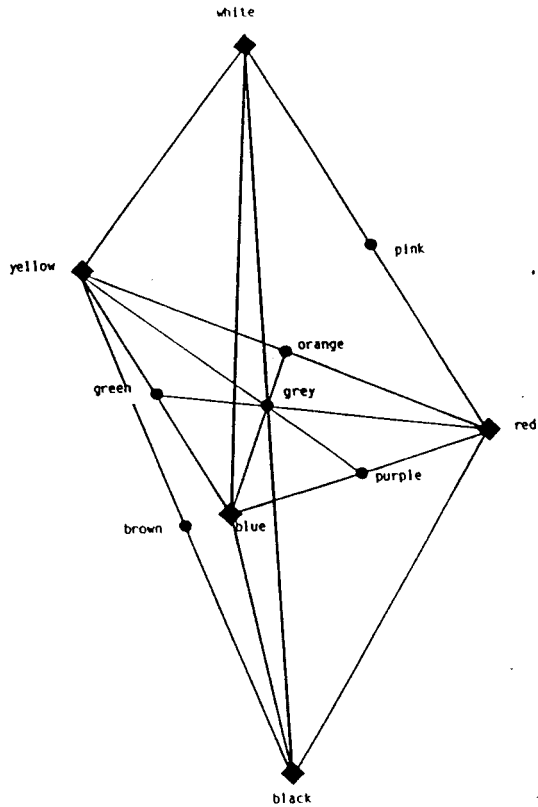


FIGURE 3: Model of the Colour Lexicon Incorporating Perceptual Distance

But while the apices and mid-points both have the potential to be encoded as basic terms, they also reflect the differing lexical status of the primary basic terms (the apices) and the secondary basics (the mid-points). Thus, this model can account for the difference in behaviour between red, yellow, and blue versus green, purple, and orange as the model assigns them different status.

However, there are still gaps on the model. Why is there not a basic term at the mid-point of *white* and *yellow*, or *white* and *blue*, or *red* and *black*? I proposed a principle of maximal perceptual distance to account for these gaps. To predict whether a colour has the potential to become basic, we must note the following four facts:

- 1) If we lighten a dark colour we produce the potential for a basic colour term (e.g. *pink*).
- 2) If we darken a light colour we produce the potential for a basic colour term (e.g. *brown*).
- 3) If we lighten a light colour we do not produce the potential for a basic colour term (e.g. *canary yellow*).
- 4) If we darken a dark colour we do not produce the potential for a basic colour term (e.g. *burgundy*).

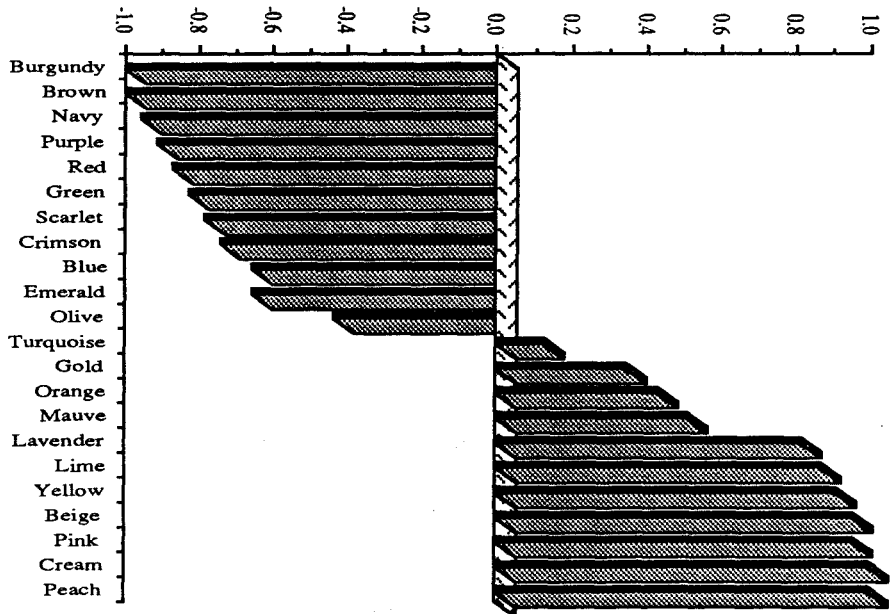
The principle of perceptual distance is graphically represented in Figure 3 (the central plane of Figure 2 has been inclined thereby altering some of the distances between colour points).

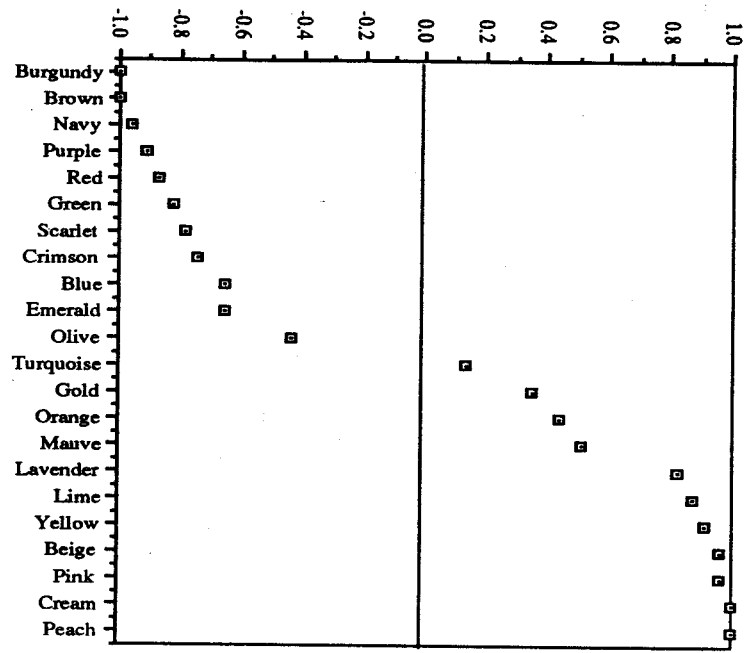
Now, this accounts for three of the four gaps but it does not account for the gap between white and blue. Here we are lightening a dark colour and we would expect a basic colour term. This, however, is an accidental gap in the lexicalization of English colour space. Languages such as Russian and Ukrainian have basic terms for light blue. The model, thus, allows for this. This fact can also be accounted for by the fact that Russian 'blue' is focussed closer to black, thus increasing the distance from blue to white. Consequently, there is greater perceptual distance and another basic term can be encoded.

I had, however, no empirical justification for assuming that the subjects actually treated light and dark colours differently. I needed some support for maintaining that this was, in fact, a psychologically relevant distinction.

In order to determine whether this was the case, I presented the same list of 22 colour terms to forty-six subjects, and asked them to sort the terms into two categories labelled Light and Dark. A score of +1 was assigned if the term was placed in the *light* column, and a score of -1 was assigned if the term was placed in the *dark* column. Mean scores were then calculated for each colour term. The closer the mean score was to zero, the more variability in placement. That is to say, if twenty-three subjects placed *gold* in the *dark* column and twenty-three subjects placed *gold* in the *light* column, the mean score would be zero. The following are the results⁴:

⁴ See Appendix A for the numerical data.





References

- Archibald, J. (1989). A lexical model of color space. In: R. Corrigan, F. Eckman, & M. Noonan, eds. *Linguistic Categorization*. Dordrecht: John Benjamins.
- Berlin, Brent and Paul Kay. 1969. *Basic Color Terms: Their Universality and Evolution*. University of California Press.
- De Valois, R.L., I. Abramov, and G.H. Jacobs. 1966. "Analysis of response patterns of LGN cells". *Journal of the Optical Society of America*, 56: 966—977.
- Gottesman, Greg. 1991. *Laundry and Ironing*. In: *College Survival*. New York: Prentice Hall.
- Kay, Paul and Chad McDaniel. 1978. "The linguistic significance of the meanings of basic colour terms". *Language*, 54. 610—646.
- Mervis, Carolyn and E.M. Roth. 1981. "The internal structure of basic and non-basic colour categories". *Language*, 57. 384—405.
- Nishisato, S. 1980. *Analysis of Categorical Data: Dual Scaling and its Applications*. University of Toronto Press.

Tide detergent (box). 1991. Cincinnati: Procter & Gamble.

APPENDIX A: MEAN SCORES

| | |
|-----------|--------|
| Burgundy | -1.000 |
| Brown | -1.000 |
| Navy | -0.957 |
| Purple | -0.913 |
| Red | -0.870 |
| Green | -0.826 |
| Scarlet | -0.783 |
| Crimson | -0.739 |
| Blue | -0.652 |
| Emerald | -0.652 |
| Olive | -0.435 |
| Turquoise | 0.130 |
| Gold | 0.348 |
| Orange | 0.435 |
| Mauve | 0.511 |
| Lavender | 0.822 |
| Lime | 0.870 |
| Yellow | 0.913 |
| Beige | 0.957 |
| Pink | 0.957 |
| Cream | 1.000 |
| Peach | 1.000 |

APPENDIX B

By: Ken Shirriff, Oakland, California

Diverse languages classify colours in generally the same way. Why do languages follow these rules? My hypothesis is that cultures find it necessary to develop words for colours in order to do their washing. That is, "language follows laundry." This is a bold claim, but the evidence is compelling: The rules of laundry directly account for the rules of colour terms.

Previous Research

Berlin and Kay (1969) examined 98 languages from several families and found that although languages have different numbers of basic words for colours, the colours described by these words fall into a universal system of eleven categories. Moreover, languages with fewer than eleven basic colour terms obey the following rules (Berlin and Kay, 1969):

1. All languages contain terms for white and black. (To be precise, these terms distinguish light shades from dark shades)
2. If a language contains three terms for colours, then it contains a term for red.
3. If a language contains four terms, then it contains a term for either green or yellow, but not both.
4. If a language contains five terms, then it contains terms for both green and yellow.
5. If a language contains six terms, then it contains a term for blue.
6. If a language contains seven terms, then it contains a term for brown.
7. If a language contains eight or more terms, then it contains a term for purple, pink, orange, or some combination.

Laundry and Colour Terms

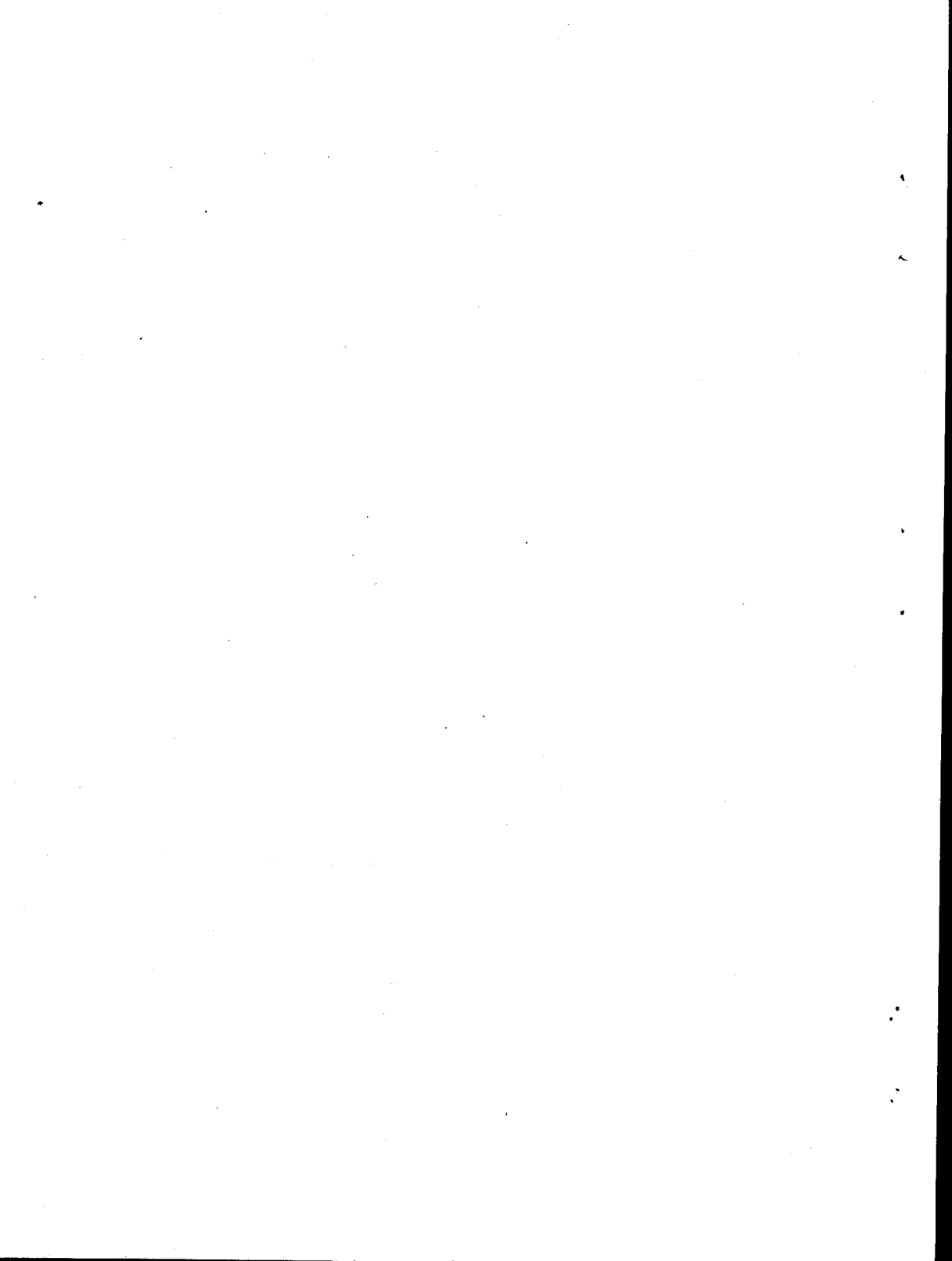
A typical detergent box (Tide, 1991) states that colours must be separated for laundry. This is elaborated by Gottesman (1991), who gives a basic rule of laundry: "Always separate darks and lights". An obvious conclusion is that in order to wash clothing, a culture must first have words to distinguish darks and lights. This is in perfect agreement with colour rule 1.

The second rule for laundry separation is: "Never wash red with anything even remotely white" (Gottesman, 1991). According to my hypothesis, cultures must next develop a word to describe red. This is in agreement with rule 2.

For more advanced laundry, bright colours such as green and yellow should be washed separately. This corresponds to rules 3 and 4.

Next, cultures discover that washing blue jeans separately is beneficial, resulting in rule 5. Finally, the remaining colours are named, resulting in rules 6 and 7.

Although I consider the merits of my hypothesis to be obvious, additional research can be done to confirm it. An examination of primary dyes might reveal why greens and yellows are separated. Anthropological studies could prove the match between a culture's word for colour and its washing habits.



Mechanisms of Language Change
CONSTRAINTS ON REINTERPRETATION PROCESSES
as evidenced in
Particle Reinterpretation Data from Homeric Greek

by Cynthia Caswell
University of Calgary

Abstract

Lightfoot (1980) claims that the mechanism for syntactic change, **reinterpretation**, lacks constraints. An in depth study of particle reinterpretation in Ancient Greek will be presented to demonstrate semantic constraints on the process. Particle reinterpretation involves several steps, whereby adverbial particles are associated with Kase-phrases (KP) and are eventually reinterpreted as the case-assigning head of the phrase. The constraints on the process are specific and easily identifiable, but they should not be attributed to the "the very broad constraints of UG", as Lightfoot claimed.

Familiarity with Lehmann and Venneman's (1974) theory of word-order change and Baker's (1988) theory of incorporation is helpful for understanding the issue, but not absolutely necessary to follow the argument as presented here.

O. Introduction

Proto-syntax is probably the most controversial area of study in historical linguistics. The discovery of Hittite radically changed assumptions about Proto Indo-European word order. Whereas PIE had been assumed to be SVO by the earliest historical syntacticians (late 1800's), the predominant SOV order in Hittite provoked a reanalysis of the Sanskrit, Greek, Latin and Germanic data. Sanskrit showed predominant OV patterns and in the other languages, relics of OV order were also found. In particular, a comparison of Latin to the daughter languages showed a reasonably clear pattern of change, from mixed VO and OV ordering in Latin to a more consistent VO ordering in the daughter languages (Lehmann; 1974: 238; Lightfoot; 1980: 36). Consequently, Lehmann (1974), Vennemann (1974) and others introduced a theory of word-order change based on Greenberg's typological universals (the continuum for pre- (OV) and post- (VO) specification tendencies). Changes from OV to VO ordering involve reinterpretation and are constrained by Behagel's Law, which states that languages tend towards *continuous constituents*. This law naturally feeds into a Preference Law for Word-order based on consistency of specification. It is claimed that once the verb-object order has changed, the other constituents will change order also, so that the direction of specification in the language is consistent (Murray; 1989; Bellusci; 1991: 24-40).

Lightfoot's major arguments against this theory of word-order change are: 1) lack of a sufficient data base, which would justify the extension of the results to other language families

(Lightfoot: 36); and 2) the lack of constraints (other than the very broad constraints of UG) on changes in the grammar brought about by reinterpretation processes. Because any given generation within a language community is cut off from the earlier history of the language, he argues that no successful reconstruction of proto-syntax can be achieved without a data-base of proto-sentences (Lightfoot: 37,38).

While Lightfoot's complaints about extending the theory to other languages and the lack of a sufficient data-base of proto-sentences seem well justified, his argument that there are no constraints on the reinterpretation process is open to debate. There are syntactic changes that can be posited for Indo-European families that do have a significant data-base, from which we may draw some conclusions about reinterpretation. One such syntactic change is the transition from *particles to prepositions*. The majority of my data will be drawn from changes that are documented from the Greek language and that are found within the literary tradition of Homer.

Interestingly enough, the Homeric tradition (like much poetry) preserves archaic features of the language, some of which are now attributed to pre-Mycenean ancestry (Palmer; 1980: 97-101; Horrocks; 1980: 148-53). While I can not argue that the Greek bards were aware of the origins of the syntactic variations maintained within the tradition, nevertheless, these variations existed in writing and were performed publicly for centuries amongst the Greek populations. Consequently, a path of syntactic change can be traced, by comparing the Homeric documents with the earlier Mycenaean tablets and the works of the later Classical poets and prose writers. Secondary data from other languages is also available for a comparison of particle reinterpretation within the Indo-European family.

So, I am attempting to approach the issue of constraints on reinterpretation from an empirical basis. Although the reinterpretation process is not considered to be a conscious activity of the language community, this should not rule out the identification of linguistic constraints on its occurrence. Before these constraints can be assessed, however, it is necessary to have an in depth description of the state of "particles" within the Homeric tradition, in order to understand its value as a data-base for assessing this type of syntactic change. The summary in Section 1 draws heavily from the work of Horrocks (1980) *Space and Time in Homer*, which is undoubtedly the most thorough and in depth discussion of the process available. His work also offers a thorough range of support data from other Indo-European languages.

1. THE FUNCTION OF PARTICLES IN HOMERIC GREEK

1.1 Ambiguity in Traditional Analyses

Within the traditional analytic framework, there has been considerable ambiguity in determining the function of particles in Homeric Greek. Three roles for the particle are identified: adverb, preposition and preverb. Horrocks criticizes Chantraine's (1953)¹ analysis of the

¹ Horrocks gives a summary and criticism of Chantraine (1953) before presenting his own criteria for assessing particle types and functions. Chantraine's view is considered to be typical of the traditional view of the account, which appears in most Greek grammars. (Horrocks 1980:10)

particles, claiming that it is contradictory for Chantraine to consider “the evolution of prepositions as a process which is simultaneously both complete and still in progress in Homeric Greek” (Horrocks; 1980: 12-3). Horrocks illustrates how the ambiguity of the traditional terminology becomes apparent when examining the particles in context. The following examples demonstrate this ambiguity, using the particle *amphi* (Horrocks: 9).² In example (1) the particle *amphi* would be labelled as an adverb, according to the framework of traditional analyses, yet it clearly shows prepositional tendencies.

- (1) O 266-7 ἀμφὶ δὲ χαῖται / ὄμοις ἄισσονται
 amphi de khaitai / omois aisontai
 about - and - long- streaming mane / his shoulders [dat] - floats
 “and his long streaming mane floats about his shoulders”

In (2a) the particle is probably a case of preverb in tmesis,³ for despite its prepositional position in that example, the particle occurs as a “univerbated”⁴ adverb in (2b). Note also that the compound verb (amphero) takes on more than one sense, during translation into English (compare 2a and 2b).

- (2a) Λ 482-3 ὣς ῥά τότε ἀμφῷ Ὀδυσῆα ... / Τρῶες ἔπον
 hōs ra tot amph odūsea (acc.) / trōes hepon
 even - so - then - about - Odysseus ... / Trojans set
 “even so then, (the) ...Trojans surrounded Odysseus”

- (2b) Σ 348 γάστρην μὲν τρίποδος πῦρ ἀμφεπε
 gastreen (acc) men tripodōs pyr amphepe
 belly - then - cauldron - fire - aboutset
 “then the fire played about the belly of the cauldron”

In example (3) *amphi* is functioning clearly in a prepositional manner, especially when compared to example (1).

- (3) Γ 334 ἀμφὶ δῆῶρ ὄμοισιν βάλετο ξίφος ...
 amphi d]ar omois (dat) baletō ksiphōs
 about - and so - (his) shoulders - threw (he)-(his) sword
 “and so he threw his sword about his shoulders”

² The various statements of analysis are drawn from Horrocks; however, any errors of application in the broader framework of reinterpretation presented here are my own.

³ Webster (1988) defines tmesis as: “separation of parts of a compound by the intervention of one or more words.” English example: *what place soever* vs. *whatsoever place*.

⁴ Univerbation, in historical theory, is a process of compounding or fusion, which may or may not be subject to constraints, depending on the theorist’s position.

Collectively, these examples demonstrate that the traditional terminology can not be applied unambiguously when analyzing the Homeric text. Horrocks argues instead for a single role "particle". He considers the function of particles to be essentially adverbial; although they demonstrate both *adverbial* and *prepositional* tendencies. Prepositional particles either explicitly or implicitly modify a Noun Phrase inflected for case, which will be termed here KASE-phrase (KP; Travis; 1986⁵). Adverbial particles stand only in relationship to a verb and have no specifying⁶ relationship to a KASE-phrase. Whether or not the particle specifies KASE is the distinguishing factor between the two functional tendencies.

1.2 A Survey of Particle Distribution

Horrocks' argument is supported by a survey of the particles as they are found in general clause and preverb environments. The (traditionally termed) independent adverbs and prepositions are actually found in parallel environments throughout the text. Both can occur in all of the clause types listed in Table 1, and they both function in optional adjuncts.

TABLE 1: BASIC CLAUSE TYPES IN HOMERIC GREEK

Adapted from Horrocks (1980: 14,15)

1) Nominal & Copular Clauses⁷

S + Csubject

B 204: [οὐκ] ἀγαθὸν πολυκοιρανίη
(uuk) agathon polukoiranīe
[not] a good thing [is] many-lords

S + Cplace
(state of being)

A 515: [ἐπεὶ οὐ τοι] ἐπι δέος
(epeē uu toi) epi deos
[since there is nothing] to/for fear

⁵ In more recent developments in Generative Grammar the distinction between lexical and functional categories has been revived and KASE as an abstract functional category heads an NP and receives specific morphological case from the verb to mark on that NP. For languages such as German and Greek, where the morphological case is added to the stem as a suffix not a prefix, this would support KASE being manifested as a *postposition*, which is in agreement with the notion of a historic SOV word order. The tendency for the prepositional particles to become fixed, preceding the KASE phrase is in agreement with the changing trend to SVO word order. Postpositions do occur in both Homeric and Classical Greek, but they are rare in comparison to the prepositional forms.

⁶ The term "specify" is used in a semantic and a syntactic sense. The particle takes up the specifier position of the KASE-phrase (X-Bar Theory) and narrows the meaning of the KASE relation. The term is not used in the sense employed in historical theories of word-order change.

⁷ Horrocks argues, according to Kahn (1973), that the difference between nominal and copular clauses is trivial, because omission of the copula is so frequent in Greek. c.f. Horrocks (1980:15)

| | | |
|--|--------|--|
| S + Csubject + Vcop | β 29· | [ῆ] οἱ προγενέστεροί εἰσιν (εε) ἡοὶ προγενεστεροὶ εεσιν [or] those-who older (ones) are |
| S + Cplace + Vcop | A 63: | [τῷ] ὄναρ ἐκ Διὸς ἐστίν (l) ὄναρ ἐκ διὸς ἐστίν [a] dream from Zeus is (source) |
| 2) <u>Intransitive Clause</u> S + Vintrans | O 101: | ἡ [δὲ] γέλασσε hee (de) gelase she [but] laughed |
| 3) <u>Simple Transitive Clause</u> S + Odirect + Vtrans | O 328: | ἄνῆρ ἔλεν ἄνδρα aneer helen andra man (nom) slaughtered man (acc) |
| 4) <u>Complex Transitive Clause</u> S + Odirect + Cobject + Vc-trans | θ 20: | μῖν μακρότερον... θῆκεν min makroteron ... theeken him taller ... (she) made |
| S + Odirect + Cplace + Vc-trans | θ 441: | ἄρματα [δῶ] ἄμ βωμοῖσι τίθει harmata (d) am bowmoisi tithee the chariot's gear [and] upon a stand - he-set |
| 5) <u>Di-transitive Clause</u> S + Odirect + Oindir + Vdi-trans | ρ 287: | ἧ ... κακῶ ἀνθρώποισι δίδωσιν hee ... kak anthrooisi didosin which ... evils (upon) men (dat) gives |

a = lower case letter - source is the Odessey

A = capital letter - source is the Iliad (Same throughout this document)

S = subject

O = object

V = verb

C = complement

[] = non obligatory clause element(s)

place: includes source, goal (also when a person) location and state (of being)

The following data, which illustrate the phenomena of parallel environments, employ a variety of particles. I have attempted to demonstrate the two syntactic functions with a semantically matched pair of particles, but where the example pair does not include a semantic match, the reader may assume that both the adverbial and prepositional functions actually exist for each particle in the pair (from Horrocks: 16 - 18):

(4a) Prepositional Phrase as Complement in Verbless Copula clause

F 110 ἀλλῶ ἐπι τοι καὶ ἐμοὶ Θάνατος καὶ μοῖρα κραταιή
 all epí toi kai emoi thanatós kai moira krataiē
 yet - over - the(dat)- also- me(dat)- death(nom)- and - fate(nom) - hangs
 “yet over me also, hang death and fate”

(4b) Adverb as Complement in Verbless Copula clause

p 315 οὐδῶ ἐπί Φειδῶ
 uud epí pheidō
 neither [is] (there) over sparing (nom)
 “(they) show no restraint”

(5a) Prepositional Phrase as Complement in Overt Copula clause

O 607 ἀφλοισμὸς δὲ περὶ στόμα γίγνεται
 aphloismós de peri stoma gignetai
 1(nom) - and - about - (his) mouth (acc) came into being
 “and he foamed at (around) the mouth”

(5b) Adverb as Complement in Overt Copula clause

v 343 ἔνθα δῶ ἅνα σταφυλαὶ παντοῖαι ἔασιν
 entha dana staphūlai pantōiai easin
 there - and - upon(them)- clusters (nom) - of every kind - are
 “and there, upon them, are all kinds of clusters”

(6a) Prepositional Adjunct of Place

B 4 ὀλέσει δὲ πολέας ἐπὶ νηυσὶν Ἀχαιῶν
 olesei de poleas epí neūsin akaiōn
 (he)destroy (subj) - and -many (acc) - beside- (the) ships(dat) - Achaeans (gen)
 “and (how) he might destroy many (soldiers) beside the ships of the Achaeans”

(6b) Adverbial Adjunct of Place

a 273 Θεοὶ δῶ ἐπι μάρτυροι ἔστων
 theoi d epí martūroi estōn
 gods (nom) - and - at (it) - witnesses - be (imper)
 “and let the gods be witnesses at there (the assembly)”

(7a) Prepositional Complement in Complex transitive clause
 δ 58 παρά δέ σφι τίθει χρύσεια κύπελλα
 para de sphi tithei khrúseia kúpella
 by - and - them (dat) - set (he)- golden - goblets (acc)
 "and he set golden goblets by them."

(7b) Adverbial Complement in Complex transitive clause
 Φ 364 ὑπὸ τὲ ξύλα κάγκανα κεῖται
 hýpo te ksúla kánkana keetai
 under (there) - and - firewood - dry (acc) - is set
 "and dry firewood is set under there"

1.3 Prepositional Tendencies of Particles

It is important to note that the independent adverbial particles in examples (4b), (5b) (6b) and (7b) are all implicitly prepositional or pro-prepositional⁸ forms; that is: they have a covert KP that can be recovered from the context. These forms are similar to the modern German forms *davon* or *dadurch*, or the older English forms (surviving only in legal language today) *thereon* and *thereto* (Horrocks: 18). Again, the prepositional phrases are not fully prepositional, rather the prepositional particle specifies the complement KASE-phrases. This is apparent because in Homeric Greek adverbial KASE-phrases are able to stand alone without particle support, as in the following examples (Horrocks: 19):

(8) Locative - expressed with dative case (Adjunct)
 Δ 423 πόντῳ μὲν τε πρῶτα κορύσσεται
 pōntō (men te) prōta korúsetai
 the open sea (dat) - at the first (acc) - (it) is gathered into a crest
 "Upon the open sea is it (the wave) at first gathered into a crest"

(9) Goal expressed with accusative case (Complement to V-of-motion)
 A 322 ἔρχεσθον κλισίην
 erkhesthōn klisiēn
 go (imperative) the hut (acc)
 "Go to the hut...!"

⁸ In the framework of Generative Grammar, pro-forms are covert noun phrases. The importance of the pro-form to Horrocks argument is that pro does not receive case. Consequently, in these forms, the particle is implicitly functioning like a preposition, yet it is not case assigning.

(10) Path expressed with accusative case (Complement & Adjunct)

γ 71 ποθεν πλειθῶ ὑγρά κέλευθα
pothen pleith hūgra keleūtha
where (acc) - sail (you) - watery - ways (acc)
"To (C) where are you sailing over (A) the watery ways?"

(11) Source expressed with genitive case (Adjunct)

Ο 655 Ἀργεῖοι δὲ νεῶν μὲν ἐχώρησαν
argeeoi de neōn men ekhōrēsan
the Argives(nom)- and- ships(gen) - gave way
"and the Argives gave way (retreated) from the ships"

This optionality of the particle in the KASE-phrase is further indicated by the ability of the clitic co-ordinating particle to intervene between the preposition and the rest of the phrase, as in example (7a). In the majority of the data presented here, however, the clitic actually precedes the prepositional particle. The fact that the specifying particle is generally preferred indicates that the stage of development is approaching a full prepositional status.

Complement particles which have traditionally been termed "preverb" particles also demonstrate prepositional tendencies, as indicated by examples (12a) and (12b), where the particle simultaneously specifies both KP and the verb (taken from Horrocks: 21):

(12a) Preverb Particle in tmesis, before the KASE-phrase

π 337 πάις ἐκ Πύλου ἦλθεν
pais ek pūluu eelthen
son(nom)- out- from Pylos(gen)-came
"(her) son came (out) from Pylos"

(12b) Preverb Particle

υ 371 ἐξεῆλθε δόμων
ekseelthe domōn
out-from came (he)- the halls (gen)
"He came out from the halls"

Prepositional particles are also constituents within *prepositional verbs*. English correlates to these prepositional verbs would be *look at*, *rely on*. Compared to the forms in (12a and b), the relationship between the particle and the verb in prepositional verb constituents is much more idiosyncratic and the verb is also non-copular. With a verb of motion (as in ex. 12: *come*), almost any other spatial particle may be used. (i.e. in, out, from, to, etc.) In contrast, the prepositional verbs are restricted in particle combination; as demonstrated in the form "speak to" (Horrocks: 24):

(13) Prepositional Verb

- υ 128 πρὸς δὲ Εὐρύκλειαν ἔειπεν (*ανα.../*δια.../*παρα...)
prōs d eurŭklicean eipein (ana.../ dia.../ para...)
to - and - Euryclyia - spoke (he) (* up .../ * by.../ * along ...)
“And he spoke to Euryclyia...”

Although the particle within prepositional verb constituents is more closely connected with the semantics of the verb than the KASE-phrase, it may still be viewed as specifying both constituents equally. Because the Adverb phrase (Particle+KP) is a complement, the particle may stand immediately before the verb and separate from the KP or it may stand independent of both the verb and the KP (Horrocks: 25):

(14) Particle immediately before Prepositional Verb

- A 206 τὸν δὲ αὐτὴ προσέειπε θεὰ
ton daute prōs eipe thea
him (acc) - then- she to - said - (the) goddess
“Then the goddess said to him”

(15) Particle independent of both Verb and KASE-phrase

- δ 803 καὶ μιν πρὸς μῦθον ἔειπεν
kai min prōs mŭthon eipein
and - her - to - spoke - said (she)
“and spoke to her and said..”
or “and she spoke, and said to her”

1.4 Adverbial Tendencies of Particles

There are also particles which occur pre-verbally but never co-occur with a KASE-phrase, either overtly or covertly. These are grouped into two subtypes, depending on whether or not the particle may stand in tmesis or not. Those that can undergo tmesis are constituents in *phrasal verbs* (c.f. 16 a & b). Those particle + verb combinations which *can not undergo tmesis* are considered *true compounds* in Horrocks' framework (c.f. 17a & b). The adverbial particles in true compounds are non-resultive in meaning. They add an aspectual sense to the verb, indicating completeness (fullness) of the action, but do not indicate a change in state for the affected entity (Horrocks; 1980: 26):

(16a) Particle + Phrasal Verb

- E 758 ἀπόλεσε λαὸν Ἀχαιῶν
apōlese laon akhaiōn
away-kill a host (acc) Achaeans (gen)
“he has destroyed a host of the Achaeans”

(16b) Particle in tmesis (separation) from verb

Ω 609 τῷ δῶρα ... ἀπὸ πάντας ὄλεσαν
τω̄ dara —apo pantas ὄlessan
they - then, so — away - all of them - kill
“so then, they destroyed all of them.”

(17a) Compound verb

ζ 49 ἄφαρ δῶπεθαύμασ' ὄνειρον
aphar darēthaūmas ὄneiron
immediately - and - at-wondered (she) - (her) dream (acc)
“And immediately she wondered at (was amazed by) her dream”

(17b) Ungrammatical when particle in tmesis (Horrocks: 26)

(ζ 49) * ἄφαρ δ' ἀπὸ ὄνειρον ἐθαύμασῃ
* apha r d ap ὄneiron eithaūmas
* 1 1 - and — at- (her) dream (acc) - wondered (she)

Finally, phrasal verb units should be contrasted with structures where independent prepositional particles occur as complements to copular or complex transitive verbs (c.f. again examples 5a and 7a). In phrasal verb units, the verb is not a copula or complex transitive and the particle has a non-literal or abstract sense, which indicates achievement of a *new state* or *resultative condition*, rather than the new location of the affected entity or a resultant location (Horrocks: 27):

(18a) Prepositional particle with complex transitive verb

Κ 72 ἀπέπεμπεν ἀδελφεὸν
apēpempēn adelphēon
away-sent (he) (his) brother(acc)
“...he sent away his brother...” (change of spatial location)

(18b) Adverbial particle in phrasal verb

τ 230 ὁ μὲν λάε νεβρὸν ἀπάγχων
ho men lae nebrōn apēkhōn
he (dog(nom)) - but- was pinning - the fawn - (acc) away strangling (it)
“...but the dog was pinning the fawn and strangling it (to death).”
(abstract sense)

1.5 Summary of Particle Distribution

Horrocks reduces his classification of the particles to the summary format presented in Table 2, where the types of particle + verb combinations are listed vertically, and the distinguishing criteria are listed horizontally. It is not necessary for the line of argumentation pursued here, to

TABLE 2: GREEK PARTICLE DISTRIBUTION
Adapted from Horrocks (1980: 29)

| | 1. P/P+KP omissible | 2. particle prepositional | 3. adverbial Q-Form | 4. pronominal Q-form | 5. free-choice particle | 6. partical preverbal | 7. particle in tmesis | EX. |
|--------------------------|------------------------|------------------------------|------------------------|-------------------------|----------------------------|--------------------------|--------------------------|----------------|
| 1. Adjunct Kase P | + | + | + | 0 | + | 0 | 0 | 6a 8-11 |
| 2. Complement Kase P | 0 | + | + | 0 | + | + | + | 4a 5a 7a |
| 3. Prepositional Verb | P only omissible | + | 0 | + | 0 | + | + | 12a,b 14,15 |
| 4. Phrasal Verb | 0 | 0 | 0 | + | 0 | + | + | 16a,b |
| 5. Compound Verb | 0 | 0 | 0 | + | 0 | + | 0 | 17a |

P = Particle
Kase P / KP = Kase Phrase
Q = Question

discuss his distinguishing criteria more thoroughly. From this point onward, I will be examining more closely the semantic nature of the particles in the first three categories: 1) Adjunct Kase-phrase; 2) Complement Kase-phrase; and 3) Prepositional verb units. These are the units where Horrocks has demonstrated that the particle has prepositional tendencies, yet it has not been fully reinterpreted as a preposition.

The *preverb position* which is identified as a role in the traditional assessment of the particles, has been shown by Horrocks to exhibit both prepositional and adverbial tendencies. This "preverbal" position of the particle is relevant for discussions of univertation (compounding), which is currently identified by syntacticians as the process of incorporation (Baker; 1988). This process overlaps and interacts with the particle reinterpretation process, complicating the analysis. Horrocks views particle movement by both prepositional and adverbial forms into the preverb positions as a unified syntactic process. He also views the position of particles in phrasal verbs, as having been generated in base structure, while *true compounds* are considered lexical combinations (Horrocks: 50-84). There are some problems with Horrocks' interpretation of the syntactic process, in terms of the constraints placed upon incorporation, as stated in Baker's theory (1988). However, Baker's theory is designed to explain a synchronic process, and it is appropriate

to consider whether or not normal syntactic constraints are violated when a language community is reinterpreting and restructuring syntactic functions. We will return to these questions after a consideration of the semantic constraints on the reinterpretation process.

In keeping with the introductory questions about constraints on reinterpretation as a mechanism of language change, Section 2 will now outline the general characteristics that constrain Greek adverbial reinterpretation. In Section 3, more exact stages for the reinterpretation process will be outlined, stating the evidence that indicates when each stage has been reached. Support for the reinterpretation process as an Indo-European phenomenon is provided through summaries of word-order data from Sanskrit, Hittite, Homeric and Classical Greek, and with secondary references to prepositional development in Mycenaean.

2. CONSTRAINTS ON REINTERPRETATION

2.1 Semantic Constraints

It is important to note the semantic characteristics of the particles that undergo reinterpretation. These particles, as independent adverbs, are limited to the group which indicate temporal and spatial location or direction of motion. When the particles take on a specifying relationship with the KASE-phrase, they associate with KPs that are also bearing some temporal or spatial relationship to the verb (complement), verb phrase or main clause (adjunct) (c.f. examples 8-11). Particle + KP units are able to express location, motion and orientations that involve proximity of participants or relationship to either a goal or source. These relations apply in both the spatial and temporal domains (c.f Horrocks: 180-287 for an in depth treatment of the semantics of Greek Prepositions). Although the use of particles to express temporal notions is more abstract than the locative uses, the temporal use is consistent and predictable across languages.

Non-literal and idiomatic uses of the particles in phrasal verbs were perhaps originally "motivated" by spatial schemas (c.f Lakoff; 1987: 416-61 - on analysis of metaphor and figurative language involving spatial prepositions); however, these are quickly subject to semantic drift. Eventually the meaning of the preposition no longer has a transparent relationship to one of the KASE functions it is associated with.

This restricted semantic field⁹ of the particles does constrain the reinterpretation process, but it is also apparent that the **semantic ambiguity** created by the conflation of **case functions** within the case system (paradigm)¹⁰ is likely responsible for the creation of the new function of the particle as a specifier of KASE.

2.2 Conflation of the Morphological Case System

The following (Table 3) is the set of singular and plural case-endings posited for PIE by comparative researchers. I have omitted the dual forms, because they already show a high degree of syncretism even before the daughter languages separated from PIE.

Table 3
PIE Case Endings
(Adapted from Palmer; 1980: 267)

| | Singular | Plural |
|--------------|----------------------|--------------------|
| Nominative | - s, Ø | - es |
| Vocative | - Ø | - es |
| Accusative | - m* | - ns* |
| Genitive | - es/-os/-s | - om/-o:m |
| Ablative | - es/-os/-s, -ed/-od | - bh(y)os, mos |
| Dative | - ei | - h(y)os, mos |
| Instrumental | - e/-o, bhmi/-mi | - bhmis/-mis, o:is |
| Locative | - i | su |

* Syllabic forms also

As Table 3 indicates, the dative and ablative case endings are considered to be the same for PIE; however, in Homeric Greek, the dative and instrumental cases have conflated with the locative case, and the ablative case (source) is conflated with the genitive case. The conflation is attributed to the phonological and functional similarity of these cases (Smyth; 1959: 312 - Par. 1279). However, the phonological similarity is likely the true source of the change. Phonological reduction of the case endings actually created functional ambiguity through **composite cases**; and

⁹ A semantic field is a group of related concepts that usually have a set of identifiable lexical labels. They can usually be organized hierarchically.

¹⁰ The term KASE is used when speaking of the syntactic position at the head of the phrase; the more familiar term CASE is used when referring to the actual morphological realization of a specific case or discussion of the cases as a paradigm.

this should be identified as the motivation for the new function of the particles: specifying KASE to prevent the loss of semantic distinctions. The presence of a prepositional particle in the phrase is preferred, though not obligatory, in Homer, because the particle narrows the meaning of the composite cases (Smyth: 365, 374, 378).

By analogy, the tendency of the particle to be associated with composite cases would spread to the entire case paradigm (particle + KASE), occurring first in those uses where the non-composite cases are functioning to indicate spatial location or motion or temporal relations, then extending to more abstract uses of the particles.

For example, accusative case in Greek is used to indicate a direct object complement or motion towards a location. Consequently, after adverbial reinterpretation, the direct object is marked by accusative case alone, whereas the relational idea of "motion" is indicated by a preposition (eis) + accusative case. The composite 'dative' case alone usually indicates the recipient, indirect object or beneficiary (dative functions of the eight case system), but in combination with a preposition (en), it is locative in function. Even after the conflation and reinterpretation processes occurred, instrumental uses of the dative case can still occur with or without a preposition, since the meaning of the noun phrase itself is sufficient to disambiguate the sense of the case.

The semantic ambiguity caused by phonological reduction of case endings can be seen as the primary motivation for the adverbial particles to develop prepositional tendencies. However, the new functions of the particles create greater ambiguity in the total system, as **both particles and cases have composite functions**. According to Horrocks, the particles are not able to become fully prepositional (reinterpreted as Preposition + KASE) until the independent adverbial forms are renewed in their morphology, making it feasible for the multiple functions to be distinct once more.

2.3 Renewal of Independent Adverbial Forms

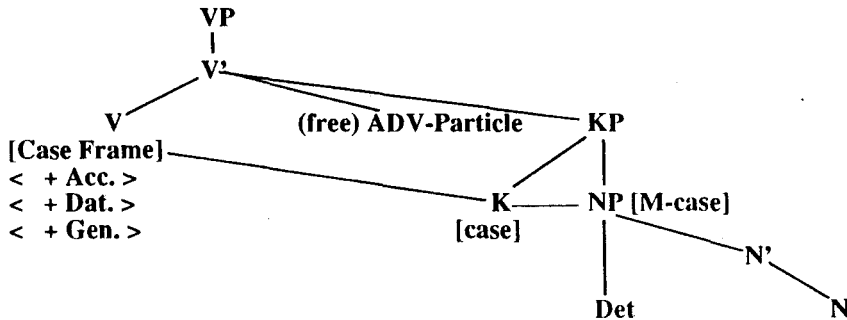
Horrocks cites Kurylowicz (1964) in arguing that the development of both prepositions and univerbation in IE languages such as Sanskrit (Vedic) and Greek were dependent upon a formal renewal of the old adverbial forms in their *primary independent function* (Horrocks: 117-18). These primary functions are deictic locatives, such as: here, there, outside, etc. Table 4 gives examples of the old and renewed forms:

Table 4
Primary Adverbial Form Renewal
 (Adapted from Horrocks; 1980:118;142)

| Old | Vedic | | Greek | | Mycenean |
|------|---------|---------------------|-------|-------------------------------------|-----------|
| | Renewed | Gloss ¹¹ | Old | Renewed | Renewed |
| úd | uccá: | “on both sides” | amphi | amphotero:then, amphotero:se | apoterote |
| ní | niccá: | “outside” | ek | ektos, ektosthen | |
| abhí | abhítah | “here” or “there” | en | entha, enthen, enthade, evthende | ete |

Once the particles have renewed forms to take up their primary functions, the old forms are then free to be used as KASE specifiers or to compound with the verb. The final step in preposition development is the reinterpretation of the relationship between the particle and the abstract KASE position which heads the phrase, as in Figure 1C. The particle is reinterpreted as being head of the phrase. Horrocks and most Greek grammarians would agree that it also takes over the case-assigning property from the verb.

Figure 1A) KASE ASSIGNED BY THE VERB'S CASE-FRAME



¹¹ I am assuming that Horrocks has chosen Vedic adverbials equivalent to the Greek forms. He does not include glosses in many of his examples. Provided this assumption is correct, then the gloss applies to both the Greek and Vedic. If the assumption is not correct, then the gloss is applicable only for the Greek forms.

- 1) Initial change from OV to VO has occurred
- 2) Adverbial Particles indicating spatial location and motion are likely complements to the Verb, locating them under V'
- 3) See Footnote 12 for clarification on the concept of 'case-frame'
- 4) M-case = Morphologically-manifested Kase

Figure 1B) ASSOCIATION OF THE ADVERBIAL PARTICLE WITH KASE

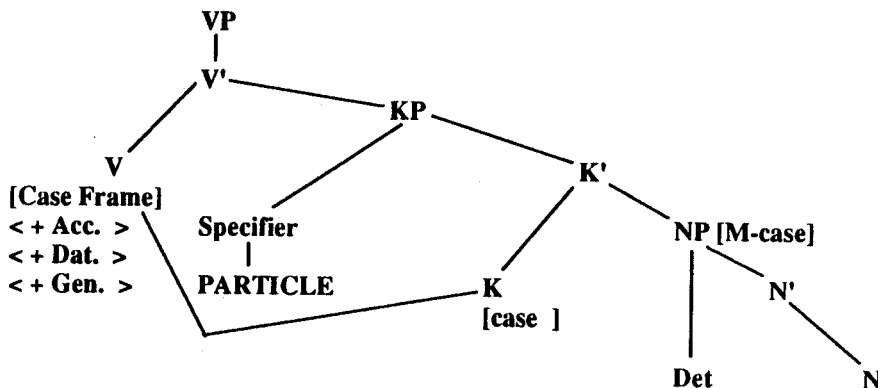
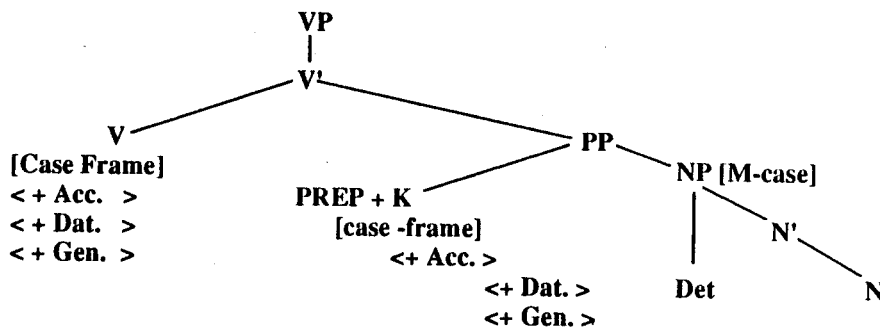


Figure 1C) REINTERPRETATION OF THE PARTICLE AS A KASE ASSIGNING PREPOSITION



Once reinterpretation has occurred, the position of the particle becomes fixed, as in Figure 1C. Since the particle heads the phrase, it must always occur with it. The fixation of the

particle in the prepositional rather than a postpositional location indicates that, although Greek employs a great deal of variation in phrase order, it has switched to a predominantly VO word-ordering. OV ordering is still frequent in subordinate clauses. In order to get an adverbial reading rather than a prepositional reading of the particle, it has also become necessary to fix the adverbial position relative to the verb, in preverb position. This in turn restricts the availability of the univerbation process to the adverbial reading.

In summary, reinterpretation was preceded by the renewal of the primary adverbial forms, which was likely preceded by the conflation of the case system that created the third syntactic function for the particles. For example, before the renewal, the Greek particle "ek" - occurs in: a) its primary function - "outside" (locative); 2) its adverbial function - "out" (direction); and 3) prepositional specification of the genitive case - "from" (formerly the ablative case function).

Since these factors can be clearly identified as motivating the reinterpretation process, it is now possible to set in order the stages of the process and to present word order evidence which indicates that the process is completed in Classical Greek.

3. PARTICLE REINTERPRETATION & WORD-ORDER:

3.1 Stages of Reinterpretation

The stages in the reinterpretation process may be summarized as follows:

1) Conflation of the case system creates ambiguity. It seems that semantic relations indicating spatial and temporal location, orientation or motion are particularly susceptible to ambiguity when the case system conflates.

2) Particles take on a specifying function in relation to the ambiguous KASE-phrases. At this early stage, any compatible pair of particle and case-ending may co-occur (Horrocks: 119). This would mean that particles would be specifying case in spatial (and perhaps temporal) adjuncts, and in complements of verbs of motion.

Both of these stages are previous to the data that Horrocks presents, since all of the daughter languages which he surveyed, and perhaps even PIE have gone beyond this stage. He begins the stages with step two and assumes that these constraints (stage 2) are justifiable for the earliest stage of the process, even though the stage is not attested (Horrocks: 120). As indirect evidence, it is valuable to note that, in Greek, the particles which associate with only one case specify the 'spatial' case distinctions which would have been lost between PIE and Greek (such as 'apo + genitive' to indicate the 'ablative' function; or 'en plus dative' to indicate 'location' - c.f. Reading Greek: 1988: 290-92).

3) The semantic relationship between particle and KASE-phrase causes them to be syntactically grouped together. Then, by reinterpretation, purely syntactic co-occurrence restrictions can develop between particles and case-endings that are semantically independent, since one of the particle's functions is case-specification. This situation is manifested in the prepositional and phrasal-verb constituents and is the stage of development which is attested in Homeric Greek. According to Horrocks, it is the stage which should be attributed to PIE (Horrocks: 121). Some of the word-order evidence that Horrocks presents to support this claim will be presented in Section 3.2.

4) The primary function of the particles is renewed, as demonstrated in Table 4. This step must precede the final stage of reinterpretation.

5) Reinterpretation of the prepositional particle as (head of the phrase and) Kase-assigner occurs with the parallel (or consequential) reinterpretation of the preverb position as strictly "adverbial". Enclitics can not intervene between either the preposition and KP or the adverbial particle and verb units. Adjunct adverbial clauses no longer appear without a prepositional head. This is the state of development found in Classical Greek.

The following examples indicate that adverbial particles and prepositions are distinct functions in Classical Greek. The compound (incorporated) form of preposition + verb, was possible in Homeric Greek (c.f. example 19a) with two possible interpretations. In Classical Greek, however, the adverbial nature of the compound must distinguished and a doubling structure is employed to achieve the adverbial interpretation of the particle (as in 19b):

19a) **Incorporated Preposition (possible in Homer only):**
οἱ ἄνδρες ἀφαιροῦσι τὰ ὅπλα τῶν πολεμίων
hoi andres - aphairouusi - ta hopla - ton polemion
the men - from-take - the weapons - the enemy
nominative - 3rd.pl. pres. active - accusative - genitive
"The men take the weapons from the enemy."
or "The men take the weapons away from the enemy."

19b) **Doubling Structure (Required in Classical) & Incorporated Adverb:**
οἱ ἄνδρες ἀφαιροῦσι τὰ ὅπλα ἀπὸ τῶν πολεμίων
hoi andres - aphairouusi - ta hopla - apo ton polemion
the men - away-take - the weapons - from the enemy
nominative - 3rd.pl. pres. active - accusative - Prep + genitive
"The men take away the weapons from the enemy."

Compare these to examples 12a & b where the particle simultaneously specifies both the verb and the Kase-phrase. It is controversial whether or not the particle may have taken over headship of the phrase at stage 3. Headship and case-assigning abilities are linked together in Horrocks' analysis at stage 5. However, according to the constraints in Incorporation Theory, the particle must already be a head in order to incorporate (Baker, 1988: 51-54) with the verb, which is head of the verb-phrase and incorporation of the "prepositional particle" does occur in Homeric Greek.

Horrocks suggests that if the prepositional and adverbial functions were distinct syntactically, rather than composite functions of a single syntactic entity of particle, then the prepositional particle would not be able to move out of the Kase-phrase into preverbal position and undergo univerbation or incorporation (c.f. examples 2b and 12b in this paper; Horrocks: 122). However, this interpretation seems to anticipate the full adverbial status of the preverb position, which is not complete until stage five. Also, there is plenty of evidence from other languages that

Table 5: Enclitic Positioning

Sanskrit (Horrocks: 93-95)

Type (a) P (E)...V

prā vām.....bhārante

"they are brought forth...to you"

Type (a) V (E) ...P*

bādhante...abhimāñnam āpa

"they drive off adversaries"

Type (b) {C/.}(E)...PV

te...prā bhārāmasi

"we bring forth...to you"

Type (c) {P/C/.}(E)V...

antāh pāsyanti vṛjinā

"they see into evil"

Type (d) {C/(V)/.}(E)(V)...P NP_{do} (...)(V)
- one of the V positions is obligatory

ānuvratām āpa jāyām arodhau

"I have driven away a devoted wife"

Hittite (Horrocks: 96)

Type (a) P (E)...V

šer-wa-ši šarnikmi

"I shall make restitution for him"

Type (b) {C/.}(E)...PV

nu-kán tamedani kuedanikki andan
paitteni

"And you go in to someone else"

Homeric Greek (Horrocks: 101)

Type (a) P (E)...V

an de subootees histato

"then the swineherd stoodup"

Type (b) {C/.}(E)...PV

opse de dee Menelaos *an* istato

"and he saw Menelaos standup"

Type (d) {C/.}(E)...P NP_{do} (V)

hai d'apo (men) siton eaireon

"these began to take away... food"

Type (e) P (E)V...

apo d'oolese laous

"(away) and others he has slain"

Type (f) PV(E)...

ap ekrupsen *de* moi hippous

"but hid away my horses"

Type (g) {C/.}(E)PV...

su men nun hoi parathes kseineia

"but you now set before (her)
entertainment."

V = Verb

P = Particle

E = one or series of enclitics (connectives or pronouns)

C= sentence initial connective or pronoun

. = space to be filled by a single constituent

..... space to be filled by one or more constituents

() = optionality of elements enclosed

{ } = exclusive choice between elements enclosed

prepositions can undergo incorporation processes (c.f. Baker; 1988: 229-304) and the complexity of the case-frame¹² of Greek verbs means that, potentially, a verb could assign the case associated with an incorporated preposition, even if it is not accusative case (c.f. again - the case-frames indicated in Figure 1).

This means that between the stage indicated in Figure 1B and that of Figure 1C, there could be an intermediate stage where the verb and the prepositional particle employ a mode of co-operative government, since the particle cannot always be unambiguously interpreted as the nearer governor (as Baker's theory would require). The inability of the prepositional particle to incorporate in Classical Greek may not be the best evidence that the particle has become a head or a case-assigner. Instead, the binding of the adverbial particle to the verb to disambiguate its adverbial function from the prepositional function could itself be the reason why prepositional-particle incorporation ceases in Classical Greek. The preverb position has been reinterpreted as indicating the adverbial function of the ambiguous particle forms and is no longer available for the preposition incorporation process.

However, if one does not consider the notion of co-operative government to be theoretically feasible, there could also be an intermediate stage where the particle could be a head and still not be a case-assigner. The intervention of enclitics between the particle head and NP should still be possible at this stage. So the absence of intervening enclitics is actually the best evidence that full prepositional status (or two syntactically distinct functions) has been achieved¹³. Consequently, univerbation (incorporation) and the completion of particle reinterpretation could be seen as parallel reinterpretations, which are achieved in stage five.

3.2 Word-order Evidence and Clitic Placement

Table 5 indicates clitic placement relative to adverbial particles in phrasal verbs in Vedic Sanskrit, Hittite and Homeric Greek. The Greek word orders apply equally for prepositional particles in Homeric Greek that have been moved out of a complement Kase-phrase (Horrocks: 112), that is: incorporated "prepositional" particles.

In each of the languages it is clear that the enclitic may still intervene between the verb and particle, but forms where the particle univerbates in preverb position are also attested. According to the stages outlined previously, these major Indo-European languages all show

¹² In most of the languages that Baker examines, the verb can only assign accusative case (structural). Greek verbs can assign genitive and dative as well as accusative case structurally to their innermost objects (indicated by the ability of these objects to undergo passivization), so consequently, they could assign the appropriate case (other than accusative) to maintain the meaning of an incorporated preposition (Smyth, 1959, pp. 355-56, 396). This is important because the meaning of the preposition changes depending on the case it is combined with.

¹³ My thanks to Eithne Guilfoyle (University of Calgary) for her clarification of this point, in conversation.

particle development at about stages 3 and 4. (Data on renewal of primary adverbial forms — stage 4 — is not available for Hittite in Table 4.)

4. CONCLUSIONS

Horrocks argues from comparative word order data, that prepositional tendencies in particle function must have been already in progress in late PIE. The fact that the 8-case system of PIE (Table 3) shows syncretism might also support this view, providing that one accepts the idea that syntax can actually project backward from the state of the daughter languages to PIE.

Horrocks also argues extensively, that preposition development in Mycenaean is complete and comparable to Classic Greek. However, glosses were not available for the Mycenaean data, so I was not able to assess it or present any of it here. On the basis of his analysis of Mycenaean, Horrocks argues that Homeric Greek preserves very ancient Greek forms of syntax, that are earlier than the state of prepositions in Mycenaean and that these forms would be very close to the stage of particle development in late PIE (Horrocks: 108, 109). So Horrocks leans towards the view that the particle reinterpretation process is an “inherited” feature of the daughter languages, although he acknowledges that it could just as easily be a parallel process within the family.

It is important to note, however, that the motivation for and constraints on the reinterpretation process would allow for some variation in its manifestation within the family. The difference in preposition and particle forms in English (ie: away + from) in comparison to the lexical ambiguity of Greek, indicate that the forms for both the primary and adverbial particle functions are renewable, rather than the latter function being disambiguated through a fixed preverbal position. Since conflation of the case system is the first step in the process, in principle, the particle reinterpretation process could be thwarted altogether in some of the daughter languages, by a renewal of their case systems. IE languages also exhibit differences in P+ Kase combinations, based on the extent of syncretism in their case systems and the potential redundancy that is inherent in the process (eg: Greek “ek” and “apo” - can both mean “from”). A more in depth study of the language families with regard to this phenomenon would yield more data, indicating the range of variation in IE languages that is permitted within each of the stages of the process.

In spite of the variation that is bound to occur in reinterpretation, it is also clear that this shift of function from particle to preposition is both highly motivated and highly constrained by semantic factors in its early stages. There are many other kinds of particles and connectors in Greek, yet it is this small, semantically distinct group which enters into the reinterpretation process. This shift, from dependence on morphological case marking to dependence on prepositions and word-order to mark relations, is a very frequent and major type of syntactic change. It also appears to be subject to more specific requirements than the “broad constraints of UG”, which Lightfoot asserts to be the only boundaries capable of constraining a reinterpretation process.

REFERENCES

- Baker, M. C. 1988. *Incorporation*. Chicago, USA. University of Chicago Press.
- Bellusci, D. C. 1991. Serialization Patterns in Shona Verb Morphology, *Calgary Working Papers in Linguistics*. No. 14. Fall 1991. Calgary, Canada. Department of Linguistics.
- Horrocks, Geoffrey, C. 1980. *Space and Time in Homer: Prepositional and Adverbial Particles in the Greek Epic*, Salem, New Hampshire, USA. Ayer Publishers, Inc.
- Joint Association of Classics Teachers. 1978 (1988 Printing). *Reading Greek*, New York USA. Cambridge University Press
- Lakoff, George. 1987. *Women, fire and dangerous things: What Categories Reveal about the Mind*, Chicago, USA. University of Chicago Press.
- Lehmann, W.P. 1974. *Proto-Indo-European Syntax*, Texas, USA. University of Texas Press.
- Lightfoot, David. 1980. On Reconstructing A Proto-Syntax. In: Ramat, Paolo, et. al. (eds.) *Amsterdam Studies in the Theory and History of Linguistic Science IV - Current Issues in Linguistic Theory*, Vol. 19. Amsterdam, Holland. John Benjamins B. V.
- Murray, Robert, W. 1989. Personal Communication. Calgary, Canada. Department of Linguistics. University of Calgary.
- Palmer, Leonard, R. 1980. *The Greek Language*, Cambridge, Great Britain. Faber and Faber Limited at the Cambridge University Press.
- Smyth, H. W. 1956. *Greek Grammar*, Second Edition. Harvard.
- Travis, L. 1986. The Case Filter and the ECP, *McGill Working Papers in Linguistics*, Vol.3, No. 2.
- Venneman, Theo. 1974. Topics, subjects, and word order: From SXV to SVX via TVS. In: J. M. Anderson and C. Jones (eds.) *Historical Linguistics*, John Benjamins B. V., Amsterdam, North Holland.

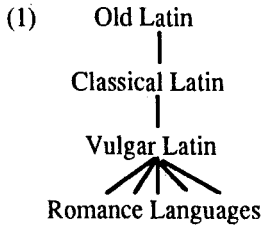
Note: Translation of the Greek into English is my own modernization of an Old English translation provided by: A. T. Murray, Harvard University Press, William Heinemann Ltd. Reprinted 1966 - *Odyssey*; Reprinted 1965 - *Iliad*.

Proto-Romance Syllable Structure

Naomi Cull
University of Calgary

1.0 Introduction

According to a traditional view (see Lindsay 1894; Muller 1929; Elcock 1960; Pei 1976), the Romance languages derived from Classical Latin (CL) with an intermediate stage of Vulgar Latin (VL), as represented in 1.¹



However, there are several reasons to believe that this was not the correct progression. First, there are certain features Classical Latin does not share with the Romance languages. For example, in Classical Latin *cuius* was invariable, yet at an earlier stage, Plautine Latin (c. 184 B.C.), *cuius*, *-a*, *-um* was an inflected adjective. Although the Romance languages also use *kuiu*, *-a* 'whose' as an adjective, Hall (1950:19) claims that this is a feature of Old Latin lost in Classical Latin. Second, there are features of Classical Latin which the Romance languages 'show no trace of' (Hall 1983:5), such as the passive voice, the future tense and most non-finite forms.² Third, the Romance languages also have the adjectival suffix ⁺*arju-* '(person, thing) connected with...' that Classical Latin avoided using (Hall 1983:5).³

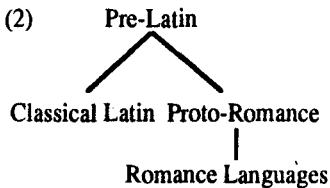
Pulgram (1973) states that the reason there are common features in Vulgar Latin (which Pulgram equates to Proto-Romance) and pre-Latin lies not in the fact that these features were quiescent while the more prestigious Classical Latin was in use, but that in fact those features shared by both Proto-Romance and pre-Latin were in continuous use from the old to the proto stage. However, since these features were regarded as sub-standard they were not reflected in the more highly esteemed Classical Latin.

¹ Based on Hall (1950:24).

² Some of these features may simply have been lost in the Romance languages.

³ Those features evident in the Romance languages but lacking in Classical Latin could also be language internal developments.

Another reason to disregard Classical Latin as the source of the Romance languages is that Classical Latin was a literary form, a language of writing, not speech. As Hall (1974:177) describes it, Classical Latin was an 'artificial' language not spoken by the masses but understood by them. Because linguists try to describe speech as it is spoken, many have tried to equate Proto-Romance (PR) with Vulgar Latin. This too is unsatisfactory as Vulgar Latin is almost identical with a later stage of Romance, Proto-Italo-Western Romance (Hall 1950:8). If Vulgar Latin was the same as Proto-Romance, we would not be able to derive the eastern or southern Romance languages which have features not traceable to Vulgar Latin. Therefore, it is necessary to push Proto-Romance farther back than Vulgar Latin.⁴ Hall (1950:24; cf. Pulgram 1975:38) argues for a tree similar to the one below.



From the branch of Proto-Romance the various Romance languages break off into subgroups and eventually we arrive at the modern Romance languages.

As can be seen in this representation, Classical Latin and Proto-Romance are sisters, Proto-Romance is not the daughter of Classical Latin. This type of representation can best explain why certain features are found in Classical Latin but not in Proto-Romance and vice versa. Of course, Proto-Romance is a reconstruction based on the comparison of the Romance languages and dialects, as Hall (1974:1) emphasizes repeatedly. Therefore, attestations of this stage are not forthcoming. The reliance on the written form however, has hindered historical linguistics in the past. Ancient Sanskrit documents led Indo-Europeanists to believe for a long time that Sanskrit and Proto-Indo-European were one and the same. This of course has been proven incorrect and reminds us not to rely solely on the written word when doing reconstruction.

In this paper I am going to argue that Pre-Latin and Proto-Romance shared the same syllable structure and that Classical Latin was innovative. I believe both Pre-Latin and Proto-Romance had the syllable structure VC\$CV and that in the case of certain consonant clusters, this shifted in Classical Latin to become V\$CCV in order to improve the syllable contact. Thus it is

⁴ Pulgram (1975:42) likewise refrains from using the term Vulgar Latin as an equivalent to Proto-Romance because Vulgar Latin is often equated with post-classical Latin and 'vulgar', meaning non-standard speech, existed prior to Classical Latin.

from this Proto-Romance and Pre-Latin syllable structure that the modern Romance languages are derived and on which some important later phonological processes are based.

From Classical Latin verse and accent assignment there is much evidence to support the claim of $V\$CCV$. For example, when a short vowel was followed by a plosive plus a liquid, the first syllable was light, e.g., $V\$CCV$, as in *tenē\$brae*, *pā\$iris* and *pō\$plus* (Allen 1973:137-138). However, there also exists evidence of an earlier heterosyllabic syllabification of a short vowel followed by a consonant and liquid, for example, an *e* in an open syllable became *ɨ*, yet in *intē\$gra* we have *ɛ*, much like that in *infēc\$ta*, whereas in *infīscit* we get *i*. This indicates then that the original structure was *intēg\$ra* and not **intē\$gra* which would have produced **intī\$gra*.

If Pre-Latin was $V\$CCV$ we would have to argue that it became $VC\$CV$ in early Latin and then returned to $V\$CCV$ once again in Classical Latin. However, it would be difficult to motivate the change from $V\$CCV$ to $VC\$CV$, but there is motivation for the change from $VC\$CV$ to $V\$CCV$ in Classical Latin in syllable contact improvement, as we will see later.

The developments of Classical Latin and Proto-Romance as I believe them to exist are shown below.

(3) Pre-Latin $VC\$CV$

| CL | PR | |
|----------|----------|----------------------------------|
| $V\$CrV$ | $VC\$rV$ | |
| $V\$CIV$ | $VC\$IV$ | |
| $V\$CGV$ | $VC\$GV$ | (where C = plosive or fricative) |

The various diachronic phonological processes discussed below which took place in Italian, Portuguese, Romanian and Catalan will show that Proto-Romance was $VC\$CV$.

2.0 Introduction to the Theory

There have been many difficulties in attempting to reconstruct a standard syllable structure for Proto-Romance. For instance, in dividing a consonant cluster which occurs word-medially, do we divide it as $VC\$CV$ or as $V\$CCV$? Murray and Vennemann (1983) (hereafter M & V) discussed the problem of Proto-Germanic syllable structure and concluded that the correct reconstruction was $VC\$CV$. Some similarities between Germanic and Romance were also reviewed. It is on the basis of this paper, as well as Murray (1987) and Vennemann (1988) and evidence from four Romance languages, that I hope to find support for the Proto-Romance syllable structure shown above. I will begin by first discussing the principles of the Preference Law theory (M & V 1983; Vennemann 1988), the framework in which I will be working. Then I will give evidence from Romance languages which strongly suggests that the correct Proto-Romance syllable structure is $VC\$CV$. Finally, I will discuss the works of other authors (Pensado 1989; Salverda de Grave 1930) which raise questions about the proto-forms and also provide some answers.

2.1 Preference Laws

The overview of the Preference Laws in this paper will be based on Vennemann (1988) with further expansion from other sources. Although Vennemann presents a number of Preference Laws, in this paper I will only offer those which are pertinent to the discussion at hand.

It is well known that syllable structure can help explain certain phonological phenomena such as stress and tone assignment (see Halle and Vergnaud 1987). Syllable structures can also provide the basis for certain sound changes, such as syllable-initial glide strengthening in Modern Spanish (Pensado 1989:128). Vennemann (1988) explains sound changes by referring to the syllable structure of a particular language and the universal Preference Laws for syllable structure.⁵

What is meant by a Preference Law? As Vennemann (1988:1) puts it: 'These laws specify the preferred syllable patterns of natural languages as well as determine the direction of syllable structure change'. *Preference* implies a better/worse condition rather than a natural/unnatural condition. Therefore, the nature of syllable structures is gradient, that is, some syllable structures are more *preferred* than others. "'X is the more preferred in terms of (a given parameter of) syllable structure, the more Y', where X is a phonological pattern and Y a gradable property of X". An improvement to a syllable structure is a syllable structure change, but a change that worsens syllable structure is not a syllable structure change; that is, the change is not motivated by syllable structure according to Vennemann, but is a change which affects syllable structure but is motivated by a different parameter. As an example Vennemann (2) mentions that vowel copations, such as syncope, must always worsen syllable structures, given that the preferred syllable structure is CV. In a sequence of CVCVCV, for example, syncope could produce CCVCV which would be a less preferred, therefore worse, syllable structure.

Diachronic improvements will generalize from the worst to the best syllable structures, as implied in 4.⁶

(4) Diachronic Maxim

Linguistic change on a given parameter does not affect a language structure as long as there exist structures in the language system that are less preferred in terms of the relevant preference law.

Synchronically, a language will not contain structures that are less preferred without also containing structures that are more preferred.

(5) Synchronic Maxim

A language system will in general not contain a structure on a given parameter without containing those structures constructible with the means of the system that are more preferred in terms of the relevant preference law.

However, changes which operate along different parameters may alter the system so that there is

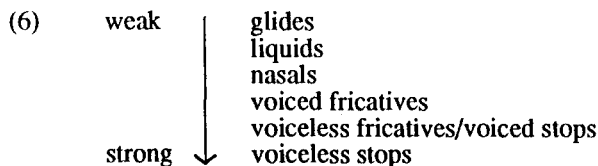
⁵ The Preference Laws are to be construed as universals, however, individual languages can develop their own language-specific tendencies which sometimes conflict with universal laws and so 'unnatural' structures sometimes arise. This is to be expected by virtue of the fact that dialects each have different influences that shape it in various ways. Even so, all languages share certain basic features that tie them together and it is those features to which universal laws can apply.

⁶ For an earlier attempt at explaining the chronology of diachronic changes see Foley (1977).

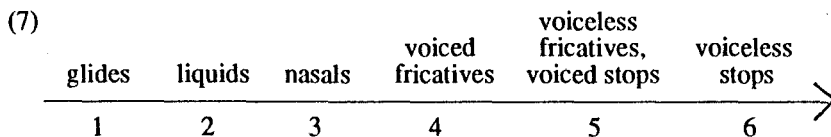
not always an even transition from less preferred to more preferred structures (4). Thus a language system is never 'perfect'.

2.1.1 Consonantal Strength

The phonetic correlates of each speech sound in a language can be placed on a Consonantal Strength Scale. The strength is measured by the 'degree of deviation from unimpeded (voiced) air flow' (Vennemann 1988:8). Segments on the scale are arranged hierarchically from weakest to strongest, as represented in 6 (Murray 1987:118).⁷



These segments can also be put on a numerical strength scale (M and V:519; cf. Hooper 1976:206).⁸



The usefulness of giving each class a numerical value will be made clear shortly.

2.1.2 The Head Law

(8) **Head Law**

A syllable head is the more preferred: (a) the closer the number of speech sounds in the head is to one, (b) the greater the Consonantal Strength value of its onset, and (c) the more sharply the Consonantal Strength drops from the onset toward the Consonantal Strength of the following syllable nucleus.

Part (a) of the Head Law says that the number of speech sounds will be reduced in a head with more than one speech sound, as shown in the Pāli example in 9 where consonant deletion has taken place.

⁷ While a Consonantal Strength scale is considered universal, there may be language-specific values given to each class. That is, in one language voiceless fricatives may be considered stronger than voiced stops rather than equal to them, as is represented in 6. As well, among a particular class certain segments may be considered stronger than others, for instance, labials may be considered stronger than dentals or vice versa.

⁸ For further justification behind such a scale see Foley (1970); Hooper (1976); and Vennemann (1972).

- (9) $\text{prajnā} > \text{paññā}$ 'knowledge'

A syllable with no head is also less preferred and so epenthesis is one means of remedying this situation, as in Northern German.

- (10) ʔalt for alt , chaʔos for cháʔos

The most preferred head is one with the greatest consonantal strength according to (b) of the Head Law. Therefore a weak syllable head is often strengthened as in 11.

- (11) Lt. januarius It. gennaio /dʒ/ 'January'
 Lt. vivere It. vivere 'to live'

Head Law (c) states that the greater the slope of the head toward the nucleus, the more preferred the head. That is, the first member of the head should have a greater consonantal strength than the next member, if there is one, and the nucleus should be weaker still. For example, given that /t/ is stronger than /d/, a head containing $\text{\$/tr}$ is more preferred than one with $\text{\$/dr}$ since the slope is greater when the voiceless plosive is the initial member of the head rather than its voiced counterpart.

Classical Greek had syllable heads consisting of a nasal plus a liquid. As the drop from a nasal to a liquid is very slight, this type of complex head is not preferred. To improve this condition either syllable head strengthening or consonant epenthesis occurred.

- (12) $\text{\$/mro-tós} > \text{brótós}$ (strengthening) 'mortal man'
 cf. $\text{\$/á-mrotós} > \text{\$/ám\$ro\$tós} > \text{ám\$pro\$tós}$ (epenthesis) 'immortal'

2.1.3 The Coda Law

- (13) Coda Law

A syllable coda is the more preferred: (a) the smaller the number of speech sounds in the coda, (b) the less the Consonantal Strength of its offset, and (c) the more sharply the Consonantal Strength drops from the offset toward the Consonantal Strength of the preceding syllable nucleus.⁹

Whereas Head Law (a) stated that the preferred number of speech sounds in the head is one, the Coda Law (a) states that zero speech sounds is preferred in the coda. Elimination of members of the coda takes place in Icelandic word-medially and word-finally.

- (14) hest+ur hest+bak hest+s 'horse'
 $\text{/hes\$tYr/}$ $\text{/hes\$bak/}$ $\text{/hes\$s/}$

⁹ I have tentatively placed the affricates with the plosives on the Consonantal Strength scale.

¹⁰ While there might appear to be a conflict between Coda law (b) and (c), the conflict is only apparent. The Coda Law states that preferentially, there should be no coda, but if there is one, it should be stronger than the preceding nucleus.

Because a weak offset is preferred, codas tend to weaken rather than strengthen as was the case for syllable heads. Coda weakening is evident in some Spanish dialects.

(15) Sp. salag > dial. salah¹¹ 'halls'

Coda Law (c) is the inverse to Head Law (c). That is, in a complex coda the strongest element should be the final speech sound of the group and the slope should be greatest from the final segment of the cluster toward the preceding speech sound toward the nucleus. This means that a complex coda of *rtʃ* would be preferable to *rdʃ* or *ltʃ*, etc. (Vennemann 1988:27).¹²

2.1.4 The Syllable Contact Law

The Syllable Contact Law (hereafter SCL) is the most crucial Preference Law with regards to this paper.

(16) Syllable Contact Law¹³

A syllable contact A\$B is the more preferred, the less the Consonantal Strength of the offset A and the greater the consonantal Strength of the onset B; more precisely-the greater the characteristic difference CS(B)-CS(A) between the Consonantal Strength of B and that of A.

Recall that in 7 the phonetic correlates were given a numerical value. A syllable contact of *rʃt* would yield an equation of $6 - 2 = 4$, where CS(B) = t = 6 and CS(A) = r = 2. If the contact were *tʃr*, the result would be -4, from $2 - 6$. The larger the difference, the better the contact; therefore, the contact of *rʃt* is preferred over *tʃr* since it yields a higher output. The German example in 17 illustrates this.

(17) Wartha /varʃta/, but Tatra /taʃtra/

In the first example, *Wartha*, there is a good syllable contact of *rʃt*. The contact in the second example, *Tatra* would be bad if it was syllabified as *tʃr* as we have just seen. However, it is tautosyllabic which is quite acceptable since it satisfies Head Law (c) (Vennemann 1988:41).

Now that we have reviewed the Preference Laws as set out by Vennemann, we can continue with the problem at hand; namely, Proto-Romance syllable structure with evidence from the histories of four Romance languages.

3.0 Italian

Murray (1987) (following Foley 1977:32) presented a Consonantal Strength scale for Romance based on the sound changes which have occurred diachronically in the Romance languages. In this scale labials are considered stronger than velars and dentals. Foley (1977) finds evidence for this from the fact that intervocalic dentals and velars are deleted in Spanish and French

¹¹ *h* patterns as a glide in terms of its consonantal strength.

¹² Of the liquids *r* is weaker than *l* as we will see shortly.

¹³ Vennemann (1988) refers to this law merely as the Contact Law. I chose the title which appears most commonly in the literature.

as shown in 18 and 19.

| | | | | | |
|------|-----|---------|-----|--------------|--------------|
| (18) | Lt. | rēgālem | Sp. | real | 'royal' |
| | Lt. | crēdo | Sp. | creo | 'I believe' |
| but | Lt. | habere | Sp. | haber [aβer] | 'to have' |
| | | | | | |
| (19) | Lt. | legere | Fr. | lire | 'to read' |
| | Lt. | amicam | Fr. | amie | 'friend' |
| | Lt. | credere | Fr. | croire | 'to believe' |
| | Lt. | vitam | Fr. | vie | 'life' |
| but | Lt. | habere | Fr. | avoir | 'to have' |
| | Lt. | ripam | Fr. | rive | 'riverbank' |

Among the liquids *l* is taken to be stronger than *r*. Evidence for this comes from Romanian where intervocalic *l* weakens to *r*, and also Spanish word-final *l* and Andalusian syllable-final *l* both weaken to *r*.

| | | | | | |
|------|------|----------|------|-------|---------|
| (20) | Lt. | solem | Rom. | soare | 'sun' |
| | Sp. | local(e) | > | lugar | 'place' |
| | And. | alma | > | arma | 'soul' |

The Consonantal Strength scale for Romance as presented by Murray (1987:119) is given below.¹⁴

| | | | | | | | | | | | | | | |
|------|----|--------|--|---|----|--------|--|-----------------|----|--------|--|---|---|---|
| (21) | a) | weak | | r | b) | weak | | velars, dentals | c) | weak | | j | u | 1 |
| | | strong | | l | | strong | | labials | | | | r | | 2 |
| | | | | | | | | | | | | l | | 3 |
| | | | | | | | | | | | | N | | 4 |
| | | | | | | | | | | | | v | | 5 |
| | | | | | | | | | | | | d | g | 6 |
| | | | | | | | | | | | | b | f | 7 |
| | | | | | | | | | | | | t | k | 8 |
| | | | | | | | | | | strong | | p | | 9 |

¹⁴ Murray has assumed that *b* and *f* are of equal strength according to the scale in 6 shown earlier.

3.1 Gemination¹⁵

Modern Standard Italian is well known for its phonemically long consonants. These long consonants or geminates are usually represented orthographically in Italian, for example, *fatto* 'fact' versus *faio* 'fate'. In the first word the sequence of <tt> is long but in the second word, the <t> is short. These long consonants are not always represented orthographically, however, as we can see in the word *figlia* [fiʎʎa] 'daughter'. If we can explain the historical motivation for the development of these long consonants in Italian as a syllable contact improvement, we will have partial evidence for the syllable structure of Proto-Romance. That is, later phonological developments in Italian should help furnish proof for the earlier syllable structure of Proto-Romance.

We can begin by first looking at gemination of consonants before *j*.

| | | | | | |
|------|-----|-----------|-----|-------------------|-----------------------|
| (22) | Lt. | sapiat | It. | sappia | 'he knows (subj.)' |
| | Lt. | rubiam | It. | robbia | 'madder' (plant name) |
| | Lt. | fugiat | It. | fuggia | 'he flees (subj.)' |
| | Lt. | vindemiam | It. | vendemmia | 'grape harvest' |
| | Lt. | filiam | It. | figlia [fiʎʎa] | 'daughter' |
| | Lt. | +pūgium | It. | puzzo [puttʰo] | 'stench' |
| | Lt. | brachium | It. | braccio [brattʰo] | 'arm' |
| | Lt. | radium | It. | raggio | 'ray' |

but

| | | | |
|-----|-------|------------------------|------------------|
| Lt. | aream | Pied. aira (via +arja) | 'threshing area' |
|-----|-------|------------------------|------------------|

Both the voiced and voiceless series of Latin plosives as well as *l* geminated before *j* but *r* did not. Gemination of these plosives also occurred before the glide *ɥ*, though less regularly for *l* and again not at all for *r*.

¹⁵ The historical development of gemination in Italian is not equal to the synchronic gemination of Italian, that is, diachronic gemination in Italian cannot be equated to *Raddoppiamento sintattico* (Nespor and Vogel 1986:165-168 and 170-175; Saltarelli (1983) refers to this process as *Rafforzamento*). N & V describe the type of gemination which takes place across word boundaries as a resyllabification process within the phonological phrase which occurs when a short stressed vowel word-finally is followed by a word-initial consonant. Through resyllabification the word-initial consonant also becomes the coda of the preceding syllable. This cannot explain the historical evolution of geminates in Italian however, since in words such as *figlio* 'son', Lt. *filius*, which have a long vowel preceding the word-internal consonant, there would be no explanation for the resyllabification according to N & V's analysis. Saltarelli (21) tries to explain diachronic gemination through his theory of synchronic gemination which is when the coda of a branching rhyme is empty, it reassociates to the coda of the following onset creating ambisyllabic segments (19). Again, this analysis does not work since he sees diachronic gemination as an ameliorative process subsequent to consonant deletion. That is, once the coda of a syllable is deleted, leaving an unassociated *C* along the CV tier, the coda reassociates to the right creating a geminate consonant. This cannot be generalized to all types of historical gemination however, since gemination did not only take place as a result of consonant deletion. Also, the example he uses, Lt. *rupta*, It. *rotta* is traditionally considered an assimilation process (Pei 1954:58), not a coda deletion process.

| | | | |
|------|------------------------|-----------|------------|
| (23) | Lt. sapuŋ | It. seppi | 'I knew' |
| | Lt. haβuŋ | It. ebbi | 'I had' |
| | Lt. poβuŋ | It. potti | 'I could' |
| | Lt. ⁺ caβuŋ | It. caddi | 'I fell' |
| | Lt. voβuŋ | It. volli | 'I wanted' |

but

| | | |
|------------|------------|---------------|
| Lt. doβuit | OIt. doβve | 'he suffered' |
| Lt. paβuit | It. paβve | 'it seemed' |

This gemination can be explained using the Preference Laws of Syllable Structure and assuming heterosyllabification of the consonant and the following glide, that is $VC\$_jV$ and $VC\$_\mu V$. The SCL given in 16 is required to explain the change from $VC\$_jV$ to $VC\$_CjV$.

The SCL states that a head with a higher consonantal value and a weaker coda in the preceding syllable is more preferred than the opposite situation. By referring to the strength scale presented in 21 and by using the equation in 16 we can see that the contact evaluation for a syllable structure of $Vr\$_jV$ is $1 - 8 = -7$, where 1 is the consonantal strength value of j and 8 is the value of r . In 22 and 23 Latin shows a poor syllable contact in that the weaker element, the glide, is in the head position and the stronger element is in the coda which is the opposite of what is most preferred according to the SCL. To rectify this, gemination occurred. That is, the consonant of the coda duplicated so that the following head would be filled by a stronger identical consonant instead of a glide thereby resulting in a more preferred syllable contact. The Proto-Romance syllable contact of $S\$_jW$ (where S = strong and W = weak) became the more preferred $S\$_S$ in modern Italian.

If syllable contact is actually the motivation behind gemination, why then do we see gemination of l but not r when r is in the coda and the following syllable head is a glide? Gemination of r does not occur because of the two liquids, r is the weaker one. The syllable contact of $r\$_j$ is more tolerable than $l\$_j$ since r is only slightly stronger than the following glide, therefore there is less motivation to geminate.¹⁶ This tendency for the worst syllable contacts to change before or instead of the better syllable contacts is explained in a corollary principle from Murray (1987:121):

- (24) The probability of a syllable structure undergoing a syllable structure improvement process increases as its evaluation decreases.

Since the contact evaluation for $l\$_j$ is -2 and -1 for $r\$_j$, syllable contact change is most likely to occur in the structure with the lower evaluation.

3.1.1 Slope Steepening

Slope steepening is a process by which the second element of syllable head weakens in order to make the slope from the initial element in the cluster towards the preceding syllable nucleus steeper. That is, the consonantal strength decreases from the initial segment of a

¹⁶ Compare to West Germanic gemination in section 7.1 below.

consonantal cluster towards the following nucleus. For example, a sequence of $\$pl$ becoming $\$pr$ or $\$pj$ would be an example of slope steepening. The Latin l became j in Italian when it was preceded by a tautosyllabic plosive.

- | | | | | | |
|------|-----|-----------------|-----|-----------------|--------------------|
| (25) | Lt. | p <u>l</u> anum | It. | p <u>j</u> ano | 'floor' |
| | Lt. | p <u>l</u> ango | It. | p <u>j</u> ango | 'I weep; I grieve' |
| | Lt. | c <u>l</u> amo | It. | ch <u>j</u> amo | 'I call' |

Slope steepening can also be seen word internally:

- | | | | | | |
|------|-----|---------------------|-----|---------------------|---------------------|
| (26) | Lt. | inc <u>l</u> udere | It. | inch <u>j</u> udere | 'to shut in' |
| | Lt. | sarc(u) <u>l</u> um | It. | sarch <u>j</u> io | 'hoe' |
| | Lt. | templ <u>l</u> um | It. | temp <u>j</u> io | 'temple' |
| | Lt. | exempl <u>l</u> um | It. | scemp <u>j</u> io | 'havoc; ruin' |
| | Lt. | confl <u>l</u> at | It. | gonf <u>j</u> ia | 'swollen; inflated' |

(27) Syllable Initial Margin Law

The preference for a syllabic structure $\$AB$, where a and b are the consonantal strength values of A and B respectively, increases with the value of a minus b .

Similar to Head Law (c) in 8, the Syllable Initial Margin Law states that a complex syllable head should consist of a stronger initial speech sound followed by a weaker second speech sound followed by an even weaker nucleus. The less a syllable head fulfills this requirement, the more likely it is for an improvement to occur. A syllable head of $\$kj$ (evaluation 7, where $k = 8$ and $j = 1$, therefore $8 - 1 = 7$) is more preferred than $\$kr$ (evaluation 6), and $\$kr$ is more preferred than $\$kl$ (evaluation 5) (Murray 1987:120); thus a syllable structure improvement will occur in the complex head of $\$kl$ prior to any improvement in either $\$kr$ or $\$kj$. This is exactly what we see occurring in 25 and 26. That is, l weakened to a glide after a syllable-initial plosive, thus improving the slope of the head. We must return again briefly to gemination.

Earlier I examined word-internal gemination before the glides j and $ɥ$. Gemination of plosives also occurred word-medially before l in Italian.

- | | | | | | |
|------|-----|---------------------|-----|---|--------------|
| (28) | Lt. | dup <u>l</u> um | It. | dopp <u>l</u> io | 'double' |
| | Lt. | oc <u>l</u> (u)lum | It. | oc <u>l</u> h <u>l</u> io | 'eye' |
| | Lt. | s <u>l</u> ub(u)lum | It. | sub <u>l</u> l <u>l</u> ia | 'chisel' |
| but | Lt. | t <u>l</u> eg(u)lum | It. | te <u>l</u> glia [te <u>l</u> l <u>l</u> a] | 'baking pan' |
| | Lt. | ad <u>l</u> o | It. | al <u>l</u> o | 'to the' |

In the first three examples in 28 gemination occurred with p , k and b , while d and g in the last two examples instead assimilated to the following consonant. This is directly in accordance with the Diachronic Maxim. Since p , k and b are all stronger than d and g on the Consonantal Strength scale,

improvement starts with the worst structures.

To reiterate, the purpose of this paper is to provide evidence for the Proto-Romance syllable structure, specifically whether an intervocalic consonant cluster should be divided as $VC\$CV$ or $V\$CCV$. The different developments that have been discussed here, gemination and slope steepening, show us that the division must have been $VC\$CV$.¹⁷

Since the weakening of l to a glide only occurred after a syllable-initial plosive we are led to believe that the correct syllable structure in 28 is $V\$CIV$. However, if we take as our starting point the heterosyllabic structure of $p\$l$, $k\$l$ and $b\$l$ we get a clearer picture of the processes that were involved. First, a poor contact of $VC\$IV$ ¹⁸ was ameliorated by gemination, which yielded $VC\$CIV$. An example of this intermediate stage can be seen in the example from dialectal Abruzzese *subbla* 'chisel' (Rohlf's 1966:348). This provided the environment for l to become j . If we begin by assuming that the plosive and the liquid are tautosyllabic we can explain the slope steepening but not the gemination. If on the other hand we start with a heterosyllabic structure both the gemination and the slope steepening are explained satisfactorily.

Weakening of r after a plosive did not occur in Italian, for example, Lt. *primum*, It. *primo* 'first' (Murray:125), because a plosive + r constitutes a fairly good syllable head and it is only the worst ones that will undergo any improvement. Word-internal slope steepening also did not occur:

| | | | | | |
|------|-----|----------|-----|---------|---------|
| (29) | Lt. | suprā | It. | sopra | 'above' |
| | Lt. | petram | It. | pietra | 'stone' |
| | Lt. | lacrimam | It. | lacrima | 'tear' |

but

| | | | | |
|-----|-----------|-----|---------|----------|
| Lt. | cōp(u)lam | It. | coppia | 'couple' |
| Lt. | vet(u)lum | It. | vecchio | 'old' |
| Lt. | oc(u)lum | It. | occhio | 'eye' |

The fact that l weakened after a plosive (after gemination took place) but r did not suggests that there was a differential syllabification in Italian; both $T\$l$ ($>T\$Tl$) and $\$Tr$ co-occurred. That is, Proto-Romance $VT\$rV >$ Italian $V\$TrV$. It might seem peculiar that both sequences of a plosive plus yod and a plosive plus a lateral liquid underwent gemination while a plosive plus r underwent resyllabification. Why would a sequence with a segment of intermediate consonantal strength (r) undergo a different process? The answer lies in the chronology of gemination and resyllabification. At an early stage in the development of Italian the very worst syllable contact, a plosive plus yod, underwent gemination in order to improve the contact. Following this the next worst contact, a plosive plus r , sustained a different but equally effective procedure, resyllabification. Finally, any remaining contacts that were deemed undesirable were once again improved upon.¹⁹ Thus the contact between a plosive plus a lateral, the least undesirable syllable contact, underwent a second stage of gemination.²⁰

¹⁷ The fact that p , k and b were heterosyllabic when preceding l implies that g and d also were heterosyllabic in this environment given that the first three segments have higher consonantal strength values than d and g , which implies that if p , k and b were still heterosyllabic, d and g must have been as well given that d and g , which make better syllable contacts with $\$l$, would not undergo any syllable contact improvements until after the worse contacts, $p\$l$, $k\$l$ and $b\$l$, had changed.

¹⁸ C denotes any of the plosives p , k or b here.

¹⁹ This will be elaborated on later.

²⁰ The change from $t\$l > t\tl did not occur in Italian as $\$tl$ is an unacceptable syllable head in any Romance language. As can be seen in It. *vecchio*, t first became k and then underwent gemination.

The fact that the two stages of gemination are interrupted by an alternate process is not unusual. Newton (1972) calls this type of rule ordering 'interdigitation'. In his paper the author is referring to two stages of obstruent deletion which are separated by a stage of glide formation in French. He states that 'what is intuitively a single phonological process is split into two parts by a second rule, so that we may symbolize the situation as A¹BA²' (41). Thus the two stages of gemination suggested for Italian are plausible.

3.2 Unnatural Syllabifications

Pensado (1989) explains the 'unnatural' plosive plus glide syllabifications in Vulgar Latin, (VC\$CV) by means of juncture and metrics in Classical Latin. That is, she sees the Romance languages as a continuation of Vulgar Latin which is itself a direct descendant of Classical Latin. I have already stated that I disagree with this supposed evolution of the Romance languages. Although I agree with her syllabification of VC\$GV, I do not agree with her argumentation for this syllabification; namely, that the 'unnatural' syllable structure of VC\$CV (in terms of the SCL) in Vulgar Latin is due to a resyllabification of V\$CCV in Classical Latin (Pensado:133).

Pensado (132) concurs with Nyman's (1982) theory of lexicalization which states that lento speech constraints acted upon unnatural syllable structures in Classical Latin containing glides which evolved through allegro speech rules. That is, syllable structure constraints of lento speech created heterosyllabic sequences of a consonant plus glide even though these types of heterosyllabic sequences had previously only been morphologically determined in Classical Latin. For example, resyllabification did not occur across morpheme boundaries in a sequence of a final vowel plus initial *muta cum liquida* even though this would create a better syllable structure, e.g., CL *āb legione* would be syllabified *b\$l* and not *\$bl*; that is, the plosive and liquid were syllabified morphologically despite the fact that a less preferred syllable structure would be the result.

However, Pensado disagrees with Nyman who concludes that this type of resyllabification, i.e., \$CG > C\$G, is unnatural and characteristic of poetic language in Classical Latin. Pensado (133) instead would claim that this resyllabification was more general and in fact was the impetus for allowing glides to occur word- and syllable-initially in Vulgar Latin where they had only occurred syllable-initially in Classical Latin and not morpheme-initially. Resyllabification according to Pensado relaxed the restriction against the distribution of glides. Thus Pensado (133) would argue that 'lexicalization contradicting phonological naturalness is in fact possible' and that while there is a tendency for natural rules to generalize, 'there seems also to be an opposite tendency to preserve or even to generalize phonologically unnatural domain restrictions' (134).

On the other hand, if we begin with VC\$CV in Old Latin as I have suggested and continue this sequence into Proto-Romance we do not have to account for the unnatural reversal of the syllable boundary in Vulgar Latin. While Pensado must explain how V\$CCV became VC\$CV in Vulgar Latin, which she herself admits is an aberrant event, we merely have to explain how Proto-Romance VC\$CV became V\$CCV in Classical Latin, which we can do by referring to a natural process of resyllabification due to poor syllable contact. Therefore, starting with VC\$CV in the proto-language simplifies the argument and does not require referral to unnatural phonological processes.

3.3 Differential Developments in Italian

Salverda de Grave (1930) (hereafter S de G) examines syllable structure in Latin and how different types of syllable structures explain the distinct developments of seemingly identical sequences of speech sounds. He does not accept Grandgent's (1927) claim that these developments are due to dialect influences or that they arise through analogy. Instead, S de G claims that these diverse developments are due to differences in syllabification (323), though he does not explain how these different syllable divisions arise.

S de G (323) posits three types of syllabifications for intervocalic consonant clusters:

- (30) I. The cluster stays together and
 a. the consonant cluster is tautosyllabic, as in It. *orecchio* from Lt. *auri\$cla* 'ear'.
 b. the consonant cluster forms one single sound and is 'intervocalic' e.g. It. *origlia* also from Lt. *auricla*.
 II. The cluster is heterosyllabic, for example, the *ng* in *pian\$go* 'I cry'.

Of course in this paper I assume that the clusters of which S de G speaks are all originally heterosyllabic.

S de G's argument is weak on a number of points. First, in Ia of 30 he argues that in *orecchio* the *k* and *l* must have been tautosyllabic in order to account for the slope steepening ($l > j$), yet he does not explain how the geminate arose in this form. Second, in Ib of 30 he claims the plosive and liquid are a single sound. He offers no description of how these two segments are syllabified, he merely states that they are intervocalic. This gives us no indication as to why *origlia*, with the geminate palatal lateral, should arise from *auricla*.

The heterosyllabic theory I have presented so far, however, can explain both the development of *origlia* and *orecchio* if we accept that the diversity in their developments can be attributed to dialect differences.²¹

| | | | |
|------------------------|----------------|-------------------|------------------|
| (31) a. <i>auricla</i> | | b. <i>auricla</i> | |
| k\$l | | k\$l | |
| g\$l | Voicing | k\$kl | Gemination |
| j\$l | Coda Weakening | k\$kj | Slope Steepening |
| j\$ʎ | Palatalization | | |
| ʎ\$ʎ | Assimilation | <u>orecchio</u> | |
| <u>origlia</u> | | | |

As 31 shows there were two different developments based on the same initial syllable structure. In the example on the left, *origlia*, the plosive first became voiced then weakened in the coda position since it was followed by a weaker syllable head. Then the liquid following the yod palatalized. Finally, the yod assimilated to the palatal lateral. In the second development, *orecchio*, the regular development of *k\$l* proceeded, that is, gemination followed by slope

²¹ Rohlfs (1966:350) also attributes this difference to dialectal variation.

steepening. The reason there were two developments could be attributed to dialect variation (much like Grandgent theorized) with the possibility that minimal pairs existed, *auriculum*, where syncope occurred after the voicing of the stop, and *auriclum*, in which there was no voicing and the development proceeded as described above (cf. Pei 1954:63).

3.3.1 Consonant + glide

I would like to show some of the different developments of consonant cluster sequences referred to in 30 and give S de G's explanation for their development and then my own.

| | | | | | | | |
|------|----|-----|-----|--------|-----|----------|------------------------|
| (32) | pj | Ia. | It. | sappia | Lt. | sa\$pjam | 'I know (pres. subj.)' |
| | | II | It. | saccia | Lt. | sap\$jam | 'I know (pres. subj.)' |

S de G explains the dialectal differences in 32 in the following manner: He maintains that in II in 32 the syllable-initial yod in the heterosyllabic sequence underwent glide strengthening, a common process in Italian, for example Lt. *majorem* > It. *maggiore* [dd^z] 'greater; elder'. Instead of the usual *d^z* though we get *t^s* in *saccia* since the yod follows a voiceless consonant. As for the *p* before the glide in II, it is in a weak syllable-final position and so is lost according to S de G. However, he does not explain why gemination, i.e., <cc> has occurred. As for Ia in 32, with the tautosyllabic *-pj-*, S de G states that *p* and *j* are so different in articulation that the *p* is maintained (324). S de G gives the same types of reasoning for differential developments involving *bj*, *mj*, *vj* and others.²² However, there is another theory which can more coherently explain the diverse developments.

We must first assume, as I have throughout this paper, that the sequence of intervocalic consonant + glide is heterosyllabic. Both of the results shown in 32 can be derived from this kind of structure, as represented in 33.

| | | | | | |
|---------|----------------------------------|----------------------|----|---------------|------------|
| (33) a. | sapiam | | b. | sapiam | |
| | p\$ji | | | p\$ji | |
| | p\$di ^z | Glide Strengthening | | p\$pi | Gemination |
| | p\$ti ^s | Voicing Assimilation | | <u>sappia</u> | |
| | t ^s \$ti ^s | Assimilation | | | |
| | <u>saccia</u> | | | | |

The difference in the developments may once again be due to dialect variation where it is possible that the etymon of *saccia* maintained its trisyllabic state longer than that of *sappia*. The developments in 33 also explain why there is a voicing distinction between *maggiore* and *saccia*, since in the latter form there was a voicing assimilation, whereas in the former there was only glide strengthening with subsequent assimilation. Since both developments can be explained with one type of syllabification, there is no need to posit another as S de G has done, for which he has no explanation.

²² See 324-327 for further examples and similar explanations from S de G on intervocalic consonant + glide developments.

3.3.2 Muta cum liquida

Salverda de Grave argues that a consonant plus a liquid were never heterosyllabified word-internally (327). That this is not the case I think has been demonstrated earlier in this paper, at least for *l*. In 3.1.1, the section on slope steepening, I said that there was a difference in Italian in the syllabification of word-medial *l* and *r* after they followed a plosive, namely $T\$l$ and $\$Tr$. However, $C\$r$ must have been present at some point in Proto-Romance. How else are we to explain Italian words with gemination before *r* such as *febbro* from Lt. *fabrum* 'blacksmith' (S de G:327)?

As I explained in 3.1.1, resyllabification did take place after an earlier stage of gemination transpired. If we start with a structure such as $T\$l$, the worst syllable contact, we know a bad syllable contact exists. One way to correct this is gemination as we saw earlier. Another way to improve a bad contact is resyllabification. In Italian resyllabification of the word-internal plosive did not occur when preceding *l* but did occur before *r* because a syllable-initial plosive plus *l* is less preferred than a syllable-initial plosive plus *r* according to the Syllable Initial Margin Law. Therefore, after the initial stage of gemination of a plosive preceding yod, resyllabification of the word-internal heterosyllabic plosive occurred when preceding *r*. When resyllabification occurred with *r* the first segments to resyllabify would be the voiceless plosives according to the Diachronic Maxim. After this stage of resyllabification another phase of gemination took place. In this second stage of gemination all the remaining syllable contacts considered undesirable were eliminated starting with the worst sequence $p\$l$ and continuing down the Consonantal Strength scale to $b\$r$. This sequence of events can explain the cases of gemination and resyllabification in Italian without making reference to differential syllabifications which S de G relies on. A summary of the changes I just described is shown below with the changes listed in the order they occurred.

Table 1

| Proto-Romance | Vp\$ <i>l</i> | Vb\$ <i>l</i> | Vp\$ <i>r</i> V | Vb\$ <i>r</i> V | Vp\$ <i>l</i> | Vb\$ <i>l</i> | |
|---------------|------------------------|------------------------|-------------------------|--------------------------|------------------------|------------------------|-------------------|
| | Vp\$ <i>p</i> <i>l</i> | Vb\$ <i>b</i> <i>l</i> | ----- | ----- | ----- | ----- | Gemination 1 |
| | ----- | ----- | V\$ <i>p</i> <i>r</i> V | ----- | ----- | ----- | Resyllabification |
| | ----- | ----- | ----- | Vb\$ <i>b</i> <i>r</i> V | Vp\$ <i>p</i> <i>l</i> | Vb\$ <i>b</i> <i>l</i> | Gemination 2 |
| | ----- | ----- | ----- | ----- | Vp\$ <i>p</i> <i>l</i> | Vb\$ <i>b</i> <i>l</i> | Slope Steepening |
| Italian | Vp\$ <i>p</i> <i>l</i> | Vb\$ <i>b</i> <i>l</i> | V\$ <i>p</i> <i>r</i> V | Vb\$ <i>b</i> <i>r</i> V | Vp\$ <i>p</i> <i>l</i> | Vb\$ <i>b</i> <i>l</i> | |

In this section I examined the processes of gemination and slope-steepening in Italian and how they show that a word-internal consonant cluster must have been divided as VC \$CV. The development of 'unnatural' syllable structures in Vulgar Latin as presented by Pensado (1989) was briefly reviewed and it was concluded that her results were at least inconclusive. I also reviewed work from Salverda de Grave (1930) to show that only one type of syllabification is necessary to explain all of the phonological developments in Italian. I hope to find corroborating evidence in my further examination of other Romance languages.

4.0 Portuguese

Evidence from another Romance language, Portuguese, also alludes to an original VC\$CV sequence. I would like to extend the principles referred to in the section on Italian to the Portuguese data below in order to show that the phonological processes of metathesis, coda weakening, and glide strengthening present in the historical development of Standard Portuguese support the theory that Proto-Romance did indeed have the syllable structure VC\$CV.

4.1 Metathesis²³

In Italian recall the most common means of ameliorating a poor syllable contact was gemination. In Portuguese, however, the most typical remedy was metathesis, as shown in 34.

| | | | | | |
|------|----|-------------|-----|-----------------------------|----------------------|
| (34) | VL | sēpiam | Pg. | siba ²⁴ | 'cuttlefish' |
| | VL | sapiat | Pg. | saiba ²⁵ | 'know (pres. subj.)' |
| | VL | pūtēum | Pg. | poço ²⁶ | 'well' |
| | VL | hōdie | Pg. | hoje ²⁷ | 'today' |
| | VL | +rābiam | Pg. | raiva | 'rabies, anger' |
| | VL | rūbēum | Pg. | ruivo | 'ruddy, reddish' |
| | CL | cavēam | Pg. | gaiva | ? |
| | VL | bāsiūm | Pg. | beijo ²⁸ | 'kiss' |
| | VL | bassiāre | Pg. | baixar | 'to lower' |
| | VL | querimōniam | Pg. | caramunha | ? |
| | VL | vindēmiam | Pg. | vindima | 'grape harvest' |
| | VL | pecūliūm | Pg. | pegulho ²⁹ | ? |
| | CL | januāriūm | Pg. | janeiro ³⁰ | 'January' |
| | CL | diaria | Pg. | geira | 'a yoke of land' |
| | VL | cōriūm | Pg. | coiro | 'leather, hide' |
| | VL | contrāriūm | Pg. | contrairo (old and popular) | 'contrary' |

²³ Data for the historical development of Portuguese comes from Williams (1962) which I will be abbreviating as Wms.

²⁴ VL tonic *e* + *j* > *i* or *ei*.

²⁵ The form *aipo* 'celery' cf. CL *āpiūm* is slightly irregular in that we expect intervocalic voicing of the *p* as is the case in *saiba*. It is possible *aipo* derived from ⁺*appiu* according to Williams (79) in which case we could posit degemination (which was prevalent in Portuguese, e.g. Lt. *pannum*, Pg. *pano* 'cloth'). This degemination would create another sequence of *p\$j* which would result in metathesis and produce *j\$p*. If the intervocalic voicing preceded degemination then we would expect to have two different developments, *p* > *b* and *pp* > *p*. This seems to be the case judging by the development of geminates in Portuguese, e.g. CL *abbātem*, Pg. *abade* 'abbot', CL *cippum*, Pg. *cepo* 'log; stump', CL *būccam*, Pg. *boca* 'mouth'.

²⁶ VL tonic *o* + *j* > *o*, *u*, *oi*, or *ui*.

²⁷ VL tonic *o* + *j* > *o* or *oi*.

²⁸ Also with palatalization of the *s*.

²⁹ VL tonic *u* + *j* > *u*.

³⁰ VL tonic *a* + *j* > *ai* or *ei*.

In *siba*, for example, the derivation would have been the following: *sēp\$jam* > *seb\$jam* (voicing could have occurred at this point or at some point later in the derivation) > *sej\$ba* (metathesis) > *siba*.

| | | |
|------|-----------------|----------------------|
| (35) | <i>sēp\$iam</i> | |
| | <i>p\$j</i> | |
| | <i>b\$j</i> | Intervocalic Voicing |
| | <i>ɟ\$b</i> | Metathesis |
| | <i>si\$ba</i> | Vowel Coalescence |
| | <hr/> | |
| | <i>siba</i> | |

Metathesis also occurred to a lesser extent with *y*.

| | | | | | |
|------|----|----------------|-----|---------------------------------|-------------------------------|
| (36) | VL | <i>sapŭit</i> | Pg. | <i>soube</i> ³¹ | 'I/he knew (pret.)' |
| | VL | <i>pōtŭit</i> | Pg. | <i>pōde</i> ³² | 'he/she could (pret. indic.)' |
| | VL | <i>placŭit</i> | Pg. | <i>prougue</i> | ? |
| | VL | <i>habŭit</i> | Pg. | <i>houve</i> | 'I/he had (pret.)' |
| | VL | <i>ēquam</i> | | <i>équa</i> > dial. <i>euga</i> | 'mare' |
| | VL | <i>pōstŭit</i> | Pg. | <i>pōs</i> | 'he/she put (pret. indic.)' |

The development of *houve* would be *hab\$ŭit* > *hav\$ŭit* (frication of the *b*) > *hau\$ve* > *houve* [ove].

As shown in 34 all of the consonants which underwent metathesis when followed by a heterosyllabic yod have a higher consonantal strength value than the following yod (refer to the Consonantal strength scale 21). Therefore, it is predicted by the Syllable Contact Law that a contact of a strong coda followed by a weaker head will be improved. By reversing the positions of the coda and the head, we now have a weaker coda than head which is more preferred.

In some examples in 34 the process of metathesis is concealed due to palatalization, for example, Portuguese *poço*, *hoje*, *beijo* and *baixar*. However, we know that metathesis did indeed occur due to the change in the vowel originally preceding the consonant + yod sequence. By investigating the development of certain vowels when followed by yod, we can see that metathesis must have taken place in order to explain how the vowel and the yod came into contact.³³

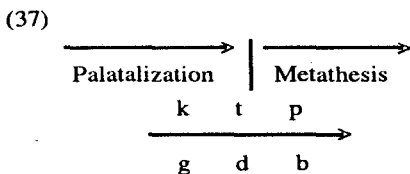
The fact that *k* and *g* appear not to have undergone metathesis when followed by a yod, as indicated by the lack of change in the preceding vowel, for example Lt. *facie*, Pg. *face* 'face', Lt. *pulēgium*, Pg. *poejo* 'pennyroyal mint', suggests that two parameters were at work, palatalization, an assimilatory process, on the one hand, and metathesis, a syllable contact improvement process, on the other. If palatalization was to apply prior to metathesis, it would in effect bleed the environment for metathesis in some instances. Thus those segments completely palatalized prior to metathesis would not undergo metathesis. It appears from the data that the velars were first in line for palatalization. This is not unexpected given that velars often undergo

³¹ VL tonic *a* + *y* > *ou* [o].

³² VL tonic *o* + *y* > *o*.

³³ The footnotes beside the examples in 43 and 44 show the developments of the vowels in question.

palatalization before the labials and dentals (Foley 1977:94). It seems that the two parameters were in competition with each other. At an early stage palatalization was productive and affected the velars first. Prior to metathesis the velars had undergone the full process of palatalization. The dentals on the other hand seem to have undergone only partial palatalization before the second parameter, metathesis, took effect. Thus those segments which had already undergone complete palatalization, *k* and *g*, were prevented from also undergoing metathesis, but those sounds which had undergone only partial palatalization were still in the environment, *t* and *d*, for metathesis to occur. This is represented below.



What 37 shows is that at the point at which metathesis becomes productive, the velars had sustained the entire process of palatalization but the dentals had only partially undergone palatalization. Thus the dentals as well as the labials were in the environment for metathesis.

| | | | |
|------|---|---------|----------|
| (38) | Stage 1: Intervocalic voicing | fač\$je | bās\$jum |
| | Stage 2: Complete Palatalization of the velars: | fa\$ce | baž\$jum |
| | Stage 3: Partial Palatalization of the dentals. | faž\$jo | baž\$jo |
| | Stage 4: Metathesis of the dentals and labials. | baž\$zo | baž\$zo |

face beijo

What does all of this mean for our reconstruction? It provides further support for an earlier syllabification of *VC\$CV* and not *V\$CCV*. While it appears in the examples above with the velars before the yod that the plosives could have been tautosyllabic with the glide, it is more likely that heterosyllabification was the true structure if we accept that the other plosives were also heterosyllabic. If the consonant and the glide in those instances of partial palatalization had been tautosyllabic there would have been no justification for the metathesis since a complex head containing a consonant and a glide is quite acceptable. However, a syllable structure of a strong coda followed by a weaker head is not acceptable, as we have seen, therefore it needs to be improved. If the dental and labial plosives were heterosyllabic, the velars must have been as well

given that a language system will not contain a less preferred structure (*p\$ʃ*) without also containing a more preferred structure (*k\$ʃ*). This was the case in Proto-Romance and metathesis in Portuguese supports this conclusion.

4.1.1 Coda Weakening

In Italian a word-internal sequence of an intervocalic plosive plus a heterosyllabic, lateral liquid first geminated and then slope steepening occurred. This did not occur in Portuguese however. Instead, the plosive in the coda weakened when preceding the *l*. Recall that coda weakening is another means of improving the syllable contact.

| | | | | | |
|------|----|-------------|-----|---|-------------|
| (39) | VL | ap̄ic(ʊ)lam | Pg. | abelha ³⁴ | 'honey bee' |
| | VL | ɔc(ʊ)lum | Pg. | ôlho | 'eye' |
| | VL | rɔt(ʊ)lum | | +roclam > Pg. rôlha ³⁵ | 'cork' |
| | VL | tæg(ʊ)lam | Pg. | telha | 'tile' |
| | VL | fab(t)lære | | fablar > fallar > Pg. falar ³⁶ | 'to speak' |

| | | |
|------|--------------|----------------------|
| (40) | ap̄ic(ʊ)lam | |
| | apic\$lam | Vowel Deletion |
| | abic\$lam | Intervocalic Voicing |
| | abij\$lam | Coda Weakening |
| | abi\$ʎa | Palatalization |
| | abe\$ʎa | Misc. Vowel Changes |
| | <hr/> abelha | |

In 39 the plosives preceding *l* all underwent weakening. The *b* in *fabülære* assimilated to the following *l*. The development of intervocalic consonants differed from those presented above.

³⁴ VL *k* > *j* when preceded by a vowel and followed by *l*. This yod had the effect of palatalizing the lateral. When the *k* was preceded by another consonant, *cl* instead became *ch*, e.g. VL *masculum* > Pg. *macho* 'male'

³⁵ *tl* > *kl* in Vulgar Latin.

³⁶ Words where *bl* > *br*, such as VL *oblīgäre* > *obrigar* 'to force', are borrowings from Spanish or are semi-learned. Thus the apparent tautosyllabification of *bl* (as evidenced by the slope steepening) is not the regular syllabification but a later development. Williams (78) does not accept previous arguments to explain the development of *bl* in *falar*, arguments such as *falar* developed under the influence of *calar* or that it is due to the dissimilation of *v* of the intermediate form ⁺*favlar*. According to Williams these arguments do not explain forms such as *taleira* 'bracket' (< ⁺*tabularia*), or *ullo* (< ⁺*ublo* < *übīllum*) and *solo* 'solo' (< ⁺*sublo* < *süb illum*). If however, we assume that the plosive and the lateral were heterosyllabic and that the plosive underwent assimilation and then deletion, these words are also explained.

| | | | | | |
|------|----|---------|-----|------------------------|-----------|
| (41) | CL | lŭpum | Pg. | lŏbo | 'wolf' |
| | CL | amīcum | Pg. | amigo | 'friend' |
| | CL | natam | Pg. | nada [ð] ³⁷ | 'nothing' |
| | CL | legŭmen | Pg. | legume | 'bean' |
| | CL | habĕre | Pg. | haver | 'to have' |

We can also compare the changes in 39 to those of an intervocalic plosive plus *r*.

| | | | | | |
|------|----|----------|-----|------------------------------------|-----------------|
| (42) | VL | apŕilem | Pg. | abril | 'April' |
| | VL | pĕtram | Pg. | pedra | 'stone' |
| | VL | lacŕimam | Pg. | lágřima | 'tear' |
| | VL | fĕbřem | Pg. | fevre ³⁸ (old and pop.) | 'fever' |
| but | VL | intĕgrum | Pg. | inteiro ³⁹ | 'whole, entire' |

Unless we take as our starting point *VC\$CV* there is no way of explaining this weakening. If the plosives in 39 were indeed tautosyllabic with the following *l* we would expect developments similar to those in 41 and 42; that is, intervocalic voicing or frication, but not coda weakening. This leads to the next point of differential syllabification, which we briefly discussed in the section on Italian.

Once again it appears that when a plosive precedes *r* it is tautosyllabic rather than heterosyllabic. When a plosive precedes a liquid there is a preference for that plosive to be in either the coda or head position, depending on the strength of the plosive and the strength of the liquid. This preference is based on the Syllable Contact Law and the Syllable Initial Margin Law.

| (43) | Contact | Evaluation | Initial Margin | Evaluation ⁴⁰ |
|------|---------|------------|----------------|--------------------------|
| | p\$ř | -7 | \$gl | 3 |
| | p\$l | -6 | \$gr | 4 |
| | k\$ř | -6 | \$dl | 3 |
| | k\$l | -5 | \$dr | 4 ⁴¹ |
| | t\$ř | -6 | \$bl | 4 |
| | t\$l | -5 | \$br | 5 |

³⁷ *t > d > ð*.

³⁸ The form *frágua* 'furnace' (< ⁺*frauga* < ⁺*fravga* < *fravega* < Lt. *fabŕicam*) shows not only the intervocalic frication of *b > v* but also shows metathesis of *r*. What it also shows is the metathesis of the *u*, that is, ⁺*frauga* > *frágua*. Given that the form ⁺*frauga* seems to represent an ideal syllable contact, a weak coda followed by a strong onset, there appears to be no justification for this metathesis. The answer could lie in the fact that this word is a borrowing from Spanish and thus does not represent the regular development of this sequence in Portuguese.

³⁹ Where *gr* remains *gr* or where *cl* and *gl* become *gr*, these words are semi-learned or borrowings according to Williams (77).

⁴⁰ Based on Murray (1987:128).

⁴¹ Even though *d* and *g* have the same consonantal strength, *g* as an onset is less preferred given that velars are inherently weaker than dentals (Foley 1977:33).

| | | | |
|--------------|----|--------------|---|
| b\$ <i>r</i> | -5 | \$ <i>l</i> | 5 |
| b\$ <i>l</i> | -4 | \$ <i>tr</i> | 6 |
| d\$ <i>r</i> | -4 | \$ <i>kl</i> | 5 |
| d\$ <i>l</i> | -3 | \$ <i>kr</i> | 6 |
| g\$ <i>r</i> | -4 | \$ <i>pl</i> | 6 |
| g\$ <i>l</i> | -3 | \$ <i>pr</i> | 7 |

Both the contact and initial margin columns show the least preferred sequences starting from the top and going downward toward the most preferred sequences. What this table shows is that *r* and a preceding plosive are preferentially tautosyllabic while *l* plus a preceding plosive are preferentially heterosyllabic. That is, when one compares a sequence of a plosive plus *r* to a sequence of a plosive plus *l*, tautosyllabic plosive plus *r* is preferred to tautosyllabic plosive plus *l* and heterosyllabic plosive plus *l* is preferred to heterosyllabic plosive plus *r*.

This table also shows that of all the plosives, *g* makes the worst syllable head. Thus it is not entirely unexpected that it remains in the coda position in *intēgrum* while the other plosives appear in the head position. That is, if the voiced velar plosive was to resyllabify and become tautosyllabic with the following *r*, a relatively poor syllable head would be the result. The *g* being in the coda position would explain why we get coda weakening as opposed to frication in *inteiro*.

The Portuguese example *dobrar* 'to duplicate', Lt. *duplāre* appears to be an exception to those forms in 39 since it seems to exhibit word-internal slope steepening which would indicate tautosyllabification instead of heterosyllabification. However, if we compare this form to Pg. *doble* 'double', Lt. *dūplus*, we see a different pattern. In the second example, *doble*, Proto-Romance stress existed on the syllable preceding the voiceless plosive. As stressed syllables tend to attract segments, this syllable was most likely closed by the plosive, e.g. *VC\$CV*. Since *dobrar* originally had stress following the plosive and liquid combination, the likely syllable structure was *V\$CCV*. Thus the syllabification and development of these two words was dependent upon stress as shown in the derivation below.

| | | | | | |
|------|----|---------------------|----|-------------------|---------------------|
| (44) | a. | dūp\$ <i>lus</i> | b. | du\$ <i>plāre</i> | |
| | | ----- | | du\$ <i>prar</i> | Slope Steepening |
| | | dub\$ <i>lu</i> | | du\$ <i>brar</i> | Voicing |
| | | dob\$ <i>le</i> | | do\$ <i>brar</i> | Misc. Vowel Changes |
| | | <hr/> | | | |
| | | doble ⁴² | | dobrar | |

4.1.2 Glide Strengthening

Yet another means of improving a sequence of *VC\$CV* where the coda is stronger than the head is head strengthening. In this case, we are talking about glide strengthening since it is the glides which appear in the head position. The example in 45 shows the strengthening of *u* to *v* when it occurs as a head.

⁴² The regular development of *bl* > *ll* > *l*, as in *falar*, would have occurred at a stage prior to voicing, therefore we do not get the development **dub\$le* > **dul\$le* > **dole* in Portuguese.

| | | | | |
|------|-----|------------------------------|-----|--------------------------------|
| (45) | Lt. | val <u>u</u> isset | Pg. | val <u>u</u> esse ‘?’ |
| | Lt. | +d <u>o</u> l <u>u</u> erunt | Pg. | d <u>o</u> l <u>u</u> erun ‘?’ |
| | Lt. | man <u>u</u> ale | Pg. | man <u>u</u> al ‘flail’ |
| | Lt. | +min <u>u</u> are | Pg. | min <u>u</u> ar ‘to decrease’ |

Because the consonantal strength of the glide is weaker than the preceding coda, the glide’s consonantal strength increases by becoming *v* in the first two examples and *gu* in the last two.

I have again tried to show in this section how phonological processes in Portuguese, such as metathesis, coda weakening, and glide strengthening, can give us an indication as to earlier syllable structure. It seems apparent, judging by the evidence put forth, that there is strong reason to believe that Proto-Romance had heterosyllabic intervocalic consonant clusters based on the data from Portuguese and Italian. In the next section I will present data from Romanian which should provide further proof for this hypothesis.

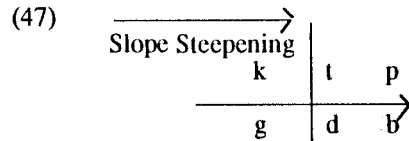
5.0 Romanian

5.1 Slope Steepening

| | | | | | |
|------|-----|-----------|-------|-------------|-------------|
| (46) | Lt. | glanda | Rom. | ghindă | ‘acorn’ |
| | Lt. | clave | Rom. | cheie | ‘key’ |
| | Lt. | plāga | Rom. | plagă | ‘wound’ |
| | Lt. | blasphēmo | NRom. | blestema | ‘to curse’ |
| | Lt. | flōre- | Rom. | înfiori(re) | ‘to flower’ |

Word-initially only the velar plosives + *l* underwent slope steepening. Why should this be so? Recall 43 in the Portuguese section where $\$pl$ made a better complex head than either $\$gl$ or $\$kl$. This was because *p* had a greater consonantal strength than either *g* or *k* ($p = 9$, $k = 8$, $g = 6$) and so is more tolerable as the initial segment of a complex head. That is, the slope of $\$pl$ is 6, $\$kl$ is 5 and $\$gl$ is 3. Therefore, the structures with the worst values, here $\$kl$ and $\$gl$, will be altered prior to more preferred structures, here $\$pl$, according to the Diachronic Maxim.

However, two segments which lie between *k* and *g* on the consonantal strength scale, *b* and *f*, do not undergo this slope steepening as expected. Perhaps this can be explained by the fact that velars are typically weak segments and have a tendency to weaken prior to labials.⁴³ Thus we get slope steepening first with the velars -*kl*- and -*gl*-, then if the phonological process is still productive, it will carry on with the dentals then labials. As 47 shows, slope steepening stopped after the velars.



Word-internally slope steepening also took place, again only with the velars.

⁴³ This fact is represented in 21 of the Italian section.

| | | | |
|------|---------------|---------------------------|------------|
| (48) | Lt. oricla | Rom. ureche | 'ear' |
| | Lt. vigilăre | Rom. veghea | 'to watch' |
| | Lt. vetulu- | Rom. vechiu ⁴⁴ | 'old' |
| but | | | |
| and | Lt. duplus | Rom. duplu | 'double' |
| | Lt. stab(u)lu | Rom. staul | 'stable' |
| | Lt. sub(u)lu | Rom. sulă | 'awl' |
| | Rom. povidlă | colloq. povirlă | ? |

Word-internally we see a variety of different developments. In the first two examples in 48 with *-kl-* and *-gl-* the plosive remains while the *l* is weakened to *j*, then deleted. With *-pl-* both the plosive and the lateral remain. Finally, with *-bl-* the plosive is deleted while the *l* remains. One possible solution to these differential developments is resyllabification and coda weakening.

The Romanian forms *ureche* and *veghea* seem to suggest tautosyllabification since the liquid after the plosive weakens to yod, thereby palatalizing the preceding velar, and then ultimately disappearing. However, if we begin with a syllable structure of *VC\$CV* (this would involve prior vowel syncope for some of the forms in 48) followed by resyllabification, we end up with a structure of *V\$CCV*. Resyllabification would start with the least preferred structures and work its way through the system. In Romanian resyllabification generalized to the voiceless plosives.

As for the remaining plosives, *b* and *d*, different developments occurred. Earlier in the Portuguese section 43 showed that *\$kl* and *\$pl* are more preferred syllable heads than are *\$dl*⁴⁵ and *\$bl*. It also showed that *d\$l* and *b\$l* make better contacts than *k\$l* and *p\$l*. It is probable then that these voiced plosives remained in the coda position rather than becoming heads through resyllabification. This placed the voiced plosives in a position of weakening and this is what happened to *b* and *d*. The syllable-final *b* weakened to a glide and colloquially *d* weakens to an *r*.

Finally, the remaining undesirable syllable contact of *g\$l* was again altered through resyllabification. Thus the last stage in our derivation, as shown below, is slope steepening which happened only with the velar plosives.

| | | | | | |
|------|-----------------------|---------|----------------------|----------|----------------------|
| (49) | | dup\$lu | orik\$la | stab\$lu | vig\$lăre |
| | Resyllabification I: | du\$plu | ori\$kla | ----- | ----- |
| | Coda Weakening: | ----- | ----- | stau\$lu | ----- |
| | Resyllabification II: | ----- | ----- | ----- | vi\$glar |
| | Slope Steepening: | ----- | ori\$kja | ----- | vi\$gia |
| | | duplu | oreche ⁴⁶ | staul | veghea ⁴⁷ |

⁴⁴ *-il-* > *-kl-*.

⁴⁵ Of course this type of syllable head is not possible in Romance, therefore coda weakening was a viable alternative.

⁴⁶ The palatalization of [k] and [g] does not appear to be an act of progressive assimilation; that is, these two plosives were palatalized by the following yod (< *l*) and not the preceding front vowel. This is shown by the fact that examples such as Rom. *mic* 'small' (Lt. *micu*), Rom. *frig* 'cold' (Lt. *frigu*) do not show palatalization (also see Nandris: 146 & 150).

⁴⁷ While palatalization of *l* > *j* does take place when *l* is syllable-initial, e.g., Lt. *linu*, Rom. *in* 'flax', Lt. *lepore*, Rom. *iepure* 'rabbit, hare' (see Nandris: 140), this change only occurs when *l* is followed by *i*; therefore, it appears that this is a case of assimilation and not a change based on consonantal strength.

A comparison of Italian and Romanian syllabification shows that in Italian $VT\$rV > V\TrV and $VT\$lV$ remained, that is, there was no resyllabification. In Romanian though resyllabification generalized with the voiceless plosives and liquids, that is, $VT\$LV > V\TLV .

5.1.1 Affrication

Another bit of evidence which indicates heterosyllabification, or $VC\$CV$, is the development of $-dj-$ and $-tj-$. These sequences had two different but parallel developments, and each development was based on stress assignment and syllable structure. Word stress on the vowel preceding the plosive tended to close the stressed syllable; that is, $\check{V}d\$j$ and $\check{V}t\$j$ were the result. If however stress was on the vowel following d and t this tended to attract d and t to the head position of the stressed syllable; that is, $\$dj\check{V}$ and $\$tj\check{V}$ resulted (Nandris:128). The outcome for both tautosyllabification and heterosyllabification is affrication but the end products differ, as shown in 50 and 51.

- (50) $\check{V}d\$jV > d^z$ $\$dj\check{V} > \d^z
 $\check{V}t\$jV > t^s$ $\$tj\check{V} > \t^s
- (51) Lt. mēdiu Rom. miez ($d^z > z$) 'core; kernel'
 Lt. adjutare Rom. ajuta ($d^z > \check{z}$) 'to help'

(cf. Rom. jos, Lt. deo(r)sum 'low; down')

- Lt. negotiu Rom. negoț (t^s) 'trade'
 Lt. +fetiolu Rom. fecior (t^s) 'boy; son'

Thus it appears we have direct evidence for an earlier $VC\$CV$ syllabification, in the case of stress on the first syllable.

5.1.2 Metathesis

The voiced labials b and m , as well as f , underwent another method of syllable contact improvement.⁴⁸ When these elements appeared in the coda position and were followed by yod, the yod metathesized thereby producing a more preferred sequence of $W\$S$ rather than $S\$W$.

- (52) Lt. scabia ORom. zgaibă ?
 Lt. diffamiat Rom. defăima 'defame'/?
 Lt. cofea Rom. coif 'helmet'

There are many processes which conspire to create a more preferred syllable structure. The processes reviewed above, slope steepening, coda weakening, metathesis and varying phonological developments based on different syllabifications, can all be explained when we assume an earlier syllabification of $VC\$CV$.

⁴⁸ The only Romanian word available with a sequence of $p + j$ available was *apropia(re)*, Lt. *appropiāre* 'to draw, bring near'. The lack of any change suggests that the i in this form may have remained syllabic, thus no change was required.

6.0 Catalan

6.1 Gemination

Catalan, like Italian, employed gemination as a means of improving a syllable contact of VC\$CV.⁴⁹

| | | | | | |
|------|-----|-------------|-------|-------------------------------|-----------|
| (53) | Lt. | diabòlus | Cat. | diable [diabblə] | ? |
| cf. | Lt. | nob(i)le | Rous. | [nobblə] ⁵⁰ | 'noble' |
| | Lt. | dūplus | Cat. | doble [dobblə] ⁵¹ | 'double' |
| | Lt. | dūplus | Rous. | [dubblə] | 'double' |
| | Lt. | regulum | Cat. | regla [rregglə] ⁵² | 'rule' |
| | Lt. | sēcāle | Cat. | segle [segglə] | 'century' |
| | | ? | Cat. | tecla [tekklə] ⁵³ | 'key' |
| cf. | Lt. | +joc(u)lare | Rous. | [žuggláj] | 'to joke' |

As can be seen in the examples in 53, when a strong plosive in the coda position meets a weaker liquid in the head position, gemination is the result. This gemination occurs in order to satisfy the Syllable Contact Law. After gemination a strong coda and an equally strong head come in contact, a situation preferable to the original.

Stress seems to play a factor in the syllabification and gemination process in Catalan. Huber (1929:11) mentions that when the vowel preceding the plosive is stressed, gemination is the result, but when stress follows the plosive, there is no gemination. The first situation, in which stress precedes the plosive, would result in a closed syllable as stressed syllables tend to attract segments to them. Thus we would have VC\$CV. When the stress follows the plosive, the result is a complex head in the stressed syllable; that is, V\$CCV. This would explain the varying results seen in *dóble* [-bbl-] and *doblar* [dubláj] 'to double' (Huber:11). In *doble* the stress precedes the word-internal plosive and gemination is the result. In *doblar*, where stress comes after the intervocalic plosive, there is no gemination (cf. section 4.1.1 on Portuguese and 5.1.1 on Romanian).⁵⁴

While the syllabification for an intervocalic plosive plus a lateral liquid was heterosyllabic, that for a plosive plus *r* appears to be tautosyllabic based on the developments of the plosive; namely intervocalic weakening.

⁴⁹ The change of *r* > *rr* syllable-initially, as in Lt. *rūbeus*, Cat. *roig* [rrot^s], is a means of onset strengthening and not syllable contact induced gemination.

⁵⁰ Rousillon is a dialect of Catalan.

⁵¹ Intervocalic voicing could have taken place prior to or after gemination. The result would be identical.

⁵² Cf. Cat. *reglar* [rægláj] 'to rule'.

⁵³ Cf. Cat. *teclat* [təklát] 'keyboard'.

⁵⁴ The non-geminate sequence in *doblar* remains controversial as the *Diccionari ortogràfic i de pronúncia* (1990) does have [dubbláj].

| | | | |
|------|--|------|---|
| (54) | Intervocalic weakening: VCV (where C represents any plosive) | | |
| Lt. | invltāre | Cat. | envidar [əmbid̪a] ⁵⁵ 'to invite' |
| Lt. | arripare | Cat. | arribar [əriβa] ⁵⁶ 'to arrive' |
| Lt. | fricant | Cat. | fregan [fr̪əγan] ⁵⁷ 'they rub' |
| Lt. | hibernu- | Cat. | hivern [iβɛrn] 'winter' |
| Lt. | nūda | Cat. | nua [nuə] ⁵⁸ 'naked (f.sg.)' |
| | VCrV | | |
| Lt. | petra | Cat. | pedra [peðrə] ⁵⁹ 'stone' |
| Lt. | supra | Cat. | sobre [soβrə] 'over, above' |
| Lt. | +acrus | Cat. | agre [aγrə] 'sour' |
| Lt. | febre | Cat. | febre [feβrə] 'fever' |
| Lt. | quadrāta | Cat. | cairat [kəjrat] 'beam' |

Comparing the examples in 54 with those in 53 we can see that the type of development was primarily based upon the syllable structure. When the plosive and liquid were heterosyllabic (53), gemination took place, however, when the plosive and liquid were tautosyllabic (54), intervocalic weakening occurred. Because the *r* is weaker than the *l*, and therefore makes a worse contact, heterosyllabification was less tolerated and resyllabification occurred. A complex head of \$Cr (where C represents any plosive) is more preferred to \$Cl according to the Head Law, thus original C\$l did not resyllabify. Instead, gemination took place which also resulted in a better syllable structure.

6.1.1 Coda Weakening

Besides gemination, coda weakening was also utilized in Catalan as a syllable contact improvement process. The SCL is satisfied by weakening a coda with a stronger consonantal strength than the following syllable head.

| | | | | | |
|------|-----|-----------|-------|------------------------------|----------------|
| (55) | Lt. | scop(u)lu | Cat. | escull [əskuɫ] ⁶⁰ | 'rock' |
| cf. | Lt. | scōpūlus | Rous. | [əskuɫ] | 'rock' |
| | Lt. | tabula | Cat. | taula [taɯlə] | 'table, board' |
| | Lt. | oculus | Cat. | ull [uɫ] | 'eye' |
| | Lt. | vetulus | Cat. | vell [veɫ] ⁶¹ | 'old' |
| | Lt. | coagulu | Cat. | coall [koəɫ] | 'clabber' |

⁵⁵ *t* > *d* > *ð*.

⁵⁶ *p* > *b* > *β*.

⁵⁷ *k* > *g* > *γ*.

⁵⁸ The transition was most likely *d* > *ð* > *∅*.

⁵⁹ There are also forms such as Cat. *mare* 'mother', Lt. *matre* and Cat. *pere* 'father', Lt. *patre*, in which *t* > *ð* > *∅* and earlier forms with *-pr-* > *-br-* > *ɸ*, e.g. Lt. *cupru*, +*cubru* > Cat. *coure* ?[kouɾə] 'copper'. As well the Catalan word *cairat* also indicates coda weakening. These forms could indicate an earlier heterosyllabic sequence which would explain the coda weakening.

⁶⁰ *-pl-* > *-kl-*.

⁶¹ Recall that *-tl-* also became *-kl-*.

In the sequences *-kl-* (from original *-kl-*, *-pl-* and *-tl-*) and *-gl-* the stop weakened to a yod, palatalized the following *l*, as indicated by the [ʎ] in all but one of the examples above, and then deleted. In *taula* the plosive also weakened to a glide but this had no effect on the following consonant. The weakening of a plosive to a glide produced a more preferred sequence of *WSS* as opposed to the original *S\$W* sequence. The examples in 54 as compared to those in 55 once again suggest that a difference in the syllable structure can account for the disparate developments.

The reader may be wondering why the same segments, *p*, *b*, *t*, *k* and *g* in the same environments underwent different developments. It has been suggested (Fouché:159) that in learned words the process of gemination took place but in non-learned words coda weakening occurred instead. Regardless of which development took place, both gemination and coda weakening have the same objective: to satisfy the SCL.

6.1.2 Metathesis

There were several developments which took place when a consonant originally in the coda position was followed by a yod. One of these developments was metathesis.

| | | | | | |
|------|-----|----------|------|-----------------------------|------------|
| (56) | Lt. | basium | Cat. | bes [bes] ⁶² | 'kiss' |
| | Lt. | bassiāre | Cat. | baixar [bæʃa] ⁶³ | 'to lower' |
| | Lt. | corium | Cat. | cuir [kujr] | 'leather' |

Once the yod and plosive metathesized, due to the poor syllable contact, the yod coalesced with the preceding vowel.

6.1.3 Onset Strengthening

When certain plosives preceded the yod, onset strengthening took place with subsequent coda deletion.

| | | | | | |
|------|-----|-------|------|--------------------------|-----------|
| (57) | Lt. | rubeu | Cat. | roig [rut ^s] | 'red' |
| | Lt. | podiu | Cat. | [put ^s] | 'hillock' |

With the sequences of *d\$J* and *b\$J* there seems to have been a strengthening first of the yod, which can also be seen word-initially in Cat. *jove* [ʒoβə] 'young', Lt. *jūvēnis*. After the yod was strengthened, the stop in the coda position was deleted.

| | | | |
|------|-----------------------|-----------------------|-----------------------|
| (58) | | rub\$ju | pod\$ju |
| | Glide Strengthening: | rub\$d ^z u | pod\$d ^z u |
| | Coda Deletion: | ru\$d ^z u | po\$d ^z u |
| | Final Vowel Deletion: | rud ^z | pod ^z |
| | Final Devoicing: | rut ^s | pot ^s |
| | (Vowel Alteration): | rut ^s | put ^s |
| | | rut ^s | put ^s |

⁶² s > z; Catalan has final devoicing.

⁶³ Palatalization took place with *-ss-*.

The development of $p\$/j$ differs from the plosives above. In some cases it appears that this sequence may have resyllabified, that is, $p\$/j > \pj , e.g., Lt. *sepia*, Cat. *sepia* [sɛpiə] ‘cuttlefish’. Since p has the greatest consonantal strength of any of the plosives and yod is one of the weakest segments, it follows that $p\$/j$ will make the worst contact. However, if resyllabification occurs $\$/pj$ is formed, resulting in the most preferred complex head. As we have repeatedly seen, change starts with the worst structures; therefore, it is expected that if a sequence of a heterosyllabic consonant plus yod is going to resyllabify, resyllabification will start with the least preferred.

In some other words it seems that i does not lose its syllabicity and thus a form like Lt. *sapiat* remains trisyllabic, as in Cat. *sapiga* ‘that he know’.⁶⁴

This final branch of Romance shows many of the processes we have already seen in previous sections of this paper, metathesis, coda weakening, glide strengthening and gemination. In order to explain the developments systematically and economically, we need to begin with a syllable structure that can account for all of the changes. The evidence points to $VC\$/CV$.

7.0 Proto-Indo-European Syllable Structure⁶⁵

The data from the four Romance languages above presents evidence for Proto-Romance $VC\$/CV$. Evidence for this syllable structure also exists in Sanskrit, Proto-Germanic, Pre-Latin and Ancient Greek. This means that $VC\$/CV$ can likely be pushed back farther than just Proto-Romance and into PIE. This is a very strong claim, but one not without support.

7.1 Germanic

Murray and Vennemann (1983) argued for $VC\$/CV$ in Proto-Germanic with verification for this argument coming partly from West Germanic gemination. Examples of this are shown below.

| | | | | | | | |
|------|-----|-----------------|----|----------|----|----------|---------------|
| (59) | Go. | sat $\$/$ jan | OS | settian | OE | settan | ‘to set’ |
| | Go. | -skap $\$/$ jan | OS | skeppian | OE | scieppan | ‘to create’ |
| | Go. | kun $\$/$ jis | OS | kunnies | | | ‘race (gen.)’ |
| | Go. | hal $\$/$ ja | OS | hellia | | | ‘hell’ |
| | Go. | akrs | OS | akkar | | | ‘acre’ |
| | ON | ep $\$/$ le | OE | æppel | | | ‘apple’ |

The data shows that before yod, all consonants except r geminated. This is due to the fact that of all the consonants, r has one of the lowest consonantal strengths. Gemination started with the poorest contacts and worked its way toward better contacts. The worst contact $T\$/j$ has a contact evaluation of -6 whereas $r\$/j$ is only -1. Given $r\$/j$'s better evaluation, gemination did not generalize to this sequence. Similarly, it was only the poorest sequence of $T\$/l$ (contact evaluation -4) which geminated (cf. Italian 3.1). Once again gemination was utilized to improve a sequence of $VC\$/CV$ ($S\$/W$).

⁶⁴ The g in this word most likely arose through glide strengthening, e.g. $pi\$/a > pi\$/ja$ (with yod insertion) $> pi\$/ga$ (glide strengthening).

⁶⁵ This section is based on my paper entitled *Proto-Indo-European Syllable Structure*.

7.1.1 Sanskrit

In Sanskrit verse there was a light/heavy syllable distinction. In a sequence of -VC \bar{C} CV- the first consonant closed the first syllable, which means that the two consonants were heterosyllabic.

- (60) Verse construction⁶⁶
- | | |
|--------------|--------------|
| <u>Light</u> | <u>Heavy</u> |
| V\$CV | V\$VC |
| | VC\$CV |

Therefore, for the purposes of verse construction, there existed a heavy syllable type such as VC\$CV.

There were also accounts by ancient Indian grammarians that intervocalic consonant clusters tended to be heterosyllabic. Thus a word like *pitre* would be syllabified as *pit\$re* and not *pi\$tre*, and *mukta-* was syllabified as *muk\$ta* not *mu\$ktā* (Varma 1961).

Gemination also occurred in Sanskrit. Forms such as *addya* < *ad\$ya* 'today', *puttreṇa* < *put\$reṇa* 'son (instr.)' and *pretya* < *pret\$ya* 'post-mortual state' (Varma 1961:62-64; from Vennemann 1988:34) indicate that a poor syllable contact of a plosive plus a glide or a plosive plus a liquid was ameliorated through gemination. As in West Germanic, gemination in Sanskrit reflects a VC\$CV pattern.

7.1.2 Ancient Greek

Different Greek dialects show different syllabification patterns. In Homeric Greek we see the same type of syllabification evident in Sanskrit; that is -VC \bar{C} CV- was syllabified as VC\$CV (Murray 1988:213). However, in Attic Greek an intervocalic stop plus a liquid is tautosyllabic as is a voiceless stop plus a nasal. Otherwise you still have VC\$CV. This is represented below.

- (61) Ancient Greek Syllable Structure⁶⁷
- | | | |
|--------------|----------------|------------------|
| <u>Homer</u> | <u>Attic</u> | |
| VC\$CV | V\$TLV, V\$DLV | |
| | V\$TNV | otherwise VC\$CV |

A piece of phonological evidence from Greek that reflects a pre-Greek pattern of VC\$CV is ω versus o in comparative and superlative adjectives ending in $o\zeta$. In the comparative and superlative adjectives ending in $o\zeta$, the vowel of the syllable following a heavy syllable would be ω . If however, the preceding syllable was light, the vowel of the following syllable would be o (Murray 1988:215; based on Hermann 1923).

- (62)
- | | | | |
|------|--------------|-----------|---------------|
| T\$T | λεπτότατος | (λεπτός | 'thin, fine') |
| T\$r | μακρότατος | (μακρός | 'long, big') |
| D\$N | κεδνότατος | (κεδνός | 'careful') |
| D\$I | εκπαγλότατος | (εκπαγλός | 'terrible') |

⁶⁶ Murray (1988:211).

⁶⁷ Murray (1988:213).

In the first example in 62 then we have to assume that $-\pi\tau-$, or $-pt-$, is heterosyllabic given that the vowel following this cluster is o and not ω . This is true of all of the underlined clusters in 62.

7.1.3 Pre-Latin

We have already seen in the introduction to this paper that there was evidence for $VC\$CV$ in Pre-Latin. Recall that *intēgra* should have been *intigra* if the plosive and liquid were tautosyllabic, leaving the preceding vowel in an open syllable, as in *infiscit*. This indicates that the syllable structure was $VC\$CV$ at a point earlier than Classical Latin. Evidence for tautosyllabification exists in the Classical Latin words *tenēbrae* and *pāstris* since in Classical Latin a short vowel followed by a plosive and a liquid is considered to be in a light syllable.

7.1.4 Summary

Sanskrit, pre-Latin and pre-Greek all indicate a common syllable structure of $VC\$CV$. Changes of this sequence in Latin and Greek are due to poor syllable contact. In both Latin and Greek we get tautosyllabification, although it has generalized further in Greek. Sanskrit remains unchanged and is therefore considered archaic. These poor syllable contacts are reflected in the consonantal strength evaluations in 64. The scale below differs somewhat from the one we saw earlier.⁶⁸

(63) Consonantal Strength Scale
L N D T

1 2 3 3.5
→

| (64) Structure | Evaluation | Skt. | Lt. | Gk. |
|----------------|------------|------|-----|-----|
| T\$L | -2.5 | - | + | + |
| D\$L | -2 | - | + | + |
| T\$N | -1.5 | - | - | + |
| D\$N | -1 | - | - | - |

(‘+’ indicates resyllabification)

The plus and minus signs show which language underwent resyllabification with a particular heterosyllabic consonant cluster. The syllable structures beneath the word ‘Structure’ in the diagram above are those posited for PIE.

The original PIE structures remained heterosyllabic in Sanskrit while there was resyllabification in Latin and in Greek. What was considered tolerable in Sanskrit was not tolerable in Latin and even less tolerable in Greek which again shows us the language-specific variability in syllable structure tolerance.

Evidence from the various branches of Indo-European all reflect an earlier Proto-Indo-European syllable structure of $VC\$CV$. Gemination in West Germanic and Sanskrit can be accounted for by poor syllable contact improvement as can the resyllabifications in Latin and Greek. Vowel quality in Latin, as well as vowel quality in Greek, and verse construction and

⁶⁸ The scale in 63 and the evaluations in 64 are from Murray (1988:216-217).

ancient grammarian reports in Sanskrit support a reconstructed Proto-Indo-European syllable sequence of VC\$CV.

8.0 Conclusion

I have used various developments from a wide spectrum of Romance languages (Italian, Portuguese, Romanian and Catalan) to show that the ancestor of the Romance languages, Proto-Romance, must have begun with a syllable structure of VC\$CV and not V\$CCV, as is commonly believed. Due to the pressures applied to this original form, as outlined by the Preference Laws, certain changes were forced. Improvement of the syllable structure can explain the historical phonological processes which I have described above. These changes cannot be explained quite so neatly, if at all, starting with Proto-Romance V\$CCV. Thus the cumulative evidence of several Romance languages confirms the argument for VC\$CV in Proto-Romance. Evidence from other Indo-European languages suggests that the syllable structure of Proto-Romance posited here was inherited from Proto-Indo-European. A summary of the major changes in the Romance languages I discussed is presented in Table 2.⁶⁹

Table 2

| OL VC\$CV > PR VC\$CV > | Italian | Portuguese | Romanian | Catalan |
|----------------------------|-------------------------------------|--------------------------------------|----------|--|
| Vp\$jV | Vp\$piV | Vj\$pv | | |
| Vt\$jV | Vt\$tiV (with palatalization) | Vj\$tiV (with palatalization) | | |
| Vk\$kiV | Vk\$kiV | V\$C (complete palatalization) | | |
| Vb\$jV | Vb\$biV | Vj\$bV | Vj\$bV | Vb\$jV (with onset strengthening and coda deletion) |
| Vd\$jV | Vd\$diV (with palatalization) | Vj\$dV (with palatalization) | | Vd\$jV (with onset strengthening and coda deletion) |
| Vg\$giV | V\$C (with palatalization) | (complete palatalization) | | |
| Vp\$rv | V\$prV | V\$prV | | V\$prV |
| Vt\$rv | V\$trV | V\$trV | | V\$trV |

⁶⁹ Blank spaces in Table 2 indicate inconclusive evidence.

| | | | | |
|---------|---|--|---------------------------------------|--|
| Vk\$ rV | V\$krV | V\$krV | | V\$krV |
| Vb\$ rV | Vb\$brV | V\$brV | | V\$brV |
| Vd\$ rV | V\$drV | V\$drV | | Vd\$ rV (with coda weakening) |
| Vg\$ rV | V\$grV | Vg\$ rV (with coda weakening) | | |
| Vp\$ iV | Vp\$plV (with slope steepening) | Vp\$ iV/ V\$plV (with slope steepening) | V\$plV | Vp\$plV/ Vp\$ iV (with coda weakening) |
| Vt\$ iV | Vk\$kiV (with slope steepening) | Vk\$ iV (with coda weakening) | V\$kiV (with slope steepening) | Vt\$ iV (with coda weakening) |
| Vk\$ iV | Vk\$kiV (with slope steepening) | Vk\$ iV (with coda weakening) | V\$kiV (with slope steepening) | Vk\$kiV/ Vk\$ iV (with coda weakening) |
| Vb\$ iV | Vb\$b iV (with slope steepening) | Vb\$ iV (with assimilation) | Vb\$ iV (with coda weakening) | Vb\$b iV/ Vb\$ iV (with coda weakening) |
| Vd\$ iV | Vd\$ iV (with assimilation) | | Vd\$ iV (with coda weakening) | |
| Vg\$ iV | Vg\$ iV (coda weakening and assimilation) | Vg\$ iV (with coda weakening) | Vg\$ iV (with slope steepening) | Vg\$g iV |

REFERENCES

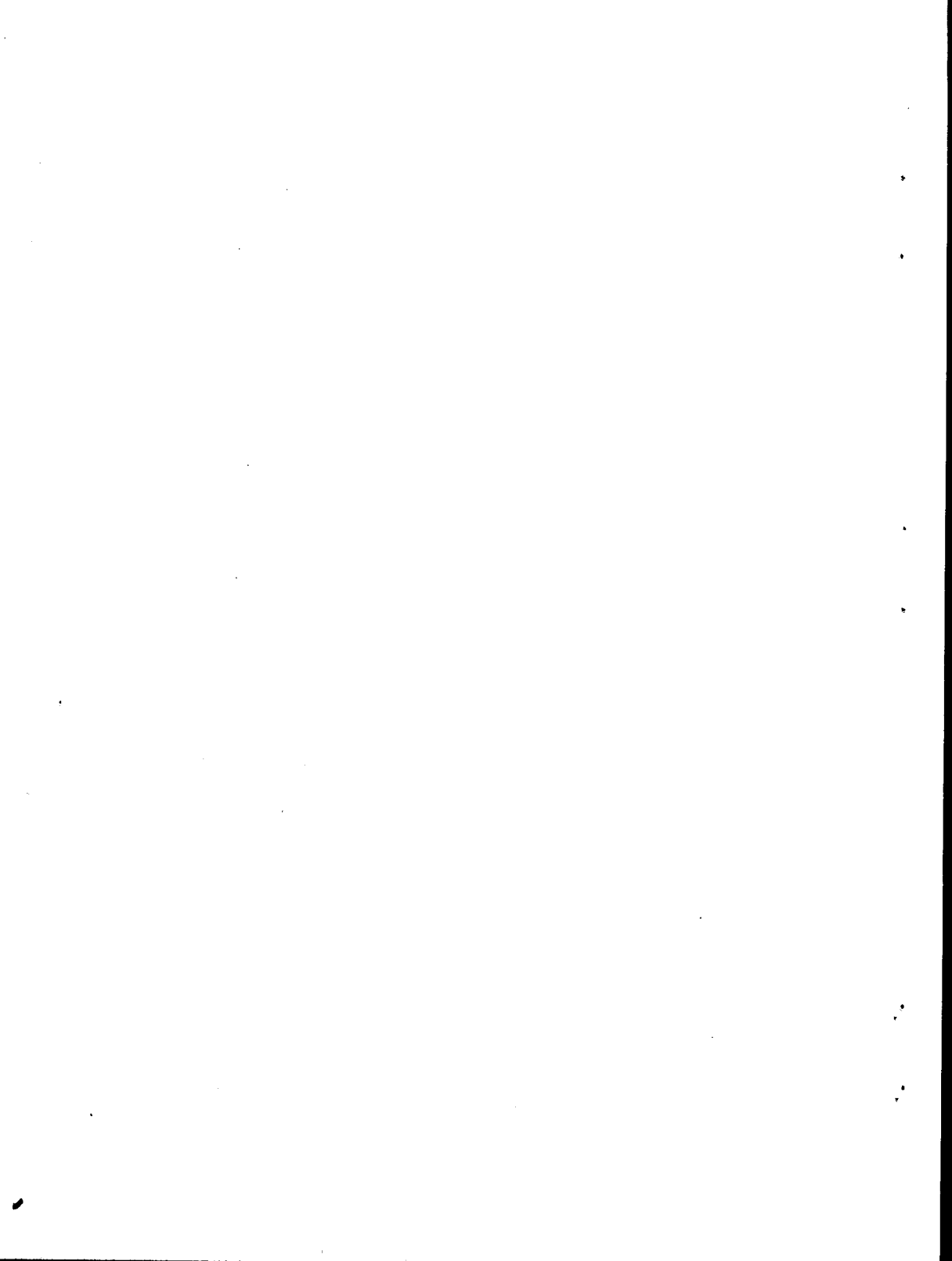
- Allen, W. Sidney. 1973. *Accent and Rhythm*. Cambridge: Cambridge University Press.
- Bruguera i Talleda, Jordi. 1990. *Diccionari ortogràfic i de pronúncia*. Barcelona: Enciclopèdia Catalana.
- Colomer, Jordi. 1973. *Diccionari anglès-català, català-anglès*. Barcelona: Llibre de Butxaca.
- Donegan, P. J. and D. Stampe. 1979. 'The syllable in phonological and prosodic structure'. *Syllables and Segments*, ed. by A. Bell and J.B. Hooper, 25-34. Amsterdam: North Holland.
- DuNay, André. 1977. *The Early History of the Rumanian Language*. Lake Bluff, Illinois: Jupiter Press.
- Elcock, W. D. 1960. *The Romance Languages*. London: Faber & Faber Limited.
- Foley, James. 1970. 'Phonological distinctive features'. *Folia Linguistica* 4. 87-92.
- _____. 1977. *Foundations of Theoretical Phonology*. Cambridge: Cambridge University Press.
- Fouché, Pierre. 1980. *Phonétique historique du roussillonnais*. Geneva: Slatkine.
- Grandgent, Charles H. 1927. *From Latin to Italian*. Cambridge: Harvard University Press.
- Hall, Robert A. Jr. 1950. 'The reconstruction of Proto-Romance'. *Language* 26. 6-27.
- _____. 1974. *External History of the Romance Languages*. New York: Elsevier.
- _____. 1976. *Proto-Romance Phonology*. New York: Elsevier.
- _____. 1983. *Proto-Romance Morphology*. Philadelphia: John Benjamins Publishing Co.
- Halle, Morris and Jean-Roger Vergnaud. 1987. *An Essay on Stress*. Cambridge: MIT Press.
- Hermann, Eduard. 1923. *Silbenbildung im griechischen und in den andern indogermanischen Sprachen*. Göttingen: Vandenhoeck and Ruprecht. 2nd ed. 1978.

- Hill, A.A. 1954. 'Juncture and syllable division in Latin'. *Language* 30. 439-447.
- Hoenigswald, H.M. 1949. 'A note on Latin prosody: initial S impure after a short vowel'. *Transactions of the American Philological Association* 80. 270-280.
- Hooper, Joan B. 1976. *An Introduction to Natural Generative Phonology*. New York: Academic Press.
- Huber, Joseph. 1929. *Katalanische Grammatik*. Heidelberg: Carl Winters Universitätsbuchhandlung.
- Lindsay, W.M. 1894. *The Latin Language*. Oxford: Clarendon Press.
- Meyer-Lübke, W. 1935. *Romanisches etymologisches Wörterbuch*. vol. 1 & 2. Heidelberg: Carl Winters Universitätsbuchhandlung.
- Muller, H.F. 1929. *A Chronology of Vulgar Latin*. Tübingen: Max Niemeyer Verlag.
- Murray, Robert W. 1987. 'Preference laws and gradient change: selected developments in Romance'. *Canadian Journal of Linguistics* 32(2). 115-132.
- _____. 1988. *Phonological Strength and Early Germanic Syllable Structure*. Munich: Wilhelm Fink Verlag.
- _____. 1991. 'Early Germanic syllable structure revisited'. *Diachronica* VIII:2. 201-238.
- Murray, Robert W. and Theo Vennemann. 1983. 'Sound change and syllable structure in Germanic phonology'. *Language* 59. 514-528.
- Nandris, Octave. 1963. *Phonétique historique du roumain*. Paris: Librairie C. Klincksieck.
- Nespor, Marina and Irene Vogel. 1986. *Prosodic Phonology*. Dordrecht: Foris Publications.
- Newton, Brian. 1972. 'Interdigitation in French phonology'. *Language Sciences*. February. 41-43.
- Nyman, M. 1982. 'Lexicalization out of casual speech: the Greek-Latin synizesis'. *Relational and Reconstructive Aspects of Grammatical Systematization. Data Oriented Studies*. Helsinki: University of Helsinki. 65-95.

- Panovf, Irina. *New Pocket Romanian Dictionary*. New York: Hippocrene Books, Inc.
- Pei, Mario A. 1954. *The Italian Language*. New York: S.F. Vanni Publishers.
- _____. 1976. *The Story of Latin and the Romance Languages*. New York: Harper & Row.
- Pensado, Carmen. 1989. 'How do unnatural syllabifications arise? The case of consonant + glide in Vulgar Latin'. *Folia Linguistica Historica* VIII 1-2. 115-142.
- Pulgram, Ernst. 1950. 'Spoken and written Latin'. *Language* 26. 458-466.
- _____. 1975. *Latin-Romance Phonology: Prosodics and Metrics*. Munich: Wilhelm Fink Verlag.
- Richardson, Elbert L., Maria de Lourdes Sá Pereira and Milton Sá Pereira. 1943. *McKay's Modern Portuguese-English and English-Portuguese Dictionary*. New York: David McKay Company, Inc.
- Rohlf, Gerhard. 1966. *Grammatica storica della lingua italiana e dei suoi dialetti*. Vol. 1. Turin.
- Saltarelli, Mario. 1983. 'The mora unit in Italian phonology'. *Folia Linguistica* 17. 7-24.
- Salverda de Grave, J.J. 1930. 'Sur l'évolution des consonnes en italien'. *Romania* 56. 321-330.
- Taylor, James L. 1970. *A Portuguese-English Dictionary*. Stanford: Stanford University Press.
- Varma, Siddheshwar. 1961. *Critical Studies in the Phonetic Observations of Indian Grammarians*. Delhi: Munshi Ram Monohar Lal. [Reprint of the 1929 edition, London, The Royal Asiatic Society.]
- Vennemann, Theo. 1972. 'On the theory of syllabic phonology'. *Linguistische Berichte* 18. 1-18.
- _____. 1988. *Preference Laws for Syllable Structure*. Berlin: Mouton de Gruyter.
- Williams, Edwin B. 1962. *From Latin to Portuguese*. Philadelphia: University of Pennsylvania Press.

APPENDIX

| | | | |
|-------|---|-------|---|
| [j] | voiced palatal central glide | <ş> | [ʃ] in Romanian |
| [ɥ] | voiced rounded labiovelar glide | <ț> | [tʃ] in Romanian |
| [β] | voiced bilabial fricative | <x> | [ʃ] in Portuguese |
| [tʃ] | voiceless alveolar affricate | C | consonantal sound except where otherwise stated |
| [dʒ] | voiced alveo-palatal affricate | CL | Classical Latin |
| [ɣ] | voiced velar fricative | D | voiced plosive |
| [ʎ] | palatal lateral approximant | G | glide |
| [ɲ] | voiced palatal nasal | L | liquid |
| [ð] | voiced interdental fricative | N | nasal |
| [tʃ] | voiceless alveo-palatal affricate | NRom. | North Romanian, also known as Daco-Roman |
| <ă> | [ə] (between /ə/ and /ɛ:/) in Romanian | OL | Old Latin |
| <î/â> | closer and tenser than /ə/ in Romanian | PR | Proto-Romance |
| <ô> | [o] in Portuguese | T | voiceless plosive except where otherwise stated |
| <c> | [s] in Portuguese | V | vocalic sound |
| <ç> | [ʃ] in Portuguese | VL | Vulgar Latin |
| <ch> | voiceless palatalized velar stop when followed by <i/e> in Romanian | ∨ | short vowel |
| <g> | [z] when followed by <i/e> in Portuguese | ∨ | long vowel |
| <gh> | voiced palatalized velar stop when followed by <i/e> in Romanian | + | reconstructed form when raised and preceding a word |
| <j> | [ʒ] in Portuguese | + | morpheme boundary |
| <lh> | [ʎ] in Portuguese | * | hypothetical or incorrect form |
| <nh> | [ɲ] in Portuguese | \$ | syllable division |
| | | > | becomes |
| | | < | derives from |



A Government and Binding Approach to Bella Coola

Michael Durrant
University of Calgary

INTRODUCTION

In the time leading up to the work presented in this paper I was struck by the relative lack of Government & Binding (GB) treatments for American Indian languages. This paper will remedy the situation somewhat by investigating the adequacy of the GB framework for a language which has hitherto been neglected by current versions of the theory. Bella Coola is an isolated member of the Salish family, located on the Canadian west coast.¹ It seems to have branched off before any other languages of the family, the two main branches being Coast Salish and Interior Salish. Bella Coola shows characteristics of both branches, but should not be considered any closer to Proto-Salish than any other language. The Salish family shares many characteristics with the neighbouring Wakashan and Chemakuan families, most notably VSO word order, reduplication, and a lack of clear distinction between verbs and nouns. The similarities between these language families have been attributed to a real diffusion, as genetic relations remain inconclusive.

In section 1, I will describe the salient elements of Bella Coola syntax, discussing the current options we have to account for them. Section 2 will put the language into a recent VSO typological classification (Woolford, 1991), while section 3 will discuss the general adequacy of the GB framework for Bella Coola.

1. SALIENT ASPECTS OF BELLA COOLA SYNTAX

1.1 Noun/Verb dichotomy

The most important peculiarity of the language is the (seeming) lack of distinction between nouns and verbs. Much of the debate in the Amerindian literature concerning the Salish noun/verb dichotomy supports the distinction between the two categories; it not only makes Universal Grammar more plausible, but shows the diversity of human languages that can be described working within UG. Researchers usually examine the ways the distinction manifests itself at the levels of Morphology, Syntax and Semantics. Phonology has proven useless in offering proof for the dichotomy in Bella Coola, while Morpho-syntax has been only nominally more successful. The semantic level judges whether the word is taken more as a referring expression, or as some sort of action or state of being; this is by far the weakest of the three levels and can sometimes be identified only in retrospect. The morphological level is based on the distribution of affixes to stems and is a strong source of data that may lead to a solution for the problem. In Bella Coola, morphology shows a distinction between three types of roots based on the kinds of personal inflection they may take (Nater, 1984):

* The research for this paper was supported by a Graduate Research Scholarship from the University of Calgary and by SSHRCC grant # 410-91-1956 to Eithne Guilfoyle.

¹ For transcription of Bella Coola data, the following phonemes are used: [c] voiceless palatal affricate; [t] voiceless lateral fricative; [ɬ] voiceless lateral affricate; ['] indicates that the preceding consonant will be pharyngealized.

- i) particles and clitics can take no personal marking
- ii) transitives can take bipersonal (subject/object) or passive pronoun paradigms
- iii) intransitives can take the subject/possessive paradigm

The intransitives are the most troublesome of the roots since they seem to be made up of everything from intransitive verbs to adjectives, nouns, independent pronouns, numbers, and interrogatives. All of these elements can be inflected with the paradigm in (1) below.

When this paradigm is attached to an intransitive verb, the pronoun indicates the subject (or agent) of the verb. If the paradigm is attached to a nominal form, it normally indicates possession. This requires some clarification, however; if the nominal occurs *without* determiners, it acts as the main predicate of the sentence, which implies that there is a verb present. And conversely, if a verbal form occurs *with* determiners, it attains a substantive function, usually of the form: gerund - possessive. The example in (2) shows the translation of a verbal form when inflected with the above paradigm, while (3) shows a nominal form with the same paradigm. The (b) examples show the meaning when there are determiners present.

- (2) a. ?atps-c (Newman, 1969a)
eat(intr)-1sg
'I am eating'
- b. ti-?atps-c
det-eat(intr)-1sg
'my eating'
- (3) a. staltmx-c
chief-1sg
'I am chief'
- b. ti-staltmx-c-tx
det-chief-1sg-det
'my chief'

There is thus no distinction made between nominal and verbal intransitive roots in the morphology; the distribution of this pronominal paradigm suggests that the intransitive roots are all the same category. By claiming that they are the same, there is a certain generality present in this language that cannot easily be explained within GB (or indeed, any theory which crucially refers to Noun and Verb as distinct categories). Does the form in (3a) have the same status that the form in (3b) would have without determiners, or is it specially generated to have nominal status when occurring with determiners? Could it be that there are nominal and verbal versions of each

intransitive root in the lexicon, one that occurs with determiners and the other without, or could it simply be a process of zero-derivation? These are not pleasing propositions since they attempt to take away some simplicity of the system.

What we have seen from the morphological level is that there appears to be a distinction between transitive and intransitive roots based on what types of affixes may be attached, but once derivational stems have attached, they can all possibly act the same.

In trying to reconcile these morphological facts within a greater theory of syntax, I was reminded of a certain conception of 'transformation' from an earlier version of generative grammar. In Bach (1968), it was suggested that all Noun, Verb, and Adjective Phrases were introduced in the form of relative clauses. This means that each type of phrase would have a similar underlying structure, (much as the intransitive roots in Bella Coola seem to). According to Bach's argument, the following two sentences are equivalent versions of the same sentence at Deep Structure and Surface Structure, respectively:

"I spoke to the one who was an anthropologist"

and

"I spoke to the anthropologist"

The first sentence describes the underlying form of the sentence - not only for English, but supposedly for all languages. If we adopt such a proposal for Bella Coola, the generality of the intransitives is very well explained; it is possible to translate the intransitives as nouns which are derived from underlying relative clauses. The only difference, then, between English and Bella Coola is that English has many transformations which change the underlying structure (sometimes obligatorily) to the surface realizations, while Bella Coola lacks these transformations.

This type of analysis was acceptable in the late sixties, when transformations were still in their formative stage, but this argument does not fit into current Government and Binding theory. To even suggest that English noun phrases are actually relative clauses at D-Structure would be unacceptable within current frameworks. If this argument is taken into the semantic area, however, we may be able to express some cross-linguistic phenomena. For now, it may be easiest just to take the intransitive roots as verbal nouns, but this will also cause problems (which are to be discussed in section 1.5).

It should be noted that there are, in fact, some roots which can act as both transitives and intransitives. These make up a relatively small class of roots which cannot be derivationally formed. When they are used with the transitive paradigm there is a sense of causative or agent-related activity, while, when they occur with the subject/possessive paradigm, they express more of a 'medio-passive notion' (Newman, 1969a). By suggesting that at least some roots have both transitive and intransitive forms in the lexicon, we are introduced to the problem of how to account for the transitivity and detransitivizing affixes that are very productive in the language. What effect do these affixes have on the argument structure of the verbs they are attached to? Are they base-generated with the verb, or are they actually contained within one of the verb's arguments and later incorporated into the verb? These argument-inhibiting and argument-adding affixes may even be stored, attached to their roots within the lexicon (as are the transitive/intransitive roots

discussed above); this would certainly save some trouble with changing the argument structure of the verbs, but would miss an important aspect of the language. This will be put on hold until some other things have been worked out.

1.2 Subject/Object agreement on transitive verbs

Transitive predicates in Bella Coola may occur with only their subject/object fused-form pronoun paradigm, or they may also have the lexical forms of the arguments, thus both (4a) and (b) below are completely normal in the language. The regular transitive subject/object paradigm is shown in (5) with the pronoun used in (4) in boldface.²

- (4) a. k'x-ic (Newman, 1969b)
 see(trans)-1sgS/3sgO
 "I see it/him/her"
- b. k'x-ic-c'n wa-sut-s ta-mnat-nu
 see-S/O-now det-house-3sg.poss det-son-2sg.poss
 "I see the house of your son"

The gaps in the paradigm indicate where reflexive items would occur with the reflexive morpheme, (which will not be covered in this paper). Newman (1969b) analyzed the above fused-form morphemes in order to break them down into their constituent parts. He found the system surprisingly regular, with the object forms preceding the subject forms in all but those with a second person object (where the order was reversed).³ The constituent pronouns in Newman's paradigm bear a close enough resemblance to their corresponding forms in other paradigms to say that they are pronouns.

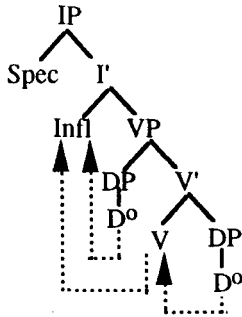
Now that the transitive paradigm has been shown to be related to other pronouns in the language, the status of these endings must be accounted for; are they the actual arguments of the verb (theta-marked in deep structure), or simply agreement markers for null arguments? The implications of this dilemma are far-reaching, for both the language and current GB theory.

If it is assumed that these suffixes are, in fact, the real arguments of the verb, they must occur in D-Structure, either as independent pronouns which make up the fused form, or as null elements which form a portmanteau morpheme that is realized only as the fused form in INFL. The first suggestion is unlikely, as the forms are not completely regular; if this were a productive process, they would show more regularity. The latter is apparently our only other option, suggesting that these forms are stored as independent morphemes that attach to the verb through a process such as incorporation when there is movement through a specifier position. The diagram below suggests one possible way that this can happen (this diagram only accounts for the 'regular' order of Verb-Object-Subject):

² There is also a causative transitive paradigm, which inflects a different class of roots. The causative roots are specified to take the causative transitive paradigm and have essentially the same meaning as regular transitive roots with the regular paradigm. The only time there is a causative meaning is when the causative paradigm inflects a non-causative root:

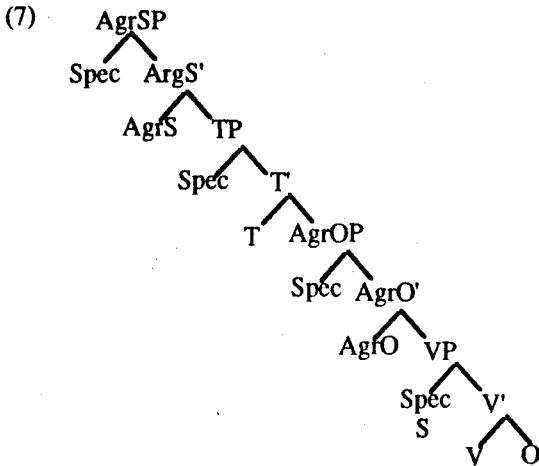
³ This may have something to do with the phenomenon of Agent Hierarchy as described by Jelinek and Demers (1983) for some Coast Salish languages, and also the thematic role hierarchy described by Randall (1988).

(6)



The only remaining problem with this proposal is the status of the lexical complements to the verb, as in (4b). If the null pronouns are the arguments of the verb, the lexical elements could not have thematic roles, and must then be classed as adjuncts. The problem with this is that adjuncts show freer word order than complements, and in this language the word order is preferably VSO when there are lexical subjects and objects. This suggests that the lexical items act more as complements than adjuncts, bringing into question the status of the pronominal suffixes as the arguments of the verb. There is also the consideration that the lexical elements intuitively come after the verb in the order: subject — object. These points indicate that it is unlikely that the pronouns are the (theta-marked) arguments of the verb, and rather that the lexical items are theta-marked (when present).

We will now investigate the hypothesis that transitive suffixes are only agreement markers. With this approach, the verb will get its marking for subject and object by being moved through the Agreement nodes on its way to INFL. The following diagram is a modification of Mahajan (1990) cited in Travis (1991) which indicates the structure of Agreement phrases within the (former) IP for a VSO language.



With this proposal, there are null arguments in the VP, but the agreement for these arguments is found within each AgrP. Some strong points in favour of this analysis include the fact that the lexical items are assigned theta roles, and their surface word order is identical to their underlying order.

In order to clarify the status of the transitive suffixes, the facts for and against accepting them as the verb's underlying arguments (as opposed to it's agreement markers) will now be summarized:

- the transitive suffixes have an implicit relationship with other pronouns of the language (which are presumably theta-marked) suggesting that they are also theta-marked.
- if we suggest that the pronouns are theta-marked, however, we lose the explanation for the typological characteristic for VSO word order with lexical items.
- AgrPs in the IP, and (permissably) null arguments for the verb, explain both where there are lexical subjects and/or objects, and where there are not. It thus seems that the AgrP analysis is the most precise, managing to account for both overt and covert arguments of the verb.

1.3 The Passive Paradigm

There is some question as to whether the passive in Salish languages is really the passive; there is no word order change, and there is no morpheme added to indicate passivization.⁴ The only way that it is identified as passive is by the addition of the passive pronoun paradigm onto the

⁴ It is possible that this passive is a manifestation of the ergative nature of the language.

transitive verb. The paradigm is shown in (8), and an example in (9).⁵

| | | | | | | | |
|-----|----|-----|--------|----|-----|--------|-----------------|
| (8) | sg | 1st | -tinic | pl | 1st | -tinit | (Newman, 1969b) |
| | | 2nd | -ct | | 2nd | -tap | |
| | | 3rd | -im | | 3rd | -tim | |

(9) knix-tim-c' (Newman, 1969a)
 eat(trans)-3pl(pass.)-now
 "they are now being eaten"

Jelinek and Demers (1983), in their study of Coast Salish languages, found that not everything can be expressed in both the active and passive voice. They mentioned that in Squamish it is impossible to say the equivalent of "the man helps you", when there is a second person patient; you must instead say "you are helped by the man" (i.e. using the passive paradigm). This has an interesting parallel in Bella Coola since the second person form of the passive paradigm is *identical* to the corresponding suffix of the transitive paradigm where there is a second person object and a third person subject. The example below can, in fact, have three different meanings.

(10) k'x-ct (Newman, 1969b)
 see(trans)-2sg(pass.) or 3S/2sgO
 —"he sees you (sg)"
 —"they see you (sg)"
 —"you (sg) are seen"

We thus see that (at least with second person patients) the passive paradigm acts as the transitive paradigm with an unspecified agent. This may only be due to suppletion within the system, or perhaps it gives insight into the psychological organization of the passive for speakers of Bella Coola. The latter idea is unlikely, in that if it was true, we would probably see the entire passive paradigm supplanted by transitive correlates.

Assuming that the passive paradigm is distinct from the transitive paradigm, how do we represent them in relation to one another? In English, some researchers believe the "be-en" morphology inhibits the assignment of accusative case to the object, causing it to move to a case position, but how can the passive paradigm in Bella Coola express the same process that we assume happens during English passivization? This bears directly on what we decided in section 1.2 - whether S/O marking on transitive verbs is agreement or a realization of arguments. With the assumption that this marking is simply agreement, we can suggest that the passive paradigm is also agreement, which only happens to indicate the patient of the verb. Perhaps the surface realization

⁵ This is the 'regular' passive paradigm; there is an equivalent 'causative' passive paradigm which serves the same function to the 'causative' transitive paradigm (see note 2).

assumed that the determiners indicate the function of a particular full word. It is then possible to specify the main verb as simply having Determiner Phrase' (DP) arguments.⁷ The question then arises as to what constitutes a DP; could it be considered as a D and a VP, or a D and an IP? In any case, it does not seem to be an NP (unless identical realizations can be considered as either noun or verb, depending on whether determiners are present — this does not seem likely, or aesthetically pleasing). Since a VP may act as a full clause, I will propose that for Bella Coola the D subcategorizes for an IP. This may not show the same pattern as other languages do within GB, but it is the most descriptively accurate and may even suggest some sort of special symmetry within the language; there have been no examples of overt complementizers in the language (CP → C IP), and it seems that they take the same complements as determiners, so why not conflate the two categories? Other possible evidence for the non-existence of a CP in the language includes the lack of movement, both NP and WH. This is, of course, only a preliminary suggestion and will have to be thoroughly investigated before anything concrete can be proposed.

2. IMPLICATIONS OF WORD ORDER CLASSIFICATION FOR BELLA COOLA

2.1 Background

The study of language based on word order typology has allowed us to compare genetically unrelated languages with the same fundamental word order in order to see what other aspects of their syntactic organization follow from this basic typology. The study of VSO languages has, for example, come up with the point that many languages have the same underlying word order (SVO). Greenberg (1963) has suggested that VSO languages should (implicationally) show the following characteristics (because they are head-initial): prepositions rather than postpositions; adjectives should follow their noun; SVO as the only alternate word order; interrogative words or phrases must occur first in an interrogative word question; the inflected auxiliary (if there is one) must occur before the main verb. These points, and others brought up by Woolford (1991) will be addressed for Bella Coola.

2.2 Implications of VSO typology

To begin with, Greenberg's universals will be investigated for Bella Coola; his first point, that there should be prepositions in the language rather than postpositions, is realized in the language fairly obviously. There are two main prepositions in the language, one indicating "to, towards", while the other one indicates "at, on". These prepositions take DPs as their complements, and the PPs themselves can be considered complements to the verb.

- (13) kt-c ʔut-ti-t'xt-tx
 fall-1sg on-det-rock-demonstr.
 "I fell on this rock"

⁷ I accept the proposal that Fukui and Speas (1986) make concerning the use of DP instead of the traditional NP. I feel that it amplifies the distinction between functional and lexical categories and demonstrates symmetry in the system.

⁸ Assuming, from section 1.2, that it is in fact the lexical items (or null elements) within the VP which bear the theta-roles (and not the 'pronouns', which are only agreement).

In Bella Coola we don't see adjectives following their heads, since adjectives form their own constituents (similar to DPs). This doesn't necessarily mean that the language is not head-initial, but just that adjectives perform a different function in this language (perhaps more along the line of appositives, where the subject or object is being modified by an equational-type phrase). As for SVO word order as the only alternate ordering, there is insufficient data to substantiate the claim. There are no attested examples of SVO word order, but it is not known whether they are not permitted or simply were not elicited. The interrogatives do come at the beginning of the interrogative clause, but only because they act as the predicate of the sentence, as can be seen in (14). When they are uninflected, as in (15), they have a modifying role.

- (14) ʔustam-nu-ks (Newman, 1969a)
 go where?-2sg-interrogative
 "where are you going"
- (15) wat-l'ks ʔatq"ut ʔac
 who?-interr. book these
 "whose books are these"

As for Greenberg's last point, the location of the auxiliary, this is not possible to confirm since there is no auxiliary in Bella Coola. In other Salish languages, however, there is an auxiliary, and it does occur before the main verb.

Comments will now be made concerning Woolford's (1991) classification of VSO languages. Her primary purpose in writing the paper was to prove that all types of VSO languages can be accounted for by proposing that the arguments of the verb are located within the VP at DS (see Fukui and Speas, 1986; Koopman and Sportiche, 1990). She then went on to show that this analysis gives such a general (and simple) account of the phenomena involved, that it accounts even for (apparently) nonconfigurational languages. She divided the VSO languages into those which seem to have flat structure within the VP, and those that have hierarchical structure. The original purpose of this paper was simply to use her tests in an attempt to determine the structure of Bella Coola VPs. This analysis follows.

Looking for evidence of a VP-internal subject, Woolford begins by looking at areas which have been investigated fairly regularly by others in the field (namely McCloskey, 1983; Sproat 1985 and Choe, 1986),

It does not appear that Bella Coola can be confirmed as either of Woolford's two classes of VSO languages (flat vs. hierarchical VP) on the basis of binding evidence; where there is ambiguity in an English sentence such as "Peter's father saw him", there is no ambiguity in the corresponding Bella Coola one. This is due to the deictic markers that make up one aspect of the determiners. The deictic system marks the possessor with the proximate (visible) form, while the possessed is marked in the distal (invisible) form.⁹ The following example demonstrates the structure.

⁹ While this may not make total logical sense (ie. it is the father who you are seeing, and thus must be visible but is, in fact, marked with distal), it is common in other Salish languages, and even occurs in Algonquian.

- (16) k'x-ic-c'n wa-sut-s ta-mnat-nu (Newman, 1969b)
 see-S/O-now distal-house-3sg.poss prox-son-2sg.poss
 "I see the house of your son"

There may be other structures where binding phenomena are visible, but at this stage, without the help of a native speaker, it is most difficult to find examples.

Woolford's next criterion which gives an idea about the internal structure of the VP is word order. In VSO languages, the following are word orders that would be generated in the base: Infl V S; Comp Neg V S. Since Bella Coola does not have an overt auxiliary or complementizer, the only remaining category which can give us a clue as to the structure of the VP is Neg. If, according to Sproat (1985) for example, the subject is generated in the Spec of IP, there will then have to be movement of the verb out of the VP, and even higher than Spec of IP in order to get the proper word ordering. The example below shows the realization of a Bella Coola sentence with negation (Newman, 1969b).

- (17) qax" ?inu-s ta-mna-c
 not be you.sg(intrans)-3sg det-son(intr)-1sg
 ("he is not being you, he who is my son")
 "you are not my son"

This example suggests that if the subject is generated in the Spec of IP, the negative particle is going to have to move out of the IP. . . but to where? Since Bella Coola does not seem to have (overt) complementizers, there is no concrete proof as to whether it has a CP, although some people claim that all languages have COMP even if it is not (always) realized. If we assume, however, that the subject is generated within the VP, we will only have to explain the movement of the verb to a position preceding the subject. This analysis may then support the analysis in section 1.2 with the diagram in (7) showing the agreement phrases within the IP. The verb is moved through each functional head (including AgrO and AgrS), and morphemes are attached to indicate what is happening at each level. We would only have to suggest an additional node for the negative phrase.

There is no proof for [Spec,IP] in Bella Coola due to the lack of movements that take place. It seems that there is neither passive-NP movement nor WH-movement. Woolford brings up [Spec,IP] in order to prove that it is empty at the D-Structure, and thus that any word order evidence indicating that it is filled signals that a movement has taken place. Since we do not see any word order evidence to suggest that [Spec,IP] is ever filled, we are left with the conclusion that the subject remains within the VP at surface structure.

Comparable to the subjects, the verbs also show a lack of movement once they have been incorporated with INFL (with agreement markers); this becomes apparent when we try to move verb-object constituents out of the VP. This does not give us any proof for a hierarchical structure within the VP, and is the first suggestion that Bella Coola may have a flat VP.

The lack of any other movements (by which we can test parasitic gap constructions, for example) indicates that there may not be an asymmetry between subjects and objects in the

language, or perhaps there is but we are just unable to see it. We are therefore unable to commit the language as either having a flat or hierarchical VP. This may seem to defeat the purpose of this analysis, but I believe it simply opens the door for more extensive study (with a native speaker). At the very least, Bella Coola does not contradict Woolford's proposal, just as the supposedly nonconfigurational languages did not.

3. ADEQUACY OF THE GB ACCOUNT OF BELLA COOLA

The most glaring problem in the proposed GB analysis for Bella Coola is the issue of what makes up the DP. Inherent in this problem lies the more fundamental difficulty of the transitivity and detransitivizing affixes that occur between the most common paradigms of the language. If we assume that both a transitive form and an intransitive form are stored in the lexicon for each root that can have both forms (along with its transitivity or detransitivizing affix), we are positing a much larger lexicon than is necessary, even though it will explain roots that can have both transitive and intransitive forms without additional markers. On the other hand, how are we to change the argument structure of a verb during the derivation if we have only the *roots* in the lexicon? If we see the detransitivizing affix, for example, as occupying the complement to the verb in the deep structure, the phenomenon can possibly be explained by this affix being assigned the theta role by the transitive verb, and simply absorbing it so that no other DP can be assigned that theta role.

The transitivity affix is a little more difficult to explain; where can it occur in the VP to introduce another theta position? This seems patently impossible since the verb must have its theta grid when it is generated in deep structure. The only way that it would seem possible to do something like this in the syntax would be if there was an oblique marker on the object of the derived transitive verb. This does not happen. Once the verb has been transitivity affixed, it takes the same pronominal agreement paradigm and lexical complements as any "root transitive" verb; there is absolutely no difference between the two. The same is true of the detransitivizing affix, in that the pronominal markers on the verb treat both derived intransitives the same as "root intransitives". There only seem to be semantic factors which limit the transitivity or detransitivizing process.

Current proposals suggest, however, that the transitivity-changing affixes may have their own theta-grid, with specifications for inheritance of theta-roles from the verb.¹⁰ The specifications on these affixes may indicate whether certain theta-roles are blocked or absorbed. It also seems possible that new theta-roles may also be introduced with the affixes, but there isn't anything concrete concerning this. The following is an example of a typical detransitivizing affix that will demonstrate two different ways of analyzing these types of affixes.

- (18) mnck-ic (Nater, 1984)
count(trans)-1sgS/3sgO
"I will count them (objects)"

¹⁰ See Randall (1988), for instance.

- (19) mnck-m-c
count-detrans-1sg.poss
“I am busy counting”

We can either see the detransitivizing affix as actually occupying the object position in the deep structure, or as a derivational affix on the verb which changes the theta-grid of the verb, blocking the THEME role. If we go with the first suggestion, it will be assumed that the affix is some sort of unspecified (or unspecifiable) object (although it could plausibly be considered some sort of aspect marker — durative). This particular example shows the close relationship between the transitive pronominal markers and the intransitive ones; most notably the /-c/ which indicates 1st person singular with the detransitized verb also indicates 1st person singular in the fused-form /-ic/. When the elements making the fused-form of the ‘intransitive’ come together, they have the same ordering (ie. object, subject) as the transitive version, supporting this proposal. Not all the pronouns in the transitive paradigm are this transparent, however, making this analysis somewhat doubtful.

If we are going to accept the second proposal (that the affix is affecting the theta-grid of the verb), we may have to decide where the affixation is taking place: in the lexicon or in the syntax. From what Randall (1988) has suggested though, it doesn’t seem to matter; if the blocking of the THEME role is occurring in the syntax, it won’t interfere with any syntactic rules, although it may be difficult to suggest a location for the affix (unless it is simply adjoined to the verb in the base). This analysis seems very well suited for the problems in Bella Coola, but there is one possible catch: Randall states that if a theta-role is blocked, all the roles below it on the thematic hierarchy will also be blocked. Since THEME is the highest theta-role (in this study, at least), all theta-roles will be blocked, which suggests that there will be no AGENT role, while there actually is. Randall may account for this, however, by the optional Absorption of lower theta-roles, by the affix. The choice of *which* roles may be absorbed is where her proposal needs some more work.

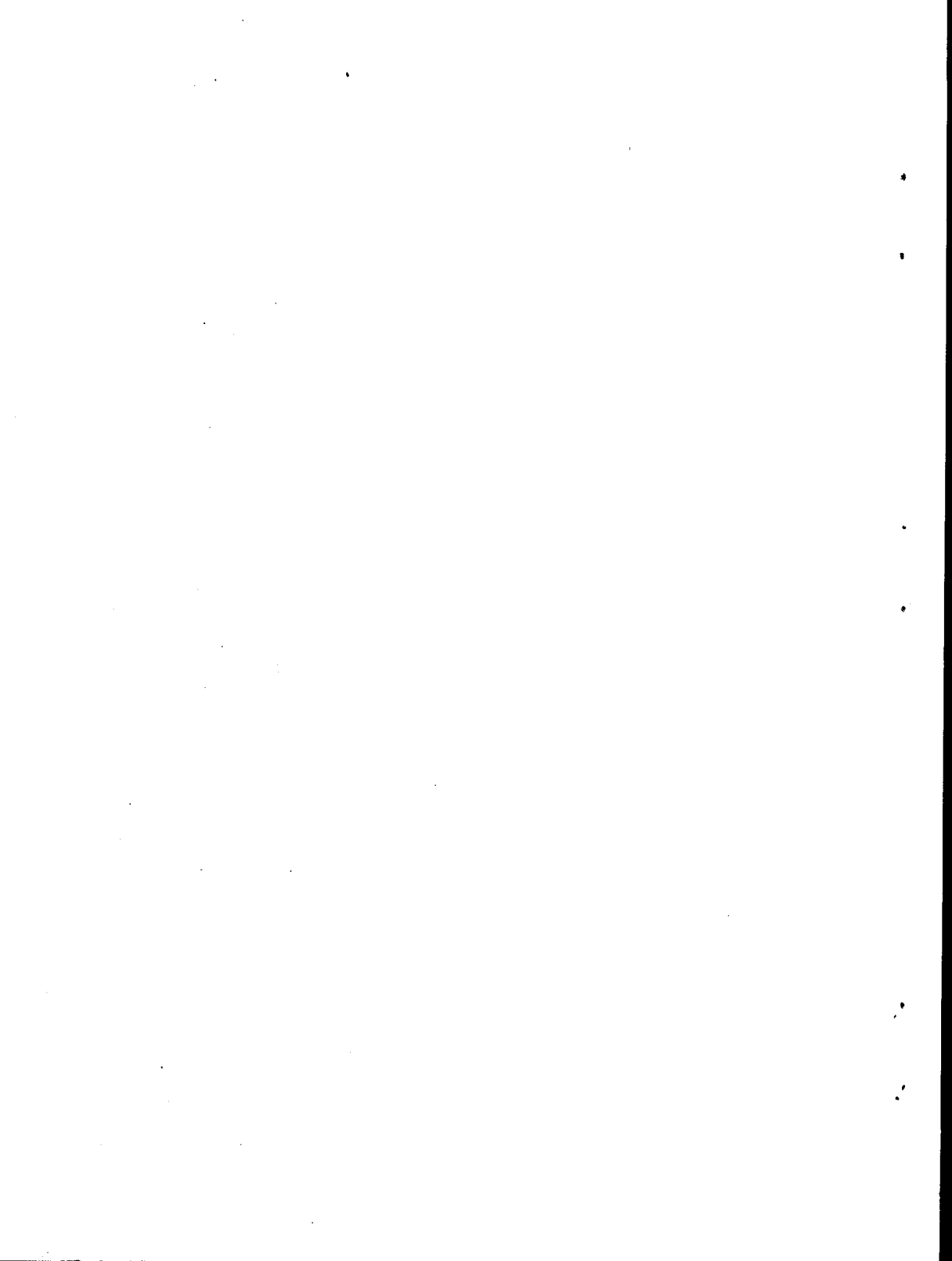
There does not seem to be an altogether obvious way to test for the two main modules of GB within Bella Coola (namely Government and Binding), but there is no reason to believe that all languages will exploit all modules of the theory. The account of Bella Coola may even be simpler without having to resort to them (although a theory usually likes to get all the support it can). In conclusion, it will take extensive study of the language with a native speaker to determine the structure of the D-Structure VP. It is difficult working with material recorded by people active in other frameworks and eras; everyone does not always think to elicit certain (possible) alternate forms from the speakers, and the analyses are invariably coloured by different objectives. There are enough points raised here to question certain aspects of current GB theory, but overall it does account fairly well for Bella Coola without offering any glaring contradictions.

REFERENCES

- Bach, Emmon. (1968). Nouns and Noun Phrases. In: Bach, Emmon, Robert T. Harms, eds. *Universals in Linguistic Theory*. New York: Holt Rinehart and Winston.
- Choe, Hyon Sook. (1986). An SVO analysis of VSO Languages and Parameterization: A Study of Berber. In: *Proceedings of the Sixteenth Annual Meeting NELS, GLSA*. Amherst: University of Massachusetts.
- Davis, Philip W. and Ross Saunders. (1984). An Expression of Coreference in Bella Coola. In: E.-D. Cook, Donna Gerds, eds. *Syntax & Semantics*, Volume 16. New York: Academic. 149-167.
- Fukui, Naoki and Margaret Speas. (1986). *Specifiers and Projection*. ms. Cambridge, Mass: MIT.
- Greenberg, Joseph. (1963). Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements. In: J. H. Greenberg, ed., *Universals of Language*. Cambridge, Mass: MIT Press.
- Jelinek, Eloise and Richard Demers. (1983). The agent hierarchy and voice in some Coast Salish Languages. *IJAL*, 49 (2). 167-185.
- Koopman, Hilda and Dominique Sportiche. (1990). *The Position of Subjects*. ms. UCLA.
- McCloskey, James. (1983). A VP in a VSO Language? In: G. Gazdar et al, eds. *Order, Concord and Constituency*. Dordrecht: Foris.
- Nater, Hank F. (1984). *The Bella Coola Language*. Canadian Ethnology Service. Paper 92.
- Newman, Stanley. (1969a). Bella Coola Grammatical Processes and Form Classes. *IJAL*, 35 (2). 175-179.
- Newman, Stanley. (1969b). Bella Coola Paradigms. *IJAL*, 35 (4). 299-306.
- Randall, Janet H. (1988). Inheritance. In: Wilkins, Wendy, ed. *Syntax & Semantics*, Volume 21. New York: Academic. 129-46.

Sproat, Richard. (1985). Welsh Syntax and VSO Structure. *Natural Language and Linguistic Theory*, 3. 173-216.

Woolford, Ellen. (1991). VP-Internal Subjects in VSO and Nonconfigurational Languages. *Linguistic Inquiry*, 22(3). Summer. 503-540.



ON THE ACQUISITION OF WH-QUESTIONS

Ken Hanna & Andrea Wilhelm
The University of Calgary

O. ABSTRACT

This paper is the account of a study carried out in 1990 in which we tried to gain further insight into the acquisition of subject and object *wh*-questions. We differentiated between two *wh*-words, *who* and *what*, for subject and object questions. The study consisted of a production and a comprehension task and was carried out with 11 children aged between 3;4 and 4;7 years.

The results suggest that children find subject questions easier to produce and comprehend than object questions. In addition, the animate question pronoun *who* was used more often than the inanimate question pronoun *what*.

1. INTRODUCTION

1.1 Research on *wh*-questions

The research on *wh*-questions has focused on a number of different issues, such as the order of acquisition of *wh*-words (for example, Tyack & Ingram 1977), the interpretation and processing of *wh*-questions by children (Ervin-Tripp 1970), and the acquisition of different syntactic types of *wh*-questions (for example, Stromswold 1988). Of this research, we will give a brief review of only the work that is most relevant to our topic. Since the occurrence of the earliest studies on *wh*-questions, the order of acquisition of *wh*-words is agreed upon. Thus, Smith (1933), Klima and Bellugi (1966), Bloom et al. (1982), Tyack and Ingram (1977) all stipulate the following order:

- (1) Order of acquisition of *wh*-words
 - where, what
 - who
 - how
 - why
 - which, whose, when

Kuczay & Brannick (1983) found roughly the same order of *wh*-words, as they first occur with auxiliaries. What we can note as relevant here is that *what* questions were generally asked before *who* questions. It is not obvious from the findings whether these interrogative pronouns first occurred in object or in subject position. The adult form of *wh*-questions, with an auxiliary and with inversion, if necessary, seems to be acquired around age 3. This is indicated by the results of Johnson (1980) who found occasional inversion with children between 1;6 and 3;0 years of age. Ingram and Tyack (1979) found 77-98% inversion in the questions of 2;00 - 3;11 year olds, and Erreich (1984) found 64% inversion with 2;5-3;00 year olds.

Let us now turn to the acquisition of subject and object questions. In the study carried out by Tyack and Ingram (1977), the results indicated that children found subject questions easier than object questions. Thus, the 100 observed children (aged 3;00 to 5;5) would give 80% correct responses to subject questions, but only 57% correct responses to object questions. Furthermore,

the children gave more correct responses to questions containing an intransitive verb (66%) than to questions containing a transitive verb (57%), and there was a tendency to interpret transitive questions as intransitive (where the sentence structure would permit it). Note that intransitive questions can never be object questions, but must always be subject (or adjunct) questions. These facts indicate a preference of subject over object questions. This preference is also reflected in one of the response strategies that Tyack and Ingram posit (p.222):

(2) Response Strategy

If you have acquired the particular question word, give an appropriate subject.

Tyack and Ingram also found an interesting pattern with respect to the interpretation of the pronouns *who* and *what* in subject and object position: *what*-subject questions were very frequently answered as if they were object questions (51.3% - with only 35.3% correct responses), and *who*-object questions were treated as *what*-object questions 23.0% of the time. According to the authors, this can be explained as an animacy error, due to the fact that animate entities are more likely to be encoded as subjects, and inanimate entities are more likely to be encoded as objects. The children seem to have paid more attention to these semantic tendencies than to the syntactic structure of the questions. However, the 23% responses to *what*-object questions as if they were *who*-object questions cannot be explained through this semantic strategy.

1.2 Some controversial findings

An article by Stromswold (1988) on the order of acquisition of subject and object *wh*-questions has led to controversy on the issue. Stromswold objects to the earlier studies on the grounds that *comprehension* tasks might not be the appropriate means to find out whether subject or object questions are acquired earlier. She focuses on the *production* of these question types, namely on the first occurrences of subject and object *who*-, *what*- and *which*- questions in the spontaneous speech transcripts of 12 children. The children were between 0;11 and 2;10 years old when the transcripts began. Stromswold found that the age of acquisition for object questions generally precedes the age of acquisition for subject questions. Likewise, object questions are in general more frequent than subject questions. The explanation for these unexpected findings is based on the syntactic structure of the question types: Children find object questions easier than subject questions because the traces of object questions are directly theta-governed by the verb, whereas the traces of subject questions are antecedent-governed by the *wh*-word.

A closer look at Stromswold's paper reveals some problems, however. First of all, some of the older children whose transcripts she examined might have produced their very first *wh*-questions before the transcripts were even started (eg. Ross: starting age 2;10, Nathan: starting age 2;6, p.108). Thus, we cannot be sure that the corpus of data that served as the base of the study is entirely reliable.

Second, once the findings on which Stromswold bases her conclusions are studied in more detail, we find that the age of acquisition for object questions is not always earlier than for subject questions (e.g., *who* subject questions were acquired 0.5 months earlier than *who* object questions). Furthermore, the differences in age of acquisition are rarely statistically significant. Thus, *what* object questions preceded *what* subject questions by only 2.5 months which, in Stromswold's own words, is "marginally significant" (p. 108). The only really significant difference is the 8 months which separate the occurrence of *which* object questions occur from

which subject questions.

The same weakness of the evidence can be observed with respect to the frequency of occurrence. *What* occurs 88% of the time in object questions and *which* 77% of the time. However, *who* only occurs in object questions 33% of the time, and occurs in subject questions for the rest of the time.

The evidence on which Stromswold bases her general conclusions is thus not as strong as one would wish, especially if one takes into account some additional factors that were found in earlier studies to be responsible for such a distribution of subject and object questions: For example, Ervin-Tripp (1970) states that *who* was recognized as [+animate] already by the time testing began (p. 90), which varied between 1;11 and 2;5 years. And the animacy strategy found by Tyack and Ingram (1977) describes children's tendency to link animate entities to subjects and inanimate entities to objects, including subject and object questions. These facts explain smoothly why, in Stromswold's findings, inanimate *what* and *which* occur so often in object questions, whereas animate *who* occurs in subject questions.

Thus, the different kinds of government are not necessarily the only or best explanation for Stromswold's findings. It is clear from the earlier studies that the different question pronouns show different syntactic and developmental behavior due to their semantic properties. We can conclude that the acquisition of subject and object questions is an unsettled issue and needs more research.

The study that we have undertaken is designed to provide more insight into this issue.

2. EXPERIMENT DESCRIPTION:

Subjects: The research experiment was carried out with eleven monolingual English children, six boys and five girls; ranging in age from three years, four months to four years, seven months. This age group was chosen because younger children would not have been able to fulfill the task. We discovered this when our youngest child could not do the production portion of the test. The combined average age of the eleven children was approximately three years nine and one-half months. The children were studied at two different day-care centres.

Tokens: The tokens that we tried to elicit in the production task and that we used as questions in the comprehension task had the following structure:

(3) tokens

a. ANIMATE SUBJECT QUESTION: Who is [V]-ing the [N]? e.g. Who is pushing the cow?

b. INANIMATE SUBJECT QUESTION: What is [V]-ing the [N]? e.g. What is hitting the car?

c. ANIMATE OBJECT QUESTION: Who is the [N] [V]-ing? e.g. Who is the bear biting?

d. INANIMATE OBJECT QUESTION: What is the [N] [V]-ing? e.g. What is the tractor pulling?

Six tokens of each type were used. For each type, half of the nouns were animate (animals) and half of the nouns were inanimate (vehicles), but the nouns within each token were either both animate or both inanimate. The verbs were all reversible in order not to give the child any semantic hints whether a subject or an object question was being elicited/asked.

Props: The props used in this study consisted mainly of pictures in which either animals or vehicles would be depicted carrying out an action. For the production portion of the test, children were shown pictures of either two animals or two vehicles involved in some type of action (e.g.: a cow kicking a horse; a car pulling a tractor). For each picture, either the subject or object would be covered to provide an opportunity for the child to initiate a question about the covered part of the picture. For the comprehension portion of the test, three animals or three vehicles were depicted in the array A B C, with A doing something to B, and B doing the same to C (e.g.: a bear (A) biting a lion (B), which in turn is biting a fox (C); a tractor (A) pulling a car (B), which in turn is pulling a truck (C)). In this part, none of the pictures were covered and only a response was required. Scoresheets were used to record all results; in addition, a tape-recorder was used on the production portion of the test in order to assist the experimenters in the verification of data.

Scoresheets

i) Production: The data sheet used for this portion of the test consisted of a listing of the twelve pictures with various responses that the child could give. The different categories of responses were; a) correct response, b) wrong *wh*-word, c) use of the form do, d) cannot answer, e) wrong verb, f) use of *who/what* is it?, g) incorrect/other (which was always completely recorded). These different categories allowed for an easy recording of the children's responses and errors.

ii) Comprehension: The data sheet for the comprehension test consisted of a listing of the twelve tokens/situations on the pictures with three different answer possibilities (correct NP, incorrect NP, incorrect NP: contained in question) that corresponded to each of the three figures on the given picture. Each response given by a child was registered under the appropriate category, allowing for an easier interpretation of the results.

Experiment: The experiment consisted of three parts: action types, production (which contained a modeling section) and comprehension. The order of the production and the comprehension part was alternated from child to child.

A) *Action Types:* For this part of the study, children were shown different pictures of both animals and vehicles engaged in various types of actions. The purpose of the test was to make sure that the children could identify the animals and vehicles and also to determine whether the children could understand the actions represented within the pictures.

B) Production:

i) Modeling: For this portion of the test, the production task was modeled between the two experimenters in order to demonstrate the required responses. One token of each type was modeled for the children, one *who* subject, one *who* object, one *what* subject and one *what* object question (e.g.: a picture of a bear biting a lion would be shown, with the lion covered over. One experimenter would say: "The bear is biting someone (and I know who); can you make up a question to find out who?" Experimenter two would respond: "Who is the bear biting?" In response to the correct question, experimenter one would uncover the picture. After each question type was modeled, an experimenter would inquire as to whether the child understood the task being performed. After modeling the fourth question an experimenter would ask the child: "Can you make up a question like experimenter two did?" After all the modeling was done, the production test was initiated.

ii) Production: The production task itself consisted of twelve pictures, three for *who* subject object questions. All twelve pictures were mixed to represent a random order of sequence for each

type of picture, so as to eliminate any learning effect. Each picture was represented in such a way that the child had to answer with a question (e.g.: The bear is biting someone. Can you make up a question to find out who?). Each response was recorded as the child uttered it.

C) *Comprehension*: The comprehension portion of the test was used to assess the child's ability to comprehend each question type. Here, as in the production test, three pictures for each question type (*who* and *what* subject and object) were shown to the child in a mixed order prepared by the experimenters. The pictures contained three figures A, B, C (either three animals or three vehicles) as described above. A subject or object question would be asked about the figure in the middle (B) and the response was recorded (e.g.: The picture would show a bear biting a lion who, in turn, is biting a cow. The child would be asked the question, "Who is the lion biting?" In this case, the correct answer would be "The cow."). The child could respond either verbally or simply by pointing to the desired figure.

3. ANALYSIS OF RESULTS

Table 3.1 represents the number of correct responses to each question type given by the children in both production and comprehension.

| Question Types | | Production | Comprehension |
|----------------|------|---------------|---------------|
| SUBJ. | who | 19/33 = 57.6% | 27/33 = 81.8% |
| | what | 13/33 = 39.4% | 19/22 = 86.4% |
| OBJ. | who | 11/33 = 33.3% | 30/33 = 90.9% |
| | what | 11/33 = 33.3% | 27/33 = 81.8% |

Table 3.1: Correct responses to Production and Comprehension Tasks.

In the comprehension part we found that one picture was ambiguous with respect to what were the agents/subjects and what were the themes/objects. This was reflected by the children's answers. We therefore excluded this picture from the comprehension task and did not count its results. This is reflected in the lower total of subject *what* questions in the comprehension part (22 instead of 33).

The table shows that the children were more successful on the comprehension (81.8 - 90.9% correct) than on the production, where only 33.3 - 57.6% correct responses were given.

In the production task, children were most successful in producing *who* subject questions (57.6%). The success rate for *what* subject questions was much lower (39.49%), and even lower for the object questions (33.3% each).

In the comprehension task, there were slight differences with respect to the question pronouns: more *what* than *who* subject questions were answered correctly (4.6% difference), and more *who* than *what* object questions were answered correctly (9.1% difference). However, these differences do not seem to be (statistically) significant, and they might even be the reverse with (an) additional child(ren).

Table 3.2 represents the overall success rates in the production of subject and object questions with respect to the two age groups: the five three-year olds (3;4 - 3;10) and the six four-year olds (4;1 - 4;7).

| Age | Subject Qu. | Object Qu. |
|---------|---------------|---------------|
| 3;4-3;6 | 12/30 = 40.0% | 13/30 = 43.3% |
| 4;1-4;7 | 21/30 = 70.0% | 15/30 = 50.0% |

Table 3.2: Correct Responses to Production Tasks by Age.

As can be seen, the four-year olds in general performed much better than the three-year olds. The latter group has a success rate of approximately 40% for subject and object questions (there is no big difference in the production of subject and object questions for the three-year olds). The four-year olds did slightly better on the object questions than the younger group (50.0% compared to 43.3%), but they did much better on the subject questions (70.0% correct). The success of the four-year olds with the production of subject questions is the most significant result of this table.

Table 3.3 represents the overall success rate for the comprehension of subject and object questions with respect to the two age groups; as illustrated above.

| Age | Subject Qu. | Object Qu. |
|----------|---------------|---------------|
| 3;4-3;10 | 20/25 = 80.0% | 25/30 = 83.3% |
| 4;1-4;7 | 26/30 = 86.7% | 32/36 = 88.9% |

Table 3.3: Correct Responses to Comprehension Tasks by Age.

The differences between subject and object questions within each age group are negligible (3.3% and 2.2%). Furthermore, the age difference is not so striking here: with both subject and object questions, the older group performed only slightly better (6.7% and 5.6%). The most significant result of this table is that all ages were very successful in the comprehension of subject and object questions.

In comparing tables 3.2 and 3.3 for the success rates of subject and object questions, it first seems that the object questions are produced/answered with more success three times (production of three-year olds, production and comprehension of four-year olds), and only one time are the subject questions produced/answered with more success (production of four-year olds). This seems to indicate that the overall success rate with object questions is higher than the success rate with subject questions. However, this is an incorrect analysis because it does not take into account that the success rates for the object questions are only minimally higher than the corresponding rates for subject questions, whereas the success rate for subject questions in one instance exceeds the corresponding rate for object questions by far. Thus, we can conclude from the comparison of tables 3.2 and 3.3 that, if there is a significant difference in the successful comprehension and production of subject and object questions, the difference is in favour of the subject questions.

Now that we have analyzed the children's success rates with respect to the different question types, *wh*-words, and age groups, let us turn to the most common errors that occurred: For the production task, the most common error was to choose the wrong *wh*-word. *Who* was chosen instead of *what* 13 out of 66 times (19.7%), and *what* was chosen for *who* 6 out of 66 times (9.1%); in sum, the error was committed 19 times out of 66 (28.8%), and *who* was chosen for *what* twice as often as vice versa. The table below shows the correlation of this error to the question type (subject/object):

| <i>who</i> for <i>what</i> in | | <i>what</i> for <i>who</i> in | |
|-------------------------------|----------|-------------------------------|----------|
| subj. qu. | obj. qu. | subj. qu. | obj. qu. |
| 6/66 | 7/66 | 2/66 | 4/66 |
| 13/66 = 19.7% | | 6/66 = 9.1% | |

Table 3.4: Production error types.

We can see that there is no important correlation between this error and the question types, but we can note a general tendency of the children to ask *who* rather than *what* questions.

The next most common error was to invert subject and object questions (24 times out of 120; 20%), e.g. to ask "Who is the cow pushing?" instead of "Who is pushing the cow?". An object question was used instead of a subject question 3 times out of 120 (0.25%), and, more strikingly, a subject question was used instead of an object question 21 out of 120 times (17.5%). The age did not make a difference for how often this error occurred. Thus, one child aged 3;6 made as many errors as a child that was one year older (4;6): both confused subject and object questions 6 times.

In some instances a child would choose the wrong verb type, e.g. pull instead of push. This error occurred 16 times out of 120 (13.3%). Fourteen (14) times of 120 (11.7%) a child was unable to respond to the elicitation task.

In the comprehension part, not many errors were made at all: 16 times out of 120 (13.2%) subject and object questions were confused. As the table below shows, there is no significant correlation between these errors and the type of *wh*-word used:

| <i>who</i> questions wrongly interpreted as: | | <i>what</i> questions wrongly interpreted as: | |
|--|----------|---|----------|
| subj. qu. | obj. qu. | subj. qu. | obj. qu. |
| 6 | 3 | 3 | 4 |
| 9/66 = 13.6% | | 7/55 = 12.7% | |

Table 3.5: Comprehension error types.

4. DISCUSSION AND CONCLUSIONS

The overall response to the tasks was good. Only one child, the youngest one, refused to do the production task. The other ten children completed it, although they did not enjoy it as much as the comprehension task. The children probably did not find the situation natural enough to ask questions (instead of "making them up"), and would have preferred to simply guess. No child had any problems understanding the comprehension task. As Table 3.3 shows, all the children performed very well on the comprehension task, whereas the younger children had problems with the production task. This is clearly indicated by Table 3.2: The younger children gave fewer than 50% correct responses. We can conclude from this that the age difference played an important role, especially with respect to the production task. The three year olds did not understand this task, or/and they had not acquired the question structures for their "active vocabulary".

We can draw a second conclusion from the data with respect to the *wh*-words used by the children. There was a tendency to use *who* more than *what*, especially for younger subjects (cf. Tables 3.1, 3.4). This tendency might partly be due to the fact that the interviewer gave a clue for *who*, where a clue for *what* would have been correct (on only two occasions). However, this behaviour is also an instance of the "animacy error" that Tyack & Ingram found. According to them, animate nouns/pronouns are preferably used as subjects because they are more agentive. Thus, there is a semantic explanation for this behaviour.

The third important conclusion concerns the choice of subject and object questions. While no significant difference between these two question types was found in the comprehension task, subject questions were definitely easier to produce than object questions. This can be seen in Table 3.1, where (*who*) subject questions were the only ones whose success rate was more than 50%. Table 3.2 shows the same phenomenon: The four year olds (i.e. the ones who understood the task better and were more successful at it) achieved 70% of the subject questions correctly as opposed to 50% of the object questions. The 20% margin demonstrates a sizeable difference between the production of subject and object questions. Furthermore, the most common and striking error to produce a subject question instead of an object question; *which* occurred 17.5% of the time.

These findings confirm the research that was done before to Stromswold by Tyack and Ingram (1977). We now found not only comprehension, but production data that confirmed the hypothesis that subject questions are easier for children than object questions. Our evidence cannot make reference to the age of acquisition of these question types (it is likely that the children produced subject and object questions before the study). However, the children's production behaviour under the somewhat artificial experimental conditions still shows a tendency of acquisition that is opposed to the tendency Stromswold found, which stated that object questions are easier (and acquired earlier) than subject questions.

The facts found in our study indicate the opposite and therefore demonstrate that the explanation suggested by Stromswold cannot be right. Since theta-governed object questions were not easier to produce for children than antecedent-governed subject questions, it is either not true that children find theta-government easier than antecedent-government, or these government relations do not hold at all. This might especially be likely for subject questions where the existence of a gap is questionable. We could then suggest an explanation for the preference of subject questions: Subject questions, in contrast to object questions, might not have a gap.

Subject questions would also be expected to be easier because, as Stromswold herself states (pp. 109 - 110), they do not require inversion, but their structure is homologous to the structure of declaratives (cf. Gazdar 1981, Gazdar et al. 1985).

There is still another reason for the preference of subject questions that has to do more with pragmatic factors and that seems to be supported by the behaviour of the children. For example, in the production task, the children were very eager to find out what was "hiding" under the cover sheet. Some children would try to guess the hidden figure before they would ask a question. This indicates that the figures under question clearly were the topic of the exchanges for the children. Topics, in turn, have been claimed to be encoded as subjects most frequently (cf. Givon, 1984). Thus, we could hypothesize that the children asked more subject questions because they were "focused" on the covered figures, i.e. because the covered figures were the topics. This, then, ties in with the fact that *who* was used more frequently than *what*: animates are more likely to be agents, and agents are more likely to be subjects, which are the major syntactic devices to encode the topic.

5. SUMMARY

Our study was concerned with the acquisition of subject and object *who* and *what* questions. In order to shed more light on this controversial issue, we tried not only to collect comprehension, but also production data. The production data turned out to be of more interest. While all the children were approximately equal in successfully understanding and answering *who* and *what* subject and object questions, some striking differences occurred in the production data. In particular, *who* was the preferred question pronoun, and subject questions were the preferred question type. These results do not confirm Stromswold's (1988) findings and hypothesis, but earlier findings such as those stated by Tyack and Ingram (1977) were confirmed.

We concluded that theta-government cannot be an explanation for the acquisition and use of *wh*-questions by children, but that syntactic, semantic and pragmatic features of subject *wh*-questions make them easier to use and probably also to acquire by children.

ACKNOWLEDGEMENTS

We would like to thank the staff and children of the Bugs Bunny and Happy daycare centres in Calgary for assisting us in this study.

We would also like to thank Joyce Hildebrand and Dr. W. D. O'Grady for their advice and their assistance with this project.

REFERENCES

- Bloom, Lois, Susan Merkin & Janet Wootten. 1982. Wh-questions: Linguistic factors that contribute to the sequence of acquisition. *Child Development* 53, 1084-92.
- Erreich, Anne. 1984. Learning how to ask: Patterns of inversion in yes-no and wh-questions. *Journal of Child Language* 11, 579-92.
- Ervin-Tripp, Susan. 1970. Discourse Agreement: How Children Answer Questions. In: Hayes, J. (ed) 1970. *Cognition and the Development of language*. New York: Wiley. 79-107.
- Gazdar, G. 1981. Unbounded dependencies and coordinate structure. *Linguistic Inquiry* 12, 155-84.
- Gazdar, G. E. Klein, G. Pullum, & I. Sag. 1985 *Generalized phrase structure grammar*. Cambridge, MA: Harvard University Press.
- Givon, Talmy. 1984. *Syntax. A functional-typological introduction*. Vol. 1 Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Ingram, David & Dorothy Tyack. 1979. Inversion of subject NP and Aux. in children's questions. *Journal of Psycholinguistics Research* 8, 33-41.
- Johnson, C. E. 1980. The Ontogenesis of Question Words in Children's Language. Paper presented at Boston University Fifth Annual Conference on Language Development (October).
- Klima, Edward & Ursula Bellugi. 1966. Syntactic regularities in the speech of children. In J. Lyons & R. J. Wales (eds.), *Psycholinguistic papers*. Edinburgh University Press.
- Kuczaj, S.A. & Nancy Brannick. 1979. Children's use of the wh-question modal auxiliary placement rule. *Journal of Experimental Child psychology* 28, 43-67.
- Smith, M. 1933. The Influence of age, sex and situation on the frequency, form and function of questions asked by preschool children. *Child Development* 4. 201-213. Quoted from Tyack & Ingram (1977).
- Stromswold, Karin. 1988. Linguistic Representations of Children's Wh-Questions. In: *Papers and Reports on Child Language Development* 27, 107-114.
- Tyack, Dorothy & David Ingram. 1977. Children's Production and Comprehension of Questions. *Journal of Child Language*, vol. 4. 211-224.
- O'Grady, W. D. 1990. Child Language Acquisition lecture notes. (Ms.)