UWWPL 24

University of Washington
Working Papers in Linguistics

Edited by Daniel J. Jinguji and Steven Moran

University of Washington ✂ Seattle, WA ✂ 2005

To order a copy of this volume or selected previous volumes, contact:

Working Papers in Linguistics
Department of Linguistics
University of Washington
Box 354340
Seattle, WA 98195-4340

206-543-2046

uwwpl@u.washington.edu

http://depts.washington.edu/uwwpl

Table of Contents

Volume 24 (2005)

Coordination and Processing

Scott Drellishak

sfd@u.washington.edu

1 Introduction

One of the ongoing research aims of modern linguistics is accounting for the range of possible phenomena in human language. In particular, generative grammarians working in Chomsky's Principles and Parameters framework have often sought to explain typological generalizations by positing principles of Universal Grammar (UG) that require them, or parameters that can take two or more values that correspond to observed language variation. Such a research program has an obvious appeal: if typological generalizations can be explained by a small set of broad, simple principles and parameters, then linguists will have gone a long way towards characterizing the precise contents of the human language faculty.

This research program, however, has tended to overlook an alternative source of explanations that can often account for typological generalizations, namely limitations of the human language processor. John Hawkins (2004) has proposed a set of processing principles that are intended to account both for preferences within languages for certain kinds of constructions, and for the distribution of typological features across languages. These processing principles offer a way of accounting for statistical universals — that is, "universals" that hold less than one hundred percent of the time — because they assert a preference for more-easily-processed structures without ruling out the alternatives. If the processing principles can account for language variation, a theory based on them is a clear improvement over a theory that requires the positing of new principles or parameters to account for newly discovered phenomena.

In this paper, I will apply Hawkins' processing principles in an attempt to account for an apparent universal in coordination strategies in the world's languages. The second

section is a description of the universal in question. The third section briefly lists and explains the processing principles of Hawkins (2004). The fourth section examines the observed universal in light of the processing principles and attempts to determine whether it can be accounted for in terms of them. The fifth section discusses the findings of this attempt and proposes possible explanations for what is found.

2   The Universal

Stassen (2000) contains a survey of noun phrase coordination in a genetically diverse sample containing 270 of the world's languages. As indicated by its title, Stassen's article broadly divides the world's languages into two groups: AND-languages, in which NP coordination is accomplished with a syntactically balanced structure (similar to those marked in English and the other Indo-European languages by *and* and its cognates), which he calls the *Coordinate Strategy*; and WITH-languages, in which NP-coordination structures are imbalanced, with one of the two coordinands marked in a way that carries comitative meaning, which he calls the *Comitative Strategy*. This paper does not focus on WITH-languages, but rather on a universal Stassen observed in the AND-languages in his sample.

Because the domain of Stassen's survey included only the coordination of two items, there are a finite number of possible morpheme orders in AND-language marking strategies. These strategies can be categorized in two ways: by the number of marked coordinands, and by the position of the marking morpheme. In some strategies, coordinated items are simply juxtaposed without marking (that is, the number of marked coordinands is zero); this is referred to as *asyndeton*. In other strategies, there is a single marking morpheme for the entire coordinated phrase; this is referred to as *monosyndeton*. In still other strategies, one marking morpheme appears for each coordinand; this is referred to as *polysyndeton*. As for position of the coordinating morphemes, they can either precede or follow each of the coordinands. Among the strategies Stassen found in his survey, various possibilities were attested. These included the very common medial monosyndeton, as found in Finnish (Uralic, Balto-Finnic[1]):

---

[1] The language classifications included here are Stassen's.

(1) Pentti ja   Pirkko

Pentti and Pirkko

'Pentti and Pirkko' (Stassen 2000:11)

Stassen also found examples of final monosyndeton, as in Pitjantjatjara (Australian, Pama-Nyungan):

(2) Henry-ku   mama   ngunytju puru

Henry-GEN mother father   and

'Henry's father and mother' (Stassen 2000:15)

Among polysyndeton strategies, Stassen found examples of languages in which the mark followed the coordinands and, more rarely, examples in which it preceded them. The former pattern can be seen in Abkhaz (North-West Caucasian):

(3) s-ànə-y   s-àbə-y

my-mother-and   my-father-and

'my mother and my father' (Stassen 2000:12)

The latter can be seen in a strategy in Sedang (Mon-Khmer) that marks coordinands with dual pronouns:

(4) préi klá   préi koa

3DU tiger 3DU turtle

'the tiger and the turtle' (Stassen 2000:17)

In spite of this variety, in all the languages in his survey, Stassen failed to find any occurrences among the AND-languages of the remaining coordination marking pattern: initial monosyndeton. As he puts it:

To round off the discussion of the various manifestations of the Coordinate Strategy, I can note that monosyndetic preposing on the first NP is not attested at all in the

sample. That is, there do not seem to be languages which conform to the AND-NP NP scheme. (Stassen 2000:15)

Given that there exist languages that exhibit the NP NP-AND strategy, it is curious that no languages exhibit a strategy with the opposite order, especially since that is just the sort of variation commonly observed between head-initial and head-final languages. Why should this typological asymmetry exist? Because the generalization is apparently exceptionless, we might be tempted to assert the existence of a universal principle to account for it; however, such a universal would have to be phrased in such a way that its existence seems improbable. First, rather than broadly applying to any type of construction, the principle would have to address coordination and nothing else. On top of that, the supposedly universal principle would be most economically phrased as a *negative* universal, so that it would specifically rule out initial monosyndeton, but allow medial and final monosyndeton, preposed and postposed polysyndeton, and asyndeton. This would leave us with a principle proposed to be universal, but phrased so narrowly that it would only apply to a particular kind of AND-coordination and to no other construction. Proposing such a principle has little or no explanatory power, because it makes no claims outside the narrow domain of coordination strategies; instead, we should look for other explanations, such as one provided by Hawkins' processing principles.

3   Hawkins's Processing Principles

John Hawkins (2004) proposes a set of performance principles in order to explain the distribution of various structures in human language. The central idea underlying these principles is his Performance-Grammar Correspondence Hypothesis (PGCH):

Grammars have conventionalized syntactic structures in proportion to their degree of preference in performance, as evidenced by patterns of selection in corpora and by ease of processing in psycholinguistic experiments. (Hawkins 2004:3)

According to this hypothesis, a structure that is preferred according to performance criteria should be more common not only within a single language but also cross-linguistically. However, for the purposes of this paper it is important to note that the

converse of this implication does not necessarily hold: alternatives that occur more often in a corpus or across the world's languages need not have a preference in performance, but might have some other cause (common origin, accident of history, etc.). In order to determine whether Stassen's coordination universal is explainable by performance factors, then, we need to evaluate it according to the principles Hawkins spells out.

The first of these principles is Minimize Domains (MiD), which Hawkins defines as follows:

> The human processor prefers to minimize the connected sequences of linguistic forms and their conventionally associated syntactic and semantic properties in which relations of combination and/or dependency are processed. The degree of this preference is proportional to the number of relations whose domains can be minimized in competing sequences of structures, and to the extent of the minimization difference in each domain. (Hawkins 2004:31)

A simple example of this principle in operation is in the preference for short prepositional phrase adjuncts before long ones in English. Consider examples (5) and (6) from Hawkins (2004:104):

(5)   The man $_{VP}$[waited $_{PP1}$[for his son] $_{PP2}$[in the cold but not unpleasant wind]]

(6)   The man $_{VP}$[waited $_{PP2}$[in the cold but not unpleasant wind] $_{PP1}$[for his son]]

Sentence (5) is preferred to (6) because the domain of the VP (i.e. the range of lexical items that must be processed in order to recognize it; see §4.1 below for a more formal definition) contains five words, from *waited* to *in*, in sentence (5), but nine words, from *waited* to *for*, in sentence (6). According to MiD, the sentence with the smaller domain is preferred.

The second of Hawkins' principles is Minimize Forms (MiF), which is based on the straightforward idea that it is easier to process less material than to process more. Hawkins' formal definition is as follows:

> The human processor prefers to minimize the formal complexity of each linguistic form F (its phoneme, morpheme, word, or phrasal units) and the number of forms with unique conventionalized property assignments, thereby assigning more properties to fewer forms. These minimizations apply in proportion to the ease with which a given property P can be assigned in processing to a given F. (Hawkins 2004:38)

MiF prefers structures with less material to those with more. For example, in a sentence in which the grammatical role of a given NP can be recognized by its position, MiF would prefer no marking to the presence of a case-marking morpheme.

The third, and most complex, of the performance principles proposed by Hawkins is Maximize On-line Processing (MaOP), which he defines as follows:

> The human processor prefers to maximize the set of properties that are assignable to each item X as X is processed, thereby increasing O(n-line) P(roperty) to U(ltimate) P(roperty) ratios. The maximization difference between competing orders and structures will be a function of the number of properties that are unassigned or misassigned to X in a structure/sequence S, compared to the number in an alternative. (Hawkins 2004:51)

Because the name of this principle is perhaps less transparent than those of the other two, it requires a bit more explanation. Hawkins' idea is that, as a sentence is being processed, various properties are being assigned to the items in the sentence. When the sentence is finished, some total number of properties has been assigned. Depending on the facts of the language, at some points during processing properties can be assigned immediately, but at other points this assignment is delayed. Consider two hypothetical SOV languages, one in which subject NPs are marked in some way, and another in which they are not. In the first language, an initial NP can be identified as the subject immediately after it occurs, while in the second, it cannot be identified until later—it might turn out to be the object of the verb in an optional-subject language, for example. According to MaOP, structures that maximize the ratio of the number of properties assignable during processing to the final number of properties are preferred. This

principle formalizes the intuition that it is easier to process sentences that do not contain ambiguous forms or garden paths.

4 Coordination and Processing

Now that we have Stassen's observed language universal and Hawkins' processing principles for evaluating universals, it remains to analyze the universal in light of each of the three principles. However, first a description of the structures that will be analyzed and a statement of some assumptions are necessary. The following sections contain a comparison of two coordination structures: initial monosyndeton, which takes the form AND NP* (i.e. a single coordinator followed by any number of noun phrases); and final monosyndeton, which takes the form NP* AND (i.e. any number of noun phrases followed by a coordinator). Other possible locations for the coordinator, including the common NP* AND NP structure, generally fall between the two peripherally marked strategies according to the processing principles. To deal with any variation arising from basic word order, the processing principles are tested on verb phrases from two hypothetical languages: one language whose basic word order is OV and another that is VO, where in each case the O is a coordinated NP. (Subjects are omitted to avoid the possibility of the subject NP being confused for a part of the object in some cases.) For each principle, therefore, four utterances will be considered:

(7)   V NP NP AND (VO, final monosyndeton)

(8)   V AND NP NP (VO, initial monosyndeton)

(9)   NP NP AND V (OV, final monosyndeton)

(10) AND NP NP V (OV, initial monosyndeton)

For simplicity, I will initially assume that the NPs are all single words, although that obviously need not be the case. In addition, where necessary I will also assume a phrase structure for coordinated phrases that is flat and in which the dominating node has the

same category as its coordinands. (11) shows an example of such a structure for two NPs with a final monosyndeton coordinator:

(11)

```
                              NP
              ┌───────────────┼───────────────┐
            NP                NP              and
```
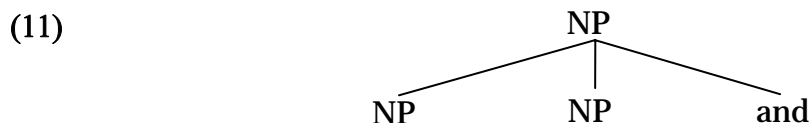
It is important to note that this structure violates two commonly held assumptions of many syntactic theories, namely X-bar structure and binary branching, and also that it does not assume the existence of a CoordP or &P maximal projection. It is assumed here, however, because it is in keeping with the sorts of phrase structures that appear in Hawkins (2004) (such as (12) below, in which the O represents a gap site), and a complete reassessment of his ideas using a different theory of syntax is beyond the scope of this paper.

(12)   [NP VP[V O]] and [NP VP[V NP]] (Hawkins 2004:94)

4.1 Minimize Domains

In order to apply MiD, we first need an understanding of what a domain is. Hawkins (2004) describes them in some detail, but for the purposes of this paper, this definition will suffice: the domain of a node M is "the smallest set of terminal and non-terminal nodes that must be parsed in order to recognize M" (Hawkins 2004:32).

Given this definition, we can apply the principle MiD to the example sentences. In each case, there are two domains to be considered: the domain of the coordinated phrase and the domain of the verb phrase (i.e. the verb and its object). Let us first consider the domain of the coordinated NP, whose extent is controlled by the recognition of the coordinated phrase and all its coordinands. In the examples in which the coordinator is initial, (8) and (10), the domain extends from the coordinator to the last item it coordinates—that is, across the entire coordinated phrase. In the examples in which the coordinator is final, (7) and (9), this remains true: the whole coordinated phrase cannot be recognized until the final coordinator is seen. Next, let us consider the domain of the verb phrase, whose extent is controlled by the recognition of the verb and all its arguments. In examples (7) and (10), in which the coordinator is on the opposite

side of the verb phrase from the verb, the domain of the VP covers all four words, because the listener cannot be said to have recognized it until hearing either the coordinator in (7) or the verb in (10). Examples (8) and (9) are slightly more problematic because it is not immediately clear whether to consider the domain to extend across the whole object to the farthest coordinated NP, or just to the coordinator. Hawkins' analysis of constituent-order preferences in head-initial and head-final languages (2004:104-111) gives us some guidance. For both word orders, he assumes that the domain for the recognition of a VP that contains a verb and an argument PP extends from the verb to the adposition, and not across the whole PP. Extending this assumption to coordination implies that the domain of the VP in (8) and (9) includes just the verb and the adjacent coordinator.

All domains, therefore, are the same size in each pair of examples that contrasts coordinator order. Accordingly, the processing principle Minimize Domains is indifferent to the difference between initial and final monosyndeton.

4.2 Minimize Forms

Applying MiF is even more straightforward. In all four examples, the amount of material is the same, and so MiF, like MiD, is indifferent between initial and final monosyndeton. In addition, it is interesting to note that according to Minimize Forms, asyndeton coordination is the most preferred strategy, since it has one less morpheme than monosyndeton coordination.

4.3 Maximize On-line Processing

Unlike the other two principles, MaOP prefers one of the coordination orders over the other. To show this, we need to consider the processing of each example sentence word by word, keeping track of which final properties have been assigned in order to calculate the On-line Property to Ultimate Property ratio (OP/UP). The ratio for each of the four examples is calculated in the following tables, using the notation of Hawkins (2004:56). Each column shows the properties (Categories, Phrases, Attachments, and Relations) that have been assigned as the sentence is recognized sequentially, as well as the current number of assigned relations and the current OP/UP ratio. NP_1 and NP_2 are used for the coordinand noun phrases, and NP_c for the coordinated NP.

Table 1: OP/UP Ratios

Example (7)	V	NP	NP	AND
Categories	V	NP	NP	Coord
Phrases	VP	NP_1	NP_2	NP_c
Attachments	VP[S]			NP_c[VP], NP_1[NP_c], NP_2[NP_c]
Relations				NP_c=OBJ-V
# Assigned	3	5	7	13
OP/UP	3/13 = 23%	5/13 = 38%	7/13 = 54%	13/13 = 100%

Example (8)	V	AND	NP	NP
Categories	V	Coord	NP	NP
Phrases	VP	NP_c	NP_1	NP_2
Attachments	VP[S]	NP_c[VP]	NP_1[NP_c]	NP_2[NP_c]
Relations		NP_c=OBJ-V		
# Assigned	3	7	10	13
OP/UP	3/13 = 23%	7/13 = 54%	10/13 = 77%	13/13 = 100%

Example (9)	NP	NP	AND	V
Categories	NP	NP	Coord	V
Phrases	NP_1	NP_2	NP_c	VP
Attachments			NP_1[NP_c], NP_2[NP_c]	NP_c[VP], VP[S]
Relations				NP_c=OBJ-V
# Assigned	2	4	8	13
OP/UP	2/13 = 15%	4/13 = 31%	8/13 = 62%	13/13 = 100%

Example (10)	AND	NP	NP	V
Categories	Coord	NP	NP	V
Phrases	NP_c	NP_1	NP_2	NP_c
Attachments		NP_1[NP_c]	NP_2[NP_c]	NP_c[VP], VP[S]
Relations				NP_c=OBJ-V
# Assigned	2	5	8	13
OP/UP	2/13 = 15%	5/13 = 38%	8/13 = 62%	13/13 = 100%

There are several caveats about these tables that should be mentioned. First, the attachment of the VP in the sentence is shown in each case, although the sentence

otherwise does not appear; however, this makes no difference in the analysis of the difference between final and initial polysyndeton, because the VP[S] attachment appears in the same place in pairs of examples contrasting that property. Second, I have assumed that the coordinated NP (NP_c) is constructed immediately upon the occurrence an initial coordinator, following Hawkins, who assumes that a PP can be constructed immediately upon the occurrence of a preposition. It could be argued in the initial monosyndeton examples that NP_c cannot actually be constructed until the appearance of the first NP lets the listener know the category of element being coordinated. This is especially true in (10), in which there is no context to give the processor a clue as to the category, although in (8) it might be argued that, since a verb has already been heard, the processor is expecting a noun phrase. However, this would affect the calculations only slightly, since it simply delays the identification of the NP_c for one word. It could also be argued that, because many languages allow coordination by juxtaposition, the occurrence of two adjacent NPs in (7) and (9) is enough for listeners to construct the NP_c. If this were true, however, then we would expect that overt coordination marking does not make the listener's task any easier, and so according to MiF, asyndeton should be the most common strategy across the world's languages, but this is not the case (but see below for further discussion of the historical origin of coordination).

Several patterns are apparent in the data. First, there is a strong preference for (7) over (8): in (8), the initial coordinator allows the assignment of two more properties after the second word and three more after the third word. Second, there is a very slight preference for (10) over (9): in (10), the initial coordinator allows the assignment of one additional property after the second word. MaOP, therefore, reveals a preference in performance for one of the coordination strategies. Interestingly, this preference is the opposite of Stassen's observed universal: MaOP prefers initial monosyndeton marking to final monosyndeton marking, either strongly or weakly depending on the basic word order of the language. Note also that this preference only increases if there are more than two coordinated elements: if we extend the tables by adding more NPs, this creates more columns in which the number of assigned properties is higher in the initial monosyndeton examples.

5 Analysis

Having applied Hawkins' three processing principles, therefore, we have come to a surprising conclusion: the principles either do not prefer one strategy over the other (MiD, MiF), or else prefer the unattested strategy to the attested one (MaOP). We cannot therefore account for Stassen's universal using Hawkins' processing principles—in fact, it appears to be a counterexample to them, because an increase in the ease of processing of a structure is associated with a *decrease* in the frequency of occurrence of that structure cross-linguistically. Clearly, we have to look elsewhere for an explanation of Stassen's universal, but where? In this section, I suggest two possible alternative explanations.

The first is, like Hawkins' principles, based on the PGCH. Recall that the PGCH attempts to correlate the frequency of language features with "their degree of preference in performance" (Hawkins 2004:3). I suggest that this principle includes a broader range of processing than what is described by MaOP—the method of calculating the OP/UP ratio is based on operations are be performed by the hearer rather than by the speaker. Perhaps ease of production should also be taken into account when evaluating the degree of preference in performance.

Considering production has two effects on a MaOP analysis above. Initial monosyndeton is no longer preferred over final monosyndeton, because speakers can construct a coordinated NP_c immediately upon beginning to pronounce it, unlike hearers, who must wait until a marker of coordination occurs. Second, production considerations may actually favor final monosyndeton. Coordination allows the inclusion of arbitrarily many phrases of a given type in a position usually occupied by a single phrase of that type. When speakers are constructing coordinated structures, especially those that coordinate longer utterances (e.g. sentences), it requires less working memory if they can decide to add another coordinand as an "afterthought". If coordination were marked initially, speakers would need to know before the first coordinand whether another will follow, but since it is, in fact, marked medially or finally, speakers can delay making this decision until the following coordinand. In other words, final monosyndeton allows the speakers to leave their options open and coordinate as an afterthought, rather than having to plan it out beforehand.

The second alternative explanation has to do with the origin of coordinate structures. Mithun (1988) discusses the various origins of coordinate constructions in the world's languages. She suggests that speakers of languages that have coordination only by juxtaposition often develop overtly marked coordination when they come into contact with speakers of languages that have overtly marked coordination, or with written language. Coordination by juxtaposition, she argues, is actually coordination marked by a special intonation contour, the "comma intonation" (Mithun 1988:332), and while this strategy is sufficient in spoken language, to be unambiguous in written language, coordination must be marked somehow. Often, the sources of the newly grammaticized coordinators are comitative markers on noun phrases (the Comitative Strategy of Stassen's WITH-languages).

Let us suppose that coordination is only recently grammaticized in many languages, that comitative marking is often its source, and that, in addition, obliquely marked noun phrases tend to follow the subjects of sentences. If so, we would expect the structures that precede coordinate structures to consist of one NP followed by another NP marked as comitative. Depending on whether the language is head-initial or head-final, this produces one of two patterns:

(13) NP NP-WITH

(14) NP WITH-NP

If structures like (13) and (14) undergo reanalysis into balanced syntactic coordination and the WITH-language becomes an AND-language, the result will be coordinate structures with the word orders attested in Stassen's survey, but not the unattested AND-NP NP order.

6   Conclusion

In this paper, I have attempted to account for a language universal observed by Stassen (2000) using principles of processing proposed by Hawkins (2004), and found that they do not adequately explain it. I have offered two alternative explanations. The first is based on the idea of taking another aspect of performance, namely production,

into account when analyzing language structures. Taking production more formally into account seems like a reasonable extension of Hawkins' ideas, but much more work, including the analysis of many more typological generalizations, would need to be done before a new production-based principle could be proposed. The second explanation is based on work on the grammaticization of coordination in Mithun (1988), and it suggests that the unattested coordination structure may be lacking because of the historical origin of coordination structures. This explanation may work for languages in which coordination is a recent development, but if Hawkins' performance-based explanation is correct, why do we still see *no* examples of AND-NP NP structures, even after, in some cases, thousands of years of language change after the grammaticization of coordination? As always in linguistic typology, there remains more work to be done, both in the collection and analysis of language data and in the formulation of theories to account for variation among the world's languages.

References

Hawkins, John A. 2004. *Efficiency and complexity in grammars*. Oxford: Oxford University Press.

Mithun, Marianne. 1988. The grammaticization of coordination. *Clause combining in grammar and discourse*, Ed. John Haiman and Sandra A. Thompson. Amsterdam: Benjamins. 331–359.

Stassen, Leon. 2000. AND-languages and WITH-languages. *Linguistic Typology*, 4: 1–54.

How UG-Provided Conceptual Structure Restricts the Possibilities for Quantification in Natural Language

David Goss-Grubbs

davidgg@u.washington.edu

1   Introduction

A common way of understanding quantification in natural language is as relations between sets. But it seems that only a highly constrained subset of such possible relations is actually grammaticized in the world's languages. In this paper I propose that at the level of Conceptual Structure, there are very few primitive quantifying relations, and that the narrow range of quantification attested in natural languages follows from the properties of those relations.

In the first section, I review the relational approach to quantification, and how the quantification attested in natural languages is much more constrained than what the theory would potentially allow for. Then, I introduce the Conceptual Structure relations AbsQuant and RelQuant, and show how they account for (at least a large fragment of) the quantifying determiners of English.
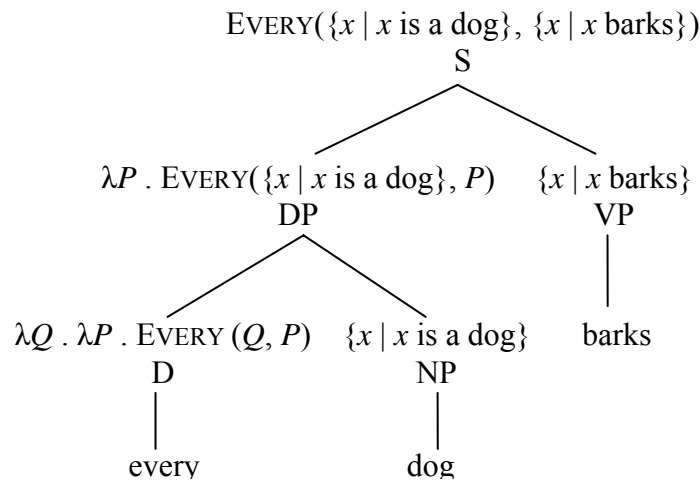
In the next section, I use these Conceptual Structure relations to account for various phenomena found cross-linguistically. I offer an explanation for the markedness of distributive-key universals, noted in Gil (1995); the violation of the principle of Quantity by the Dutch determiner *sommige*, noted in de Hoop (1995); and the ambiguity of noun phrases in Warlpiri noted in Bittner and Hale (1995).

2   The Relational Approach to Quantification

The denotations of determiners (in a language like English) can be viewed as relations between sets. The denotation of a sentence like (1a) is built up compositionally as shown in (1b).

(1)   a.   Every dog barks.

   b.
$$\text{EVERY}(\{x \mid x \text{ is a dog}\}, \{x \mid x \text{ barks}\})$$

S

$\lambda P . \text{EVERY}(\{x \mid x \text{ is a dog}\}, P)$ $\{x \mid x \text{ barks}\}$
DP   VP

$\lambda Q . \lambda P . \text{EVERY}(Q, P)$ $\{x \mid x \text{ is a dog}\}$ barks
D   NP

every   dog

A pair of sets is in the relation EVERY just in case the first set is a subset of the second set. Thus, (1a) is true just in case the set of dogs is a subset of the set of barkers, which is indeed what the sentence seems to mean.

What does a relation between sets look like? In a universe that contains exactly two individuals, *a* and *b*, the extension of EVERY is shown in (2):

(2) $\{\langle\{ \}, \{ \}\rangle, \langle\{ \}, \{a\}\rangle, \langle\{ \}, \{b\}\rangle, \langle\{ \}, \{a, b\}\rangle,$
      $\langle\{a\}, \{a\}\rangle, \langle\{a\}, \{a, b\}\rangle,$
      $\langle\{b\}, \{b\}\rangle, \langle\{b\}, \{a, b\}\rangle,$
      $\langle\{a, b\}, \{a, b\}\rangle\}$

It is a set of pairs of sets. As long as the first set in the pair is a subset of the second set, that pair will be in the extension of EVERY.

As another example, the extension of SOME is shown in (3):

(3) $\{\langle\{a\}, \{a\}\rangle, \langle\{a\}, \{a, b\}\rangle,$
      $\langle\{b\}, \{b\}\rangle, \langle\{b\}, \{a, b\}\rangle,$
      $\langle\{a, b\}, \{a\}\rangle, \langle\{a, b\}, \{b\}\rangle, \langle\{a, b\}, \{a, b\}\rangle\}$

As long as the two sets have a non-empty intersection, the pair is in the denotation of SOME. Thus, a sentence like (4) will be true just in case the set of dogs and the set of barkers have a non-empty intersection:

(4) Some dog barks.

There are quite a few logically possible determiner meanings of this type. In a universe containing *N* individuals, there are 2^N possible sets of individuals. Thus there are $(2^N)^2$ possible pairs of such sets. Finally, that means there are $2^{(2^N)^2}$ possible sets of such pairs. In other words, in a universe containing just 2 individuals, there are 2^{16} or 65,536 logically possible determiner denotations. Add one more individual, and it increases to 2^{64} possible sets of pairs of sets.

Of course, in reality we only find quite a small number of quantifiers. It has often been observed (e.g. de Swart 1998) that the determiners that actually show up universally (or near-universally) exhibit these properties:

Conservativity: A determiner meaning *DET* is conservative if, for any two sets *A* and *B*, *DET(A, B)* is true whenever *DET(A, A $\cap$ B)* is true. That is, for a conservative determiner, we don't care about things in *B* that aren't also in *A*.

Extension: A determiner meaning *DET* exhibits extension if, for any two sets *A* and *B*, if *DET(A, B)* is true in one model, then it will be true in any model with identical sets *A* and *B*. That is, for determiners that exhibit extension, we don't care about things that are in neither *A* nor *B*.

Quantity: A determiner meaning *DET* exhibits quantity if, for any two sets *A* and *B*, *DET(A, B)* is true in one permutation of the universe whenever it is true in any other permutation. That is, for determiners that exhibit quantity, we don't care *which* things are in the sets, we only care how many.

So for any conservative determiner *DET* that exhibits extension and quantity, the truth value of a sentence like (5) will depend only on the number of dogs and the number of barking dogs.

(5) *Det* dog(s) bark(s).

Barwise and Cooper (1981) propose universals to the effect that all languages have essentially quantificational NPs; and that they all have determiners whose denotations are as I have just described them.

As far as I can tell, these universals are just descriptions of what we observe in human language, but are not explained by anything in UG.

3   Accounting for the Apparent Universals

My hypothesis is that relations like EVERY and SOME are not primitives at the level of Conceptual Structure. Rather, UG provides two very general relations, from which can be built up all and only the sorts of quantifying relations that we actually find in the world's languages. These relations are listed in (6).

(6)   a.   ABSQUANT(A, B, N)

b.   RELQUANT(A, B, X)

The ABSQUANT relation (for "absolute" quantification) is a relation between three things: a set A, a set B, and an integer N. The relation holds whenever the intersection of A and B contains N members. This relation handles the so-called weak[1] determiners, like *some, many*, and the numerals. So at the level of Conceptual Structure, the representation for (4) would be (7).

(7)   ABSQUANT(A, B, N) $\wedge$ GREATER(N, 0) $\wedge$ DOG(A) $\wedge$ BARK(B)

The sentence is true if the first set (the set of dogs) and the second set (the set of barkers) have at least one element in common.

The RELQUANT relation (for "relative" quantification) is also a relation between three things: a set A, a set B, and a real number X between 0 and 1. This relation holds whenever the cardinality of $A \cap B$ divided by the cardinality of A equals X. This relation handles the so-called strong determiners, like *every* and *most*. At the level of Conceptual Structure, the (a) sentences are true whenever the conceptual structures in their (b) counterparts are true.

(8)   a.   Every dog barks

b.   RELQUANT(A, B, 1) $\wedge$ DOG(A) $\wedge$ BARK(B)

---

[1] The interaction between ABSQUANT, RELQUANT and the judgments that give rise to weak vs. strong determiners is an interesting question, but is beyond the scope of this paper.

(9)   a.   Most dogs bark

     b. $\textsc{RelQuant}(A, B, N) \wedge \textsc{greater}(N, 0.5) \wedge \textsc{Dog}(A) \wedge \textsc{Bark}(B)$

Sentence (8) is true if the number of barking dogs divided by the number of dogs is 1. That is, it is true whenever every dog is in the set of barking dogs. Sentence (9) is true if the number of barking dogs is more than half the number dogs generally.

Other determiners require both these relations in order to get their denotations. I propose the (a) sentences in (10) - (12) are associated with the (b) conceptual structures.

(10)   a.   The dog barks.

     b   *content*: $\textsc{RelQuant}(A, B, 1) \wedge \textsc{Dog}(A) \wedge \textsc{Bark}(B)$

        presupposition: $\textsc{AbsQuant}(A, E, 1)$

(11)   a.   The dogs bark.

     b   *content*: $\textsc{RelQuant}(A, B, 1) \wedge \textsc{Dog}(A) \wedge \textsc{Bark}(B)$

        presupposition: $\textsc{AbsQuant}(A, E, N) \wedge \textsc{greater}(N, 1)$

(12)   a.   Both dogs bark.

     b.   *content*: $\textsc{RelQuant}(A, B, 1) \wedge \textsc{Dog}(A) \wedge \textsc{Bark}(B)$

        presupposition: $\textsc{AbsQuant}(A, E, 2)$

That is, definites are the same, at the level of Conceptual Structure, as universals with the extra presupposition of an absolute quantification, where the second set is *E*, the set of all individuals. In effect, the presupposition is just specifying the cardinality of the first set. Singular *the* says there is one thing in the set; plural *the* says there is more than one thing in the set, and *both* says that there are exactly two things in the set.

Note that both $\textsc{AbsQuant}$ and $\textsc{RelQuant}$ are conservative. That is, they don't care about things in the second set that are not in the first set. In fact, $\textsc{AbsQuant}$ doesn't even care about things in the first set that are not in the second set. They both exhibit extension in that properties of things outside of either set cannot have any effect on whether those relations hold. They both exhibit quantity in that they are comparing only the cardinalities of sets.

In the rest of the paper, I will use this analysis of absolute and relative quantification to account for various phenomena in the quantification of natural langauges.

4   Distributive Determiners

Gil (1995) proposes a two-way distinction among types of universal quantification. They are *simple* universals and *distributive-key* universals.[2] An example from English is simple *all* versus distributive-key *every*, as seen in (13) - (14).

(13) a. All men gathered at dawn.
      b. * Every man gathered at dawn.

(14) a. All men carried three suitcases.
      b. Every man carried three suitcases.

Simple determiners can take a variety of scope relations; distributive-key determiners require distributive readings. So (13a) is fine with the collective predicate *gather*, where (13b) is bad. (14a) allows for the possibility that the men carried three suitcases each or that they carried a total of three suitcases between them. (14b) requires that they carried three each.

Many languages make the same distinction. Gil lists examples from Georgian, Tagalog, Russian, Turkish, Lezgian, and Mandarin.

Gil argues persuasively for the position that simple quantifiers are primitive, and that distributive-key quantifiers are portmanteaux which combine a simple quantifier plus some kind of additional information. Thus, distributive-key quantifiers are marked, and will be found only in languages that also have simple quantifiers.

Gil offers a number of kinds of evidence to show that simple quantifiers are basic, and distributive-key quantifiers are marked. Non-distributive readings are preferred even when distributive readings are available. In (15), the most natural reading is that the two men carried three suitcases between them (perhaps one man carried two suitcases, and the other man carried one), rather than two men each carrying three suitcases, or there being three suitcases that each of the two men carried.

---

[2] He also talks about *distributive-share* universals, but space limitations prevent me from giving them a treatment here.

(15)   Two men carried three suitcases.

This shows that there is something less natural (more marked) about the distributive reading than the non-distributive reading.

Languages often have constructions that elaborate on simple quantifiers to create a distributive reading. For instance, English (16) uses the word *apiece* to enforce a distributive reading. Maricopa (17) marks the verb with the suffix *-xper*. Turkish (18) marks the numeral in the narrow-scope NP with a suffix *-er*. Tagalog (19) uses an additional distributive-key quantifier within the narrow-scope NP. All examples are Gil's.

(16)   Two men carried three suitcases apiece

(17)   ʔipač   xvikk   ʔii   xmokm   paayperšík

man   Two   stick   three   carry-DIST.SHARE-DUAL-REAL

'Two men carried three sticks apiece'

(18)   Iki   adam   üçer   bavul   tasıdı

two   man   three-DIST.SHARE   suitcase   carry

'Two men carried three suitcases apiece'

(19)   Nagdala   ng   bawat   tatlong   maleta   ang   dalawang   lalaki

carry   DIR   all-DIST.KEY   three   suitcase   TOP   two   man

'Two men carried three suitcases apiece'

Again, this shows that distributive quantification is marked.

The distributive-key universal quantifiers typically only appear with count nouns, whereas the simple universals can appear with count nouns or mass nouns. For instance, English *all* can appear with any word that *every* can appear with, as well as with mass nouns, which *every* cannot appear with. The more restricted environment that distributive-key universals can appear in supports the idea that they are marked.

Gil proposes another universal: that distributive-key quantifiers are all universal quantifiers. He gives these English examples.

(20)   Two/some/many men $\begin{cases} \text{gathered at dawn.} \\ \text{carried three suitcases.} \end{cases}$

These determiners work fine with the collective predicate *gather*, and they have quite natural non-distributive readings. Furthermore, English contains no distributive-key counterparts for these determiners. He supports this putative universal with further data from Russian, Turkish, Georgian, Punjabi, Tagalog, and Mandarin.

Gil doesn't try to account for why these patterns might hold. I propose that scope relations are, by default, not explicitly represented at the level of Conceptual Structure. Instead, the normal situation is that the algorithm that determines truth-conditional interpretations from conceptual structures must make choices with respect to scope that are underdetermined by the conceptual structure itself.

For instance, sentences with simple universals, such as English (14a) would get a conceptual structure something like this[34]:

(21)   CARRY(A, B) $\wedge$

RELQUANT(C, A, 1) $\wedge$

MAN(C) $\wedge$

ABSQUANT(D, B, 3) $\wedge$

SUITCASE(D)

The algorithm for determining truth-conditional interpretations is free to take either quantificational relation as primary or to take them as equal, with the default being to take them as equal. The three options correspond to these three readings:

---

[3] To be complete, I need to spell out in detail the algorithm for how to get from structures like (21) to truth values. This is beyond the scope of this paper.

[4] The representation of 'bare plural' noun phrases is an interesting question, but is not addressed in this paper.

(22) a. RELQUANT($\{x \mid x$ is a man$\}$, $\{y \mid$ ABSQUANT($\{x \mid x$ is a suitcase$\}$, $\{z \mid y$ carried $z\}$, 3)$\}$, 1)

b. ABSQUANT($\{x \mid x$ is a suitcase$\}$, $\{y \mid$ RELQUANT($\{x \mid x$ is a man$\}$, $\{z \mid z$ carried $y\}$, 1)$\}$, 3)

c. RELQUANT($\{x \mid x$ is a man$\}$, $\{x \mid x$ carried suitcases$\}$, 1) $\wedge$ ABSQUANT($\{x \mid x$ is a suitcase$\}$, $\{x \mid$ men carried $x\}$, 3)

On the other hand, sentences with distributive-key universals, like English (14b), would get the same conceptual structure but with the additional information that specifies a particular scoping.[5] The reading that this additional information requires is (22a). So a language is free to have simple universal quantifiers that do not specify this additional information, but if they have structures with this additional information, they are sure to have structures without it.

I still have to answer why distributive-key universals cause the algorithm to make the choices that it does. For instance, it is only when the universally quantified NP is in the subject position that it demands wide scope.

(23) a. Every man carried three suitcases
b. Three suitcases were carried by every man

In (23a), *every man* demands wide scope. In (23b), either scope is available. In any case, one must take scope over the other. We cannot get the three-suitcases-between-them reading.

It also appears as if I may have a problem accounting for why distributive-key determiners are always universals. But really it seems like English *most* is also distributive-key. In (24), *most men* strongly prefers wide scope. At the least, the equal-scope reading is unavailable.

(24) Most men carried three suitcases.

---

[5] For the purposes of this paper, it's not important how this information is expressed.

So it looks like the scope-forcing information can live on any RELQUANT determiner, not just the universals.

5 Dutch *Sommige*

De Hoop (1995) examines the Dutch determiner *sommige*, which is often glossed as English 'some (of the)' or 'certain'. Sentence (25) means that there is a set of unicorns, characterized by some quality, known to the speaker but not necessarily to the hearer, and that the unicorns in that set are white.

(25) Sommige   eenhoorns   zijn   wit.
      some   unicorns   are   white.
      'Certain unicorns are white.'

This differs from Dutch *enkele*, which is the plain existential quantifier. Sentence (26) merely means that the number of unicorns that are white is greater than zero.

(26) Enkele   eenhoorns   zijn   wit.
      some   unicorns   are   white.
      'Some unicorns are white.'

Evidence that *sommige* is truly different from *enkele* includes the following pairs of sentences.

(27) a.   Er   bestaan   enkele   witte   eenhoorns
           there   exist   some   white   unicorns
           'There exist some white unicorns'

     b. * Er   bestaan   sommige   witte   eenhoorns.
           there   exist   some   white   unicorns

(28) a. Ik heb gisteren enkele kilometers gereisd

          I have yesterday some kilometers traveled

          'I traveled some kilometers yesterday'

     b. * Ik heb gisteren sommige kilometers gereisd

             I have yesterday some kilometers traveled

(27) shows that while *enkele* can appear in *there*-sentences, which normally allow only NPs with weak determiners, *sommige* cannot. In (28), the idea is that traveling takes a contiguous sequence of kilometers. It takes a phrase indicating some distance, here measured in kilometers. It does not take (every member of) some subset of the kilometers.

So *sommige* appears not to exhibit the property of Quantity. The truth of (25) depends not just on the number of unicorns and the number of white unicorns. Rather, its truth depends on *which* unicorns are white. How does my proposed system of absolute and relative quantification deal with *sommige*?

I propose the conceptual structure in (29a) for sentence (25), with the semantic interpretation in (29b).

(29) a. RELQUANT(A, B, 1) $\wedge$

         UNICORN(A) $\wedge$

         SALIENT-REL(A) $\wedge$

         WHITE(B)

     b. $\{x \mid x \text{ has the salient quality}\} \cap \{x \mid x \text{ is a unicorn}\} \subseteq \{x \mid x \text{ is white}\}$

That is, *sommige* not only introduces a RELQUANT, (like English *all* does), but it also introduces some contextually salient predicating relation. So the actual quantification is still accomplished through means that exhibit the property of quantity, it is just that there is an extra predication thrown into the mix as well.

So my proposal doesn't rule out determiners that don't exhibit quantity, but it does predict that they should be less common than those that do. That is, by default, a determiner will only introduce a quantifying relation (ABSQUANT or RELQUANT) or other

functional relation (e.g. GREATERTHAN) into conceptual structure. But some determiners may, in addition, introduce a predicational relation as well.

6   NP Ambiguity in Warlpiri

Bittner and Hale (1995) argue that Warlpiri has just two major syntactic categories: Noun and Verb. These two classes of words are easily distinguished from one another on the basis of morphology. The main predicate in a sentence may be expressed either by a noun or a verb. Verbs are primarily active and nouns are primarily stative. Nouns can also serve as arguments of predicates, in which case they exhibit pronominal agreement, or as secondary predicates, in which case they exhibit adjective-like agreement.

Bittner and Hale list the following uses for Warlpiri nominals, in order from most argument-like to most predicate-like:

(30) a. Pronouns, demonstratives and other indexicals
  b. Names
  c. Common nouns
  d. Expressions of quality or cardinality
  e. Expressions of psychological states
  f. Locatives and directionals

An overt NP can be a single noun, or can be constructed by putting together a head noun and one or more modifiers, as in (31). The elements of the NP need not be contiguous.

(31) Maliki   wiri-ngki
     dog   big-ERG
     'a/the big dog'

The syntax is the same, no matter what sort of nominals are used. So the single expression of cardinality *jirrima* can be a noun phrase meaning 'two (of them)' or 'the two (of them)'. Determiners do not exist as a separate syntactic category. The noun phrase (31) can either get the weak reading 'a big dog' or the strong reading 'the big dog'.

Bittner and Hale argue that this ambiguity carries over to nouns which are expressions of cardinality. For instance, the word *panu* can be either the weak 'many' or the strong 'all (of them)'. The same sort of syntactic devices are used to narrow down the choice between 'a dog' and 'the dog' as are used to narrow down the choice between *panu* 'many' and 'all' or the choice between *jirrima* 'two' and 'both'.

The weak reading can be forced with the suffix *-kari*. Example (32) shows this for common nouns. Example (33) shows this for expressions of cardinality. All examples here are Bittner and Hale's.

(32) Jarntu-**kari** ∅-∅ parnka-ja yatijarra, jarntu-**kari** kurlirra.

dog-**KARI** PRF-3s run-PST north, dog-**KARI** south

'A dog ran north, a dog (ran) south.'

(33) Panu-**kari** ka-rna-jana nya-nyi panu-**kari** ∅-li wurulyya-nu.

many-**KARI** PRS-1s-3p see-NPST many-**KARI** PRF-3p hide-PST

'I see a large group, (but) a large group went into hiding.'

When a noun appears with an obligatorily definite nominal, such as a demonstrative (in bold below), the strong reading is forced. Example (34) shows this for common nouns. Example (35) shows this for expressions of cardinality.

(34) **Yalumpu**-rra ka-rna-jana pura-mi jarntu

that-PL PRS-1s-3p follow-NPST dog

'I am following those dogs.'

(35) **Yalumpu**-rra ka-rna-jana pura-mi panu

that-PL PRS-1s-3p follow-NPST many

'I am following that large group.'

How does this fit with the system I am proposing here? I assume that in the absence of any explicit expression of cardinality, Warlpiri noun phrases just are supplied an ABSQUANT relation, which gives the weak reading. Under the right circumstances,

universal-force RELQUANT relations can be added as well, which gives the strong reading. So (36a) gets a conceptual structure like (36b).

(36) a. Maliki   wiri-ngki   ka-Ø-ju   wajilipi-nyi
          dog   big-ERG   PRS-3s$_1$-1s$_2$ chase-NPST
          'A/the big dog is chasing me'

      b. ABSQUANT(A, B, 1) $\wedge$
          BIG(A) $\wedge$
          DOG(A) $\wedge$
          CHASE(B, *me*)
          [$\wedge$ RELQUANT(A, B, 1)]

The RELQUANT predication is optional. With it, you get the definite reading. Without it, you get the indefinite reading.

If a noun phrase contains an explicit expression of cardinality, it is used in an ABSQUANT relation. As before, an optional RELQUANT may be added. An example is sentence (37).

(37) a. panu   ka-rna-jana   nya-nyi
          many   PRS-1s-3p   see-NPST
          'I see a/the large group (of them)'

      b. ABSQUANT(A, B, N) $\wedge$
          LARGE(N) $\wedge$
          SEE(*me*, B) $\wedge$
          [$\wedge$ RELQUANT(A, B, 1)]

Again, the RELQUANT part is optional. If it is added, it changes 'a large group' to 'the large group', which is often just glossed as 'all (of them)'.

7   Conclusion

I have proposed that at the level of Conceptual Structure, UG provides just a very few primitive operators that can be used for quantification. The range of quantification that is

possible using them is highly constrained, yet seems to account for the range of quantification that is actually found. I used this system to address various phenomena in quantification cross-linguistically, including the markedness of distributive-key universal quantifiers; determiners that do not exhibit quantity, such as Dutch *sommige*; and the ambiguity of noun phrases in Warlpiri, which do not use determiners for quantification at all.

References

Barwise, Jon and Robin Cooper. 1981. Generalized Quantifiers and Natural Language. *Linguistics and Philosophy*, 4: 159–219.

Bittner, Maria and Kenneth Hale. 1995. Remarks on definiteness in Warlpiri. *Quantification in Natural Languages.* Ed. Emmon Bach, Eloise Jelinek, Angelika Kratzer, and Barbara H. Partee. Dordrecht: Kluwer Academic Publishers, 81–105.

Gil, David. 1995. Universal quantifiers and distributivity. *Quantification in Natural Languages*. Ed. Emmon Bach, Eloise Jelinek, Angelika Kratzer, and Barbara H. Partee. Dordrecht: Kluwer Academic Publishers, 321–362.

de Hoop, Helen. 1995. On the characterization of the weak-strong distinction. *Quantification in Natural Languages*. Ed. Emmon Bach, Eloise Jelinek, Angelika Kratzer, and Barbara H. Partee. Dordrecht: Kluwer Academic Publishers, 421–450.

de Swart, Henriëtte. 1998. *Introduction to Natural Language Semantics*. CSLI Publications, CSLI lecture notes ; no. 80.

A Comparison Between the Development of the Chinese Writing System and Dongba Pictographs[1]

Seaver Johnson Milnor

milnor@u.washington.edu

1   Introduction

Alongside the Chinese writing system, Naxi Dongba pictographs (納西象形文字 *Naxi xiangxing wenzi,* 東巴文字 *Dongba wenzi*) stand in stark contrast.[2] Whereas Chinese is one of few known instances in which writing was invented *ex nihilo*, Dongba pictographs developed in a context of contact with other writing systems, among them Chinese and Tibetan. Yet the Dongba pictographic script does not meet all the criteria that define a writing system proper. We shall see that *Naxi xiangxing wenzi* cannot express the full range of the spoken Naxi language. This is in spite of having had the benefit of cultural contact with complete writing systems. The Dongba characters, however, meet the particular needs for which they were designed.

In the Chinese case, writing was a new invention.[3] William G. Boltz, following the earlier work of Peter A. Boodberg, argues that, as in ancient Egypt and Mesopotamia, the Chinese "invented writing according to what look like general, I am tempted to say universal, principles and patterns" (1994: 12). Likewise, the Mayans appear to have followed the same paradigm in developing hieroglyphics:

---

[1] This paper has benefited greatly from the critique of Dr. Zev Handel and two anonymous reviewers for *University of Washington Working Papers in Linguistics*.

[2] "Naxi," also written "Na-khi," refers both to an ethnic minority group native to Southwest China and their language. "Dongba," also written "dto-mba," refers to the eclectic religion of the Naxi people as well as its ritual specialists.

[3] To state that writing was independently invented in China is conventional, but has not been conclusively proven. Pulleyblank points out, "there were no literate peoples closer to China than the Indus valley from whom the idea of writing could have been transmitted" (1983: 415). See also 414-416; Boltz 1994: 34-38; Cheung 1983: 383.

(a) true writing emerges with logographic signs; (b) the first step toward "phoneticism," that is, phonetic flexibility in the use of graphs, is "rebus" writing, or what we may call "punning;" (c) phonetic complements, i.e., determinatives, arise; and (d) logographs come to be used for their sound value alone, i.e., they are "desemanticized" (Campbell 1984: 12 paraphrased in Boltz 1994: 12).

If this is the process by which writing systems emerge independently, then what are the implications for a society developing its own only after having come in contact with foreign systems? Specifically, to what extent did Dongbas follow the above stages? Let us turn first to the particular contexts from which 漢字 Hanzi (Chinese characters) and *Naxi xiangxing wenzi* arose.

1.1   Chinese Origins

There is some debate surrounding what constitutes the earliest Chinese writing. Some argue that markings on Neolithic pottery shards — unearthed at sites along the Yellow River basin, some predating even the advent of writing in Egypt and Mesopotamia — represent the formative stages of Chinese writing (Boltz 1994: 34-35). In a survey of 20th century excavations, Cheung Kwong-Yue suggests that the graphs found on pottery at two significant early sites, Banpo[4] and Jiangzhai, "allow us to propose a date of circa 4000 B.C. for the commencement of a viable, albeit primitive form of Chinese character" (1983: 383). Overturning a previously stated opinion (quoted in Boltz 1994: 37), famed archaeologist K.C. Chang concedes "that some of the pottery marks of Pan-p'o and Chiang-chai were, individually, directly ancestral to the same characters in the writing systems of the Shang and the Chou," but, nevertheless, maintains that these individual characters did not comprise a writing system (1983: 573). To the contrary, Boltz questions the possibility "that an inchoate attempt at writing would or could remain in a kind of limbo or suspended animations for several millennia before achieving the form of a true writing system" (1994: 38). Among experts, this issue is far from resolved. It may generally be the case that Chinese scholars incline to accept older dates for the

---

[4] One of the most important Chinese archaeological sites, located in modern day Xi'an County, dated approximately 4800 – 4200 B.C., and excavated in 1954 – 1957 (Cheung 1983: 323-325).

beginnings of their native writing system, but Western scholars demand a greater burden of proof. For a definitive answer, we can only await the excavation of further evidence.

1.1.1. Oracle Bone Inscriptions and Old Chinese

Oracle Bone Inscriptions (OBI, 甲骨文 *jiaguwen*) from the 商 Shang period (ca. 16th C. – 1045 B.C.) comprise the earliest Chinese collection of graphs indisputably regarded as a fully-developed writing system. These divinatory inscriptions were carved primarily on the scapulae of oxen and on turtle plastrons (Boltz 1994: 31).[5] Though the connection is rarely apparent at first glance, the characters found on oracle bones are undoubtedly ancestral to the Chinese characters used today.

Of precisely what language then are OBI a written representation? One may reply "Old Chinese;" however, this answer is not without complications.[6] Theories explicating the sound system of Old Chinese, tenuous in their own right, are based largely upon the language of the 詩經 *Shi Jing*, a heterogeneous collection of 305 poems dating ca. 800 – 500 B.C. Besides the centuries separating late-Shang OBI from the earliest *Shi Jing* poems, it may even be the case that they are unrelated languages. Though Shang characters are certainly ancestral to later Chinese *writing*, the spoken language written on OBI may well not have been ancestral to the Chinese spoken during the 周 Zhou (1045 – 221 B.C.)

A few key characteristics of Old Chinese are as follows: Unlike modern dialects, it is believed to have lacked tones but contained consonant clusters; consonant endings, which affected the pitch of words, are believed to be the source of Middle Chinese (ca. 600 A.D.) tones. Measure words (MW), derived from nouns, were not obligatory, but occasionally — as seen in OBI — appeared in phrases NOUN + NUMBER + MW. Though SVO (Subject Verb Object) word order is most common in Old Chinese, there is evidence suggesting that the underlying word order may have been SOV (Subject Object Verb) in origin (Handel 2004: 110-112).

---

[5] Interestingly enough, Dongbas are one of the few groups in the world who still practice divination using bones. See Ge 1999.

[6] One possible periodization of Old Chinese is as follows: Early, 1300 – 1100 B.C.; Middle, 1100 – 200 B.C.; and Late, 200 B.C. – 200 A.D., roughly coinciding with the Han dynasty (Handel 2004: 93).

1.2 The Naxi Context

The Naxi are one of fifty-five "minority nationalities" (少數民族 *shaoshu minzu*) recognized by the People's Republic of China.[7] Their present population of about 289,000 is largely situated in the mountainous Lijiang Naxi Autonomous Region of Yunnan province (Zhang 2000: 62). The Naxi language is a member of the Yi (a.k.a. "Loloish") branch of the Tibeto-Burman language family (Ramsey 1987: 249-250). Though Naxi is divided into two dialects, western (e.g. Lijiang) and eastern, the latter is more heterogeneous and internally less mutually intelligible. The Lijiang dialect has forty-eight consonants, nine vowels, four tones and "syntactic structure ... much the same as that of other Tibeto-Burman languages spoken in Yunnan" (266).

Writing among the Naxi is particularly interesting. Besides writing *putonghua* (普通話) with Chinese characters, they have two scripts for their own language, one phonetic and the other pictographic. Both forms of Naxi script were used in production of Dongba ritual texts. Sources disagree whether the pictographs preceded the phonetic script, or appeared later (Jackson 1979: 53).[8] I believe the most likely explanation is that given by Anthony Jackson: a phonetic script, related to that of the Yi people, emerged in the 13th century when both groups were under Mongol rule (60-61). If this was in fact the case, then the phonetic script certainly predates the pictographs.[9]

The pictographic script became ubiquitous throughout Naxi territory during the 18th and 19th centuries and was surprisingly standardized; the phonetic script, however, was more idiosyncratic and less uniform across locales. "The phonetic script was not used as the main vehicle for the ritual texts but was generally employed for spells (where the sound alone was important) and for books of divination (... as a shorthand device for colloquial Na-khi)" (Jackson 1979: 60). The Naxi phonetic script was imperfect in that it lacks diacritic marks to indicate tone — thus, as with Mandarin written in toneless *pinyin*,

---

[7] Another name often applied to Naxi people is "Moso." Resolving the Naxi/Moso distinction is an interesting question, but beyond the scope of this paper. My own understanding is that the Moso are a subset of the Naxi—reputed for the custom of "walking marriage" (走婚 *zouhun*) and matriarchal family structure — living around Lugu Lake on the Yunnan-Sichuan border. For a detailed discussion, see Jackson 1979: 275-296 and Pan 1995: 84-119.

[8] For a concise summary and appraisal of both arguments, see Pan 1995: 180-186.

[9] I am indebted to Dr. Chas McKhann, Associate Professor of Anthropology at Whitman College and expert on Naxi religion, for bringing Anthony Jackson's work to my attention.

ambiguity easily arises. Dialectical variation of course compounds the problem. Pictographs, however, "being *partly illustrative* ... can employ symbols to convey the ideas which are severally represented by one homophone but in different tones" (emphasis mine; 62).

1.2.1 Dongba Manuscripts

Over 5,000 Dongba manuscripts have been collected in libraries across the United States and Europe. The availability of so many of such texts to the western world is largely due to the efforts of explorer Joseph Francis Rock (1884–1962), a prolific collector and translator who resided in southwest China for the bulk of 1922–1949. Rock's publications, including the translations of approximately 135 Dongba texts, constitute the foundation of western Dongba studies (Pan 1995: 8-9).

From when and where did Dongba pictographs and manuscripts appear? Anthony Jackson tells us that Joseph Rock, relying on a colophon dating a text by the Chinese *tiangan-dizhi* 天干地支 sexagenary cycle, claimed Dongba texts appeared at least as early as the 16th century. However, Jackson convincingly refutes Rock's assertion and proposes circa 1750 as a more plausible date; certainly no extant Dongba text predates 1703 (Jackson 1979: 52). The political and cultural context of the early 18th century complements this interpretation with an impetus for the promulgation of indigenous pictographs. In 1723, the Qing (1644–1911) government tightened its control over minority peoples, but this did not include Manchus, the minority ethnic group comprising the Qing ruling house. Among the traumatic cultural consequences, forcing the Chinese custom of arranged marriage upon the Naxi resulted in an increased suicide rate. As Lijiang became a center of trade, the standard of living increased. With an increase both in social problems and the means to hire ritual specialists to remedy them, the Dongba religion and its textual tradition flourished. However, the greatest upsurge in Dongba text production occurred after 1830. Particularly considering Lijiang's strategic location as a trade route, the opium industry became increasingly lucrative after China's Opium Wars with Britain (1839 – 1842) and the Naxi economy benefited greatly. The population increased along with its disposable income for Dongba ceremonies (Jackson 1979: 54-55, 73; Pan 1995: 156).

Having considered socioeconomic catalysts for the Dongba tradition, now we must consider where their ritual texts came from:

> The Na-khi pictographic script consists of little stylized drawings of men, animals, trees, stones, etc., written across the page from left to right, as in Tibetan. The physical layout of the book with its three or five lines of text, the use of a pen or stylus, and even the making of the paper, all show Tibetan rather than Chinese influence.... The plain conclusion is that the Na-khi dto-mba manuscripts are modeled on Tibetan books (Jackson 1979: 60).

> Looking at all we know of the dto-mbas – their dress, their rites, and their scripts – all point to Bön-inspired sources. [Bön is a Tibetan religion rooted in pre-Buddhist animistic shamanism.] If one takes the Bön sect as an ongoing institution and then progressively strips it of its lamaseries, its temples, its books, and bans its monks from their traditional begging as a means of revenue, proscribes them from gathering together in the main towns and villages, and leaves them for a few years: what results? The answer is plainly evident: a peasant farmer with a fund of esoteric means of coping with demons – a dto-mba (68).

Jackson distills the evidence into three prerequisites for the founding fathers of the Dongba religion: (1) familiarity with Tibetan bookmaking, (2) Bön symbolism, and (3) knowledge of both written Tibetan and spoken Naxi. Such a person would have been a "Na-khi trained at a Bön lamasery."

2   What Constitutes Writing?

William Boltz defines writing "as the graphic representation of speech; and a writing system, then, as any graphic means for the systematic representation of speech" (1994: 17). "Later he says that 'the essential and indisputable feature that must be present for a graph ... to qualify as writing is phonetic representation.' Thus, ... all graphs that are not associated with pronunciation are excluded from writing" (Bottéro 1996: 575). This definition is narrow and certainly rules out graphs that otherwise may be argued to constitute "writing," but remains a convenient standard for analysis.

2.1   Notation

Boltz employs a useful system, which I will refer to as "GPS notation," to describe key attributes of graphs. The abbreviations G, P and S stand respectively for graph, phonetic value and semantic value. The three components together are arranged thus: G : [±P, ±S]. The "plus" (+) or "minus" (-) sign preceding P and S indicate where or not "the feature in question is associated with the graph" (1994: 19). "Plus" of course means that the given feature is present and "minus" that it is not. GPS notation can thus denote four possible types of graphs:[10]

(1) G : [-P, -S] This type of graph, lacking both sound and meaning, is clearly not writing. Examples may include the absentminded doodling of a bored student, scratch marks left on furniture by a misbehaving cat, an otherwise "realistic" drawing or painting so poorly executed that no person other than the artist could recognize the intended subject matter, or any graph lacking a conventionalized meaning—e.g. ▯—▯ , ⊠ , 000000.

(2) G : [-P, +S] Also not writing by Boltz's definition, lacking an associated pronunciation, graphs of this type may include symbols such as the green "Mr. Yuck" poison warning stickers placed on bottles of household cleaners to discourage children from ingesting their contents, the stamp placed upon the back of one's hand as proof of paid admission to an event, and the hexagrams and *umyang* 陰陽 (Chinese: *yinyang*) found on the flag of the Republic of Korea.

(3) G : [+P, +S] Chinese characters are of this type, with the exception of very few which in modern usage have lost their semantic association as well as the rare sub-morphemic characters (e.g. 玻 *bo* and 璃 *li*, which form the word for "glass;" 咖 *ka* and 啡 *fei*, in the transliteration for "coffee.") Boltz, by his interpretation, emphasizes, "the graph stands for the word *only* by virtue of standing for the *sound* of the word in question."

(4) G : [+P, -S] Examples of graphs with phonetic but no semantic association include the letters of the Roman alphabet and *zhuyin fuhao*. Below we will determine the GPS

---

[10] In each case, I have produced my own examples.

classification of Dongba pictographs as well as the status of the script by the standards of a writing system.

2.2   Are Naxi Pictographs Writing?

Though Chinese is certainly a writing system, the status of Naxi pictographs is debatable. *Dongba wenzi* lie right on the cusp between writing and proto-writing. Fang Guoyu and He Zhiwu, among others, tell us that the pictographs are only used among adherents of the Dongba religion, not Naxi people in general. Within their texts, the pictographs do not record every word, but rather serve as memory aids for recitation (1995: *bianyan* 1-2). As rituals became increasingly complex, pictographic mnemonic devices were employed to help the Dongbas remember the proper sequence of chants (Jackson 1979: 62). *Naxi xiangxing wenzi* texts, omitting many words from the rites they record, do not systematically represent speech and thus do not constitute a writing *system* by Boltz's definition. One could learn the spoken Naxi language, memorize the meaning and pronunciation of every pictograph in a given manuscript, and would still be unable to recite the ritual in its entirety without having first studied it under the tutelage of a Dongba.

One may then raise the possibility that it is only the *texts themselves,* in omitting certain words, that do not reflect a writing system and not a feature inherent in the pictographic script. It would be easy enough, for instance, to copy down every other couplet from a famous Tang poem, such as Li Bo's *Song you ren* 送友人 and ask a person familiar with it to recite the complete poem from the partial rendering. Such an exercise, analogous to the production of a Dongba text, would certainly not invalidate the status of Chinese as a writing system.

Disregarding the substantial effort required to write entirely in pictographs, could one not choose to produce a complete transcription of the spoken language with *Dongba wenzi*? Actually, no. "While pictographs are excellent in presenting things, they are a little less helpful in expressing certain non-visual abstract ideas, e.g. ethical doctrines, which may account for their surprising absence from the dto-mba's texts" (Jackson 1979: 62). Other features distinguishing this script from true writing systems are unread symbols "inserted into a frame only to elucidate the meaning of another symbol" and that "at

other times a drawing may be 'read' two or three times even though it appears only once" (Ramsey 1987: 266).

Let us temporarily put aside the requirements of a "writing system" and consider the slightly looser concept of "writing." Though the Dongba pictographic script as a whole does not meet the definition of a writing system, are isolated pictographs—and even a limited set of complete sentences, perhaps—writing? Two types of graphs are defined by Boltz as writing, G : [+P, +S] and G : [+P, -S]. Let us consider two simple entries from *A Glossary of Naxi Pictographs*:
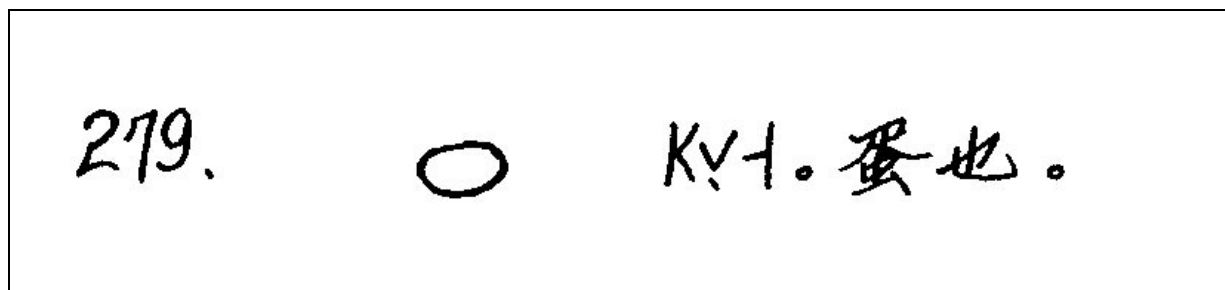
Figure 1: Egg pictograph (Fang and He 1995: 164).
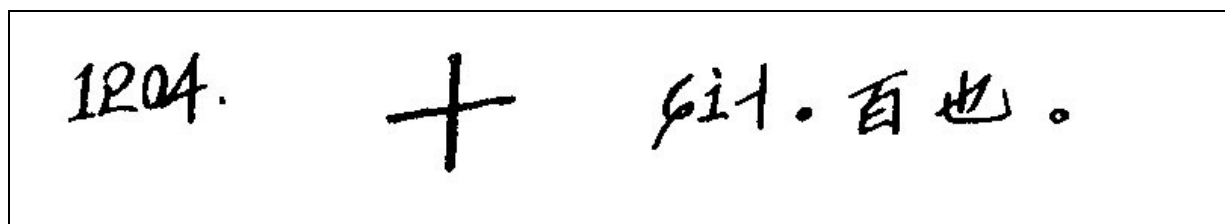
Figure 2: Numeral one hundred (Fang and He 1995: 338).

The pictograph for an egg (#279) is simply an oval.[11] Its phonetic value is represented by International Phonetic Alphabet (IPA) symbols. The diagram "–|" indicates it is pronounced with a mid level tone. The meaning is glossed as "egg" (蛋也 *dan ye*). This pictograph can thus be represented G : [+P, +S], possessing both phonetic and semantic values, and is thus a *written* character. The Dongba numeral one hundred (#1204) resembles the Chinese character 十 *shi* (ten). Likewise a [+P, +S] graph, having both a pronunciation and meaning, it too is an example of writing. Perusing the pages of the *Glossary*, it appears that every graph has an associated pronunciation, i.e. [+P], and

---

[11] The "egg" graph is nearly indistinguishable from some other ovals, such as pictograph #1208. See Fang and He 1995: 339.

accordingly meets our definition of writing. This is admittedly based upon our presumption — I believe a reasonable one — that Fang and He's interpretation of the pictographs is accurate.

Counterintuitive though it may seem, whereas individual Dongba characters are writing, the script as a whole does not comprise a writing system by our chosen definition. Each pictograph represents a spoken word, yet in the aggregate they fail to cover the entire spoken lexicon.[12]

3   Developmental Stages

As mentioned in the introduction, the first stage in the development of writing is the use of logographs, i.e., graphs that stand for words (Boltz 1994: 6). Though no longer obvious in modern characters, Chinese was pictographic in origin. The most intuitive way to write a word was to draw a picture of it. Recognizable drawings, however, are time-consuming to produce, so "there is a natural tendency for such graphs to become progressively simplified and stylized" (Norman 1988: 58-59).

The Dongba pictographic script is clearly indigenous to Naxi areas, because of the particular flora and fauna it represents (Jackson 1979: 59; Ramsey 1987: 268). From this early stage of development, Dongba pictographs diverged from Chinese and the world's other writing systems. Despite their extent of standardization, the pictographs have not been simplified nearly to the extent of any "practical" writing. Quite the opposite of Chinese characters, the meanings of numerous Dongba pictographs are immediately obvious to the untrained observer.[13] Why have Naxi pictographs not been simplified? Like OBI and *jin wen* 金文 (bronze inscriptions), the earliest Chinese characters, Dongba pictographs were used exclusively in ritual texts. Whereas Chinese characters were later applied to daily life, however, Dongba pictographs were not. Dongbas took the time to produce works of art for religious use — efficiency in production speed was not their top pragmatic concern.

---

[12] Consequently, this raises the issue — that will not be pursued here — of Chinese dialectal words for which there are no characters.

[13] Indeed, many characters are so recognizable they are used in modern art, such as the school of 現代東巴畫 *Xiandai Dongba hua* (Modern Dongba Painting), pioneered by Zhang Yunling.

3.1   Logographs

In the absence of a previous concept of writing, a crucial intellectual leap must take place — the realization that a graph can stand for a word, the name of an object, rather than the object itself. It is at this point, once the concept of "word" is realized, that a pictograph [-P, +S] becomes a logograph [+P, +S], regardless of whether the graph's form has evolved into something simple (Boltz 1994: 54). Once a graph describes a word, it obtains a phonetic value from the spoken language. I suspect in the Dongba case, however, as pictographs were devised they instantaneously became logographs, because the written word was not a new concept.[14]

3.2   Rebus Writing

There is a limit to the number of words that can be represented pictographically, as anyone who has played the game "Pictionary" must know. Abstract concepts can be represented to an extent with pictures or diagrams, as seen in the Chinese characters 上 *shang* "above" and 下 *xia* "below." Nevertheless, to fully represent all possible utterances it is necessary to write some abstract words with homophones. Employing these phonetic loans is called the "rebus" principle or paronomasia. A commonly cited Chinese example was using a logograph that pictorially represented "wheat" (麥 *mai*) to also write the homophonous verb "to come" (來 *lai*) (Norman 1988: 60-61; Boltz 1994: 60).[15]

Ramsey (1987: 267) provides three examples of rebus writing using *Naxi xiangxing wenzi*. The word "eye" is a drawing of two eyes; the graph is also used to write the homophonous word "fate." Likewise, a picture of a covered dish denotes both "food" and its homophone "sleep." There is no guarantee, however, that a homophone will be available, so frequently a near-homophone must suffice. The goral (goat antelope) pictograph is used logographically to write an aspect marker that differs in pronunciation only by tone. One may think a paronomastic borrowing, such as the Naxi word "fate," is G : [+P, -S] because its meaning is not related to what the character represents

---

[14] I have deliberately simplified Boltz's argument, finding it unnecessary to distinguish logographs from zodiographs.

[15] The characters provided are modern Chinese equivalents.

pictorially. However, Boltz notates it "G : [+P, +S, +S'] where S' designates a meaning different from S, indicating that the same graph G is used variously for a word pronounced [P] with the meaning [S], or a word with the same ... pronunciation but with the different meaning [S']" (Boltz 1994: 61). This is the consistent, logical interpretation. A series of [+P, -S] graphs would comprise the beginnings of an alphabet, syllabary, etc., whereas characters with usage extended by the rebus principle are still associated with particular words.

3.3 Determinatives

The third stage in the development of writing is "disambiguation." One disadvantage to the essential rebus principle is that it creates ambiguity; one graph is used to represent semantically unrelated words. This ambiguity can be resolved through the addition of a "determinative," also known as a "classifier," or, in the Western Sinological tradition, as a "radical."[16] This added element could be either phonetic or semantic. The latter type was used in the case of otherwise identically written rebus phonetic borrowings—e.g., including the "rain" classifier 雨 *yu* in the character 雲 *yun* "cloud" to distinguish it from 云 *yun* "to say" — and the former, for example, to differentiate "the numerous characters for types of birds" in the Chinese case (Norman 1988: 60). The radical *niao* 鳥 means "bird;" it is the semantic component in the compound graphs *e* 鵝 "goose," *ge* 鴿 "pigeon," *peng* 鵬 "phoenix," *tuo* 鴕 "ostrich," *ya* 鴉 "raven," *et cetera*.

Judging from *A Glossary of Naxi Pictographs*, the same process occurred with the Naxi script. The difficulty is finding entries for all the components of compound graphs so the reader (and author) unfamiliar with the Naxi language can make sense of them:
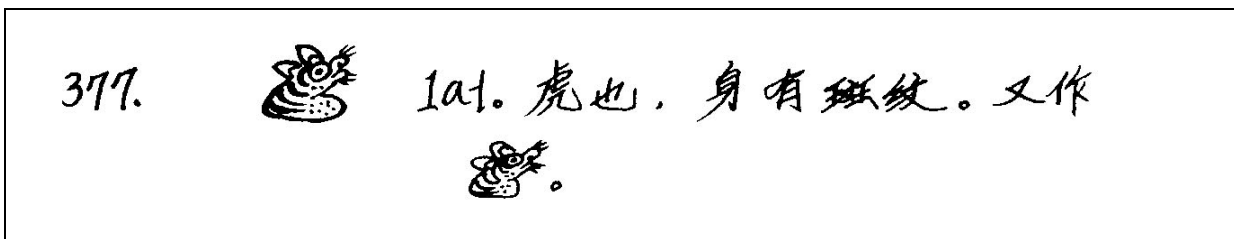

Figure 3: Tiger pictograph (Fang and He 1995: 186).

---

[16] Boltz distinguished "determinative" and "classifier," but we need not (1995: 68).
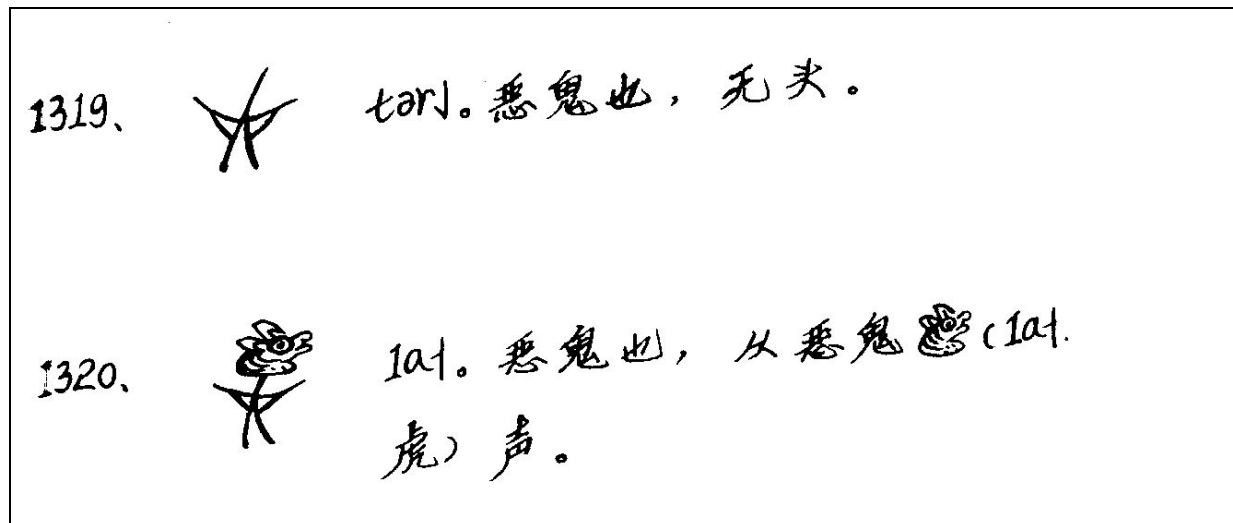
Figure 4: Evil spirits (Fang and He 1995: 359).

(#377) is defined as "tiger." Its "body has striped markings." An allograph is also provided. (#1319) is an "evil ghost, lacking a head." (#1320), also a type of "evil ghost," is homophonous with (#377). The gloss describing its structure tells us that it comes from "evil ghost" (#1319) which is thus the signific, and "tiger" (#377), not surprisingly, is the phonetic.[17]

Summing up these first three stages in the invention of writing systems worldwide, Boltz reaches a broad and exciting conclusion; this is followed by a refined explanation of the synchronicity of these processes in the Chinese case:

What we know or can reasonably infer about the origin and early development of all three great writing systems of antiquity, Egyptian, Mesopotamian, and Chinese, as well as Mayan hieroglyphics in the New World, suggest that up to this point they all evolved stage by stage according to the same basic principles. And in all four cases it is only with the determinative stage that we have a really workable, full-fledged writing system, one capable of transcribing all of the manifold complexities of real speech. The script of the Shang oracle-bone inscriptions includes characters with determinatives, showing very clearly that the writing system had already reached this stage. This is not to say that every character known in subsequent periods of written

---

[17] The format of these character entries is clearly based upon the 說文解字 *Shuowen Jiezi*. The description of the character's structure is analogous to that of a 形聲 *xingsheng* (shape and sound) gloss.

Chinese had arisen and taken its modern form by the Shang dynasty … Thus, while it may be correct to think of individual characters as having passed through these stages sequentially, for the writing system as a whole, it was undoubtedly the case that different characters were being introduced as zodiographs, being used multivalently, and acquiring determinatives all at the same time throughout the formative period of the script (68-69).

Thus, every known instance of writing being created *ex nihilo* followed the same three steps to reach the state of a full-fledged writing system. However, China diverged from the rest of the world in the fourth.

3.4 Desemanticization

Stage four in the development of writing is desemanticization. Graphs' semantic associations are lost, resulting in a purely phonetic writing system [+P, -S]. Desemantization occurred in Mesopotamia and Egypt, but — excluding modern systems such as *zhuyin fuhao* 注音符號 that never replaced *Hanzi* — not in China. As illustrated by Boltz, there were several instances of characters heading in that direction — one character used to represent multiple homophonous words even in cases where distinct characters concurrently existed — but in the end, semantics refused to separate from phonetics. Boltz offers a few reasons why this was the case. The most straightforward is that as Chinese was largely a monosyllabic language, i.e., every syllable had meaning, there was no incentive to write syllables without meaning. Put another way, as there was (with but a negligible number of exceptions) a one to one correspondence among morphemes, syllables, and characters, removing meaning would have been "an intellectual impossibility" (Boltz 1994: 168-177).

In regard to the Naxi Dongba pictographic script, the issue of desemantization is moot, considering: (1) the Naxi already have a phonetic [+P, -S] script; (2) the pictographs are used in religious rather than secular contexts,[18] so the importance of pictographic symbolism in ritual implements usurps any impetus to simplify them; and (3)

---

[18] The full truth of the matter is that Dongba religion in Lijiang now exists primarily for the demand of tourist consumption of ritual performances and souvenirs with pictographic inscriptions.

this, the only known "living" pictographic script — regardless of anyone's desire to the contrary — is nearing extinction.

4 Conclusion

Despite emerging from dissimilar contexts, the Chinese writing system and Dongba pictographs show evidence of the same three universal stages in the development of writing. However, the Naxi pictographic script neither conventionalized to the point that it could function efficiently in secular contexts nor reached the stage of development to be considered a complete writing system. Like Chinese characters, Dongba pictographs individually meet the criteria of writing, though the script as a whole falls just short of constituting a complete writing *system*. Even if the Dongba script were to survive coming decades as more than a tourist's curiosity, I think it unlikely that it would make the minor developmental leap to becoming a full-blown writing system. It arose a number of centuries ago to serve a particular ritual purpose. As its purpose need not expand to the realm of daily use among non-religious specialists — after all, literate Naxi today, as in the past, write in Mandarin Chinese — at most it will but continue to fulfill the needs of demon exorcism, amusing tourists and the like. Still, it is enticing to think that the script is sufficiently developed for a few Dongba priests or scholars to self-consciously expand it to a writing system proper, capable of expressing colloquial Naxi in its entirety, in the space of an afternoon.

References

Boltz, William G. 1993. Shuo wen chieh tzu. *Early Chinese Texts: A Bibliographical Guide*. Ed. Michael Loewe. Berkeley: The Society for the Study of Early China and the Institute of East Asian Studies. 429–442.

Boltz, William G. 1994. *The Origin and Early Development of the Chinese Writing System*. New Haven: American Oriental Society. American Oriental Series, 78.

Bottéro, Françoise. 1996. Review of *The Origin and Early Development of the Chinese Writing System*, William G. Boltz. *Journal of the American Oriental Society*, 116 (3): 574–577.

Campbell, Lyle. 1984. The Implications of Mayan Historical Linguistics for Glyphic Research. *Phoneticism in Mayan Hieroglyphic Writing*. Eds. John S. Justeson and Lyle Campbell. Albany: Institute for Mesoamerican Studies, State University of New York at Albany. 1–16.

Chang, K. C. 1983. Concluding Remarks. *The Origins of Chinese Civilization*. Ed. David N. Keightley. Berkeley: University of California Press. 565–581.

Cheung, Kwong-Yue 1983. Recent Archaeological Evidence Relating to the Origin of Chinese Characters, trans. Noel Barnard. *The Origins of Chinese Civilization*. Ed. David N. Keightley. Berkeley: University of California Press. 323–391.

Fang Guoyu and He Zhiwu (1995). *Naxi Xiangxing Wenzi Pu* [A Glossary of Naxi Pictographs]. Kunming: Yunnan Renmin Chubanshe.

Ge Agan. 1999. *Dongba Gupu Wenhua* [Dongba Bone Prognostication Culture]. Kunming: Yunnan Renmin Chubanshe.

Handel, Zev. 2004. *Course packet for CHIN 342/442 The Chinese Language*. University of Washington, Seattle.

Jackson, Anthony. 1979. *Na-khi Religion: An Analytical Appraisal of the Na-khi Ritual Texts*. The Hague: Mouton Publishers.

Norman, Jerry. 1988. *Chinese*. Cambridge: Cambridge UP.

Pan Anshi. 1995. *Reading Between Cultures: Social Anthropology and the Interpretation of Naxi (Na-khi) Religious Texts*. Ph.D. Dissertation, University of Edinburgh.

Pulleyblank, E.G. 1983. The Chinese and Their Neighbors in Prehistoric and Early Historic Times. *The Origins of Chinese Civilization*. Ed. David N. Keightley. Berkeley: University of California Press. 411–466.

Ramsey, S. Robert. 1987. *The Languages of China*. Princeton: Princeton UP.

Wang Yuanlu. 1988. *Han Gu Wenzi yu Naxi Dongba Wenzi Bijiao Yanjiu* [Ancient Han Character and Naxi Dongba Character Comparative Research]. Shanghai: Huadong Shifan Daxue Chubanshe.

Zhang Songquan. 2000. Lijiang and the Naxi People. *The Yunnan Ethnic Groups and Their Cultures*. Ed. Gan Xuechun, (Yunnan University International Exchange Series, I.) Kunming: Yunnan Renmin Chubanshe. 62–75.

Grammar, Performance, and the *Wh*-Question Typology

Jeff Stevens

jps@u.washington.edu

1   Introduction

What is the division of labor between grammar and performance in determining the character of human language? Are Universal Grammar (UG) and performance preferences in competition to optimally account for the attested phenomena of the world's languages? Or can they play complementary roles in linguistic theory? In this paper, I will argue for the latter position by investigating the *wh*-question typology as defined by Cheng (1991), in order to show how both grammar-internal mechanisms and performance preferences can contribute non-redundantly to particular linguistic phenomena. Two relevant proposals from the literature I will discuss in this paper are Hawkins' (2004) Performance-Grammar Correspondence Hypothesis (PGCH) and Cheng's (1991) Clausal Typing Hypothesis (CTH), both introduced here:

(1)   *Performance-Grammar Correspondence Hypothesis (PGCH)*
Grammars have conventionalized syntactic structures in proportion to their degree of preference in performance, as evidenced by distributional patterns of selection in corpora and by ease of processing in psycholinguistic experiments. (Hawkins 2004, p. 3)

(2)   *Clausal Typing Hypothesis (CTH)*
Every clause needs to be typed. In the case of typing a *wh*-question, either a *wh*-particle in C^0 is used or else fronting of a *wh*-word to the Spec of C^0 is used, thereby typing a clause through C^0 by Spec-head agreement. (Cheng 1991, p. 22)

In addition to the PGCH and the CTH, I will also discuss Miyagawa's (2001) proposal to account for cross-linguistic wh-phenomena by arguing for separate morphosyntactic *wh*- and Q-features as UG elements that differ in their cross-linguistic distribution.

©2006, Jeff Stevens, *University of Washington Working Papers in Linguistics*, vol. 24 (2005), eds. Daniel J. Jinguji and Steven Moran, pp 46-63, Seattle, WA.

With these proposals as background, I will argue for the following hypothesis to account both for the typological distribution of *wh*-question types as defined by the CTH and predicted by the PGCH, and for certain attested typological anomalies which, I will further argue, a UG-based account, such as Miyagawa's, can adequately explain:

(3)   *Question Strategy Determination Hypothesis (QSDH)*
      The strategy choices available to a language for typing a sentence as a question are determined by UG, while the typological distribution of the available strategies is determined by the conventionalization of performance preferences.

The QSDH concerns a specific typological generalization, as expressed by the CTH. Crucially, the strong version of the CTH (which assumes Economy of Derivation (Chomsky 1991) as a UG principle) rules out languages that either employ both Q-particles and *wh*-movement or employ neither of these two strategies for question-typing. In this paper I will discuss apparent exceptions to the CTH with a view towards explaining both why such exceptions exist and why they are typologically rare. Among the exceptions to the CTH that have been cited in the literature are sentences which employ both Q-particles and *wh*-movement, such as the Vata sentence in (4), whose analysis by Koopman (1984) I assume to be correct:

(4)   àlÓ$_i$ Kòfí yÉ t$_i$ yé lá   (Vata)
      who Kofi see PERF Q
      'who did Kofi see'
      (Koopman 1984, p. 35)

While (4) and similar data can be argued to falsify the CTH, my goal in this paper is not to challenge either the CTH or the PGCH, but simply to argue that UG can explain the existence of exceptions (such as (4) and similar data) to generalizations that follow from the CTH and the PGCH.

The structure of this paper is as follows. In section 2, I discuss the PGCH. In section 3, I discuss the CTH and its apparent exceptions. In section 4, I introduce the efficiency principles defined by Hawkins (2004) that follow from the PGCH. In section 5, I discuss the role of word order in the *wh*-question typology. In section 6, I introduce and discuss Miyagawa's (2001) proposal to account for the *wh*-question typology in terms of

morphosyntactic features. In section 7, I defend my own hypothesis, the QSDH. Section 8 is a brief summary with conclusions.

2   Grammar and Performance

According to Newmeyer (in press),"UG tells us what a possible human language is, but not what a probable human language is" (Ch. 3, p. 36). In other words, while a theory positing an innate human language faculty independent of other cognitive faculties may explain the existence of certain grammatical phenomena attested in natural languages, no such theory can fully account for the cross-linguistic abundance or rarity of such phenomena. In response to this explanatory inadequacy of UG, Hawkins (2004) presents a theory of typological generalizations based on the PGCH (repeated below), which, according to Hawkins, achieves explanatory adequacy for such generalizations:

(1)   *Performance-Grammar Correspondence Hypothesis (PGCH)*
       Grammars have conventionalized syntactic structures in proportion to their degree
       of preference in performance, as evidenced by distributional patterns of selection
       in corpora and by ease of processing in psycholinguistic experiments. (Hawkins
       2004, p. 3)

If Hawkins' theory is correct, must it supersede UG-based theories as a means to account for the facts of human language? In what follows I will argue that the task of accounting for the attested phenomena of natural languages — i.e. "possible languages" — is best suited to theories that assume an innate and autonomous UG, while the task of accounting for the cross-linguistic distribution of such phenomena — i.e. "probable languages" — is best suited to the PGCH and similar performance-based theories. I will argue for this position by investigating the *wh*-question typology as presented in Cheng 1991, cast in the light of the PGCH and its predictions.

3   Cheng 1991 and Apparent Exceptions

Cheng (1991), following a suggestion by Chomsky and Lasnik (1977), proposes that clauses must be 'typed' grammatically as declaratives, interrogatives, etc., and that a language must choose one of two strategies for 'typing' *wh*-questions, namely, either a

clause-peripheral Q-particle[1] or leftward *wh*-movement. This proposal is formalized as the Clausal Typing Hypothesis, repeated here:

(2)   *Clausal Typing Hypothesis (CTH)*
      Every clause needs to be typed. In the case of typing a *wh*-question, either a *wh*-particle [i.e. Q-particle — JS] in C^0 is used or else fronting of a *wh*-word to the Spec of C^0 is used, thereby typing a clause through C^0 by Spec-head agreement. (Cheng 1991, p. 22)

Cheng illustrates the CTH with the data in (5) and (6):

(5) $[_{CP}$ Who$_i$ $[_{IP}$ t$_i$ bought what]]?

(6)   Qiaofeng   mai-le   shenme   ne   (Mandarin)
      Qiaofeng   buy-ASP   what   Q$_{WH}$
      'What did Qiaofeng buy?'
      (Cheng 1991, p. 22)

Under Cheng's account, in (5) the pronoun *who* moves to [Spec, C] to type the clause in the scope of CP as interrogative. The pronoun *what* in (5) does not move because the clause is already typed by *who*. In contrast, in (6) the pronoun *shenme* 'what' stays in situ because the Q-particle *ne* (assumed by Cheng to be a head base-generated in C) has already typed the sentence as interrogative, making *wh*-movement unnecessary.[2,3]

    Citing the principle of Economy of Derivation from Chomsky 1991, Cheng argues that the CTH predicts the following:

(7)   No language has yes-no particles (and thus *wh*-particles) and also syntactic *wh*-movement. (Cheng 1991, p. 28)

---

[1] Where Cheng employs the term *wh-particle*, I follow Ultan (1978b) and others in employing the term *Q-particle* for clarity in later sections of this paper. Note also that Cheng distinguishes between *yes-no particles*, which mark yes-no questions, and *wh-particles*, which mark *wh*-questions: languages that employ the former will also employ the latter, although not necessarily vice-versa—a one-way implicational universal. In some languages (Japanese, Korean), but not all (Mandarin), yes-no and *wh*-particles are homophonous. I will restrict my attention to *wh*-questions in the remainder of this paper.

[2] Cheng points out that the Q-particle *ne* is optional, while arguing that *ne* has a non-overt alternate form with the same scopal and quantificational properties as *ne*.

[3] As for multiple-*wh* languages, Cheng argues that in such languages movement of additional *wh*-words is required to license each *wh*-word, and that clausal typing obtains as a secondary consequence of *wh*-movement.

According to Cheng, Economy of Derivation rules out syntactic *wh*-movement where a Q-particle has already typed a clause as interrogative. Thus, it follows from the CTH and (7) that a language exhibiting both overt *wh*-movement and Q-particles is impossible. However, such languages have in fact been attested. Bruening (2004), drawing on Ultan's (1978b) typological survey of interrogative systems in 79 randomly-chosen languages, cites 30 such languages, with varying word orders and variation between initial and final Q-particles: Agta, Albanian, Syrian Arabic, Basque, Burmese, Chontal, Fanti, Finnish, French, Louisiana French, Scottish Gaelic, Gbeya, Grebo, Guarani, Gunwinggu, Hebrew, Hungarian, Irish, Jaqaru, Klamath, Lithuanian, Malagasy, Malay, Ojibwa, Piro, Russian, Squamish, Tagalog, Twi, and Zapotec.[4] In this paper I will focus on relevant data from another such language: Vata, a Kru language spoken in the Ivory Coast whose basic word order is SOV (Koopman 1984). In what follows I will refer to any language that employs both *wh*-movement and Q-particles, regardless of basic word order or Q-clause order, as a *Vata-type language*, the better to compare such languages with *English-type languages* (which employ *wh*-movement without Q-particles) and *Japanese-type languages* (which employ Q-particles without *wh*-movement).

Consider the simple *wh*-question in (4), repeated below, and the embedded clause structure in (8):

(4) àlÓ$_i$ Kòfí yÉ t$_i$ yé   lá   (Vata)
     who Kofi see   PERF Q
     'who did Kofi see'

(8) àlÓ$_i$ n   gūgū nā   Kòfí yÉ t$_i$ yé   lá
     who you think COMP Kofi see   PERF Q
     'who do you think Kofi saw'
     (Koopman 1984, p. 35)

---

4 Bruening cites these languages partly in order to challenge the CTH. Ultan's survey simply claims that these languages employ both "question particles" and sentence-initial *wh*-words, but does not provide supporting data for all of them. I will assume Ultan's survey to be correct while also assuming, contra Bruening, that exceptions to the CTH are typologically rare, having found little data to support Ultan's claims.

Note that, according to Koopman, the *wh*-movement in (4), (8), and many similar Vata examples discussed in Koopman 1984 is obligatory and therefore cannot be analyzed as scrambling. While such data appear to pose a problem for the CTH, I will not argue here that such data are counterexamples to the PGCH, since the PGCH is intended to predict probable languages, not to constrain possible languages. I will instead attempt to show how UG can explain the existence of such data where the PGCH cannot. I will also suggest that such data lend additional support to the PGCH, since their apparent rarity may be due to parsing difficulty compared to the more widely-attested *wh*-question structures predicted by both the CTH and the PGCH, as I will discuss in section 5.

To show how UG can explain such typological exceptions as (4) and (8) as well as the more common 'possible *wh*-questions', I will consider a recent grammar-based proposal to account for the typology of *wh*-questions: namely, Miyagawa's (2001) proposal (following Hagstrom 1998) to account for cross-linguistic *wh*-phenomena by arguing for separate morphosyntactic *wh*- and Q-features as UG elements that differ in their cross-linguistic distribution: morphologically separate in Japanese-type languages, syncretic in English-type languages. Miyagawa also crucially adopts Chomsky's (2000) suggestion that the Extended Projection Principle (EPP) feature requiring overt movement of an XP to the Spec of the EPP's containing head can be generalized from Tense to other functional heads, including C.

According to Miyagawa, the *wh*-feature in *nani* 'what' in (9) below does not raise to establish clausal scope (as in English-type languages) because the accompanying Q-feature (hosted by the Q-particle *no*) has already raised to C to satisfy the EPP on C:

(9)   Taroo-ga nani-o   kat-ta   no?   (Japanese)
       Taro-NOM what-ACC buy-PAST Q
       'What did Taro buy?'
       (Miyagawa 2001, p. 311)

The *wh*-feature determines the indefinite property of *wh*-words (Kuroda 1965) while the Q-feature determines the quantificational and scopal properties of *wh*-questions (Hagstrom 1998). Miyagawa, again following Hagstrom (1998), also argues that the Q-

particle *no* is base-generated right-adjacent to *nani* (as well as other *wh*-words) and is pied-piped to C along with the Q-feature.[5,6]

Miyagawa's proposal appears to reflect the intuitive idea behind the CTH. It also suggests a possible solution to the problem for the CTH presented by the data in (4) and (8). To recapitulate the problem: the CTH predicts that a language employing both *wh*-movement and a Q-particle to type a clause as a *wh*-question should be impossible, whereas the data in (4) and (8) fit this description yet are grammatical in Vata. The possible solution is that the Q- and *wh*-features in Vata can both raise to the C projection, violating Economy of Derivation to fulfill some other requirement of Vata grammar. I will return to this possibility in section 7. In the next section, I discuss Hawkins' (2004) theory of the grammaticalization of performance preferences and its relevance to *wh*-phenomena.

4   Hawkins' Theory and its Explanatory Domain

In addition to the Performance-Grammar Correspondence Hypothesis (PGCH), introduced above in (1), Hawkins (2004) proposes three efficiency principles that follow from the PGCH, all of which are relevant to wh-phenomena, and therefore to the present discussion. These are Minimize Domains (MiD), Minimize Forms (MiF), and Maximize On-line Processing (MaOP), each summarized below:

(10) *Minimize Domains (MiD)*
       The human processor prefers to minimize the connected sequences of linguistic
       forms and their conventionally associated syntactic and semantic properties in
       which relations of combination and/or dependency are processed. The degree of
       this preference is proportional to the number of relations whose domains can be
       minimized in competing sequences or structures, and to the extent of the
       minimization difference in each domain. (Hawkins 2004, p. 32)

MiD appears to explain a significant cross-linguistic generalization involving *wh*-fronting and basic verb position, namely, that *wh*-fronting is more frequent in VSO and SVO

---

[5] The Q-particle *no*, while functioning as a question marker in clause-final position, is generally assumed in the literature on Japanese questions to be a shortened version of *no desu ka* (Hagstrom 1998). *no* is the Japanese genitive marker, which is often used to nominalize a clause; *desu* is the Japanese formal-register copula; *ka* is the Japanese formal-register Q-particle.

[6] Hagstrom presents data from Sinhala, a language of Sri Lanka which is structurally similar to Japanese but with overt Q-particles right-adjacent to *wh*-words at PF, to support his proposal.

languages than in SOV languages. When languages front *wh*-words, they form what is known as a 'filler-gap dependency', defined by Hawkins as a 'filler-gap domain' as in (11):

(11) *Filler-Gap Domain (FGD)*
An FGD consists of the smallest set of terminal and non-terminal nodes dominated by the mother of a filler and on a connected path that must be accessed for gap identification and processing; for subcategorized gaps the path connects the filler to a co-indexed subcategorizor and includes, or is extended to include, any additional arguments of the subcategorizor on which the gap depends for its processing; for non-subcategorized gaps the path connects the filler to the head category that constructs the mother node containing the co-indexed gap; all constituency relations and co-occurrence requirements holding between these nodes belong in the description of the FGD. (Hawkins 2004, p. 175)

According to Hawkins, the increasing size and complexity of FGDs as the distance increases between *wh*-fillers and their gaps (or subcategorizing verbs) accounts for the increasing dispreference for *wh*-movement in verb-final languages compared to verb-initial and verb-medial languages. For example, consider the simple English wh-question in (12):

(12) $[_{CP}$ Who$_i$ $[_{IP}$ t$_i$ greeted Mary]]?

Under Hawkins' definition of an FGD, a gap cannot be identified by the parser until its subcategorizor has been parsed, therefore the verb *greeted* in (12) must also be co-indexed with the filler along with the gap, as in (13):

(13) $[_{CP}$ Who$_i$ $[_{IP}$ t$_i$ greeted$_i$ Mary]]?

Now consider a hypothetical language with SOV order and wh-fronting (call it SOV English), where the counterpart of (13) would be (14):

(14) $[_{CP}$ Who$_i$ $[_{IP}$ t$_i$ Mary greeted$_i$]]?

Comparison of (13) and (14) should reveal the increased complexity of the FGD in (14) compared to that in (13): in the SOV structure in (14), the object intervenes between the filler *Who$_i$* and its subcategorizor *greeted$_i$*, whereas in (13) the path from filler to subcategorizor is less structurally complex and therefore easier to process. I will discuss the correlation between *wh*-movement and basic word order further in section 5.

(15) *Minimize Forms (MiF)*
The human processor prefers to minimize the formal complexity of each linguistic form F (its phoneme, morpheme, word or phrasal units) and the number of forms with unique conventionalized property assignments, thereby assigning more properties to fewer forms. These minimizations apply in proportion to the ease with which a given P can be assigned in processing to a given F. (Hawkins 2004, p. 38)

MiF appears to partially explain the cross-linguistic rarity of Vata-type languages, for reasons involving the redundancy of combining *wh*-movement with Q-particles. I will discuss this matter in section 7.

(16) *Maximize On-line Processing (MaOP)*
The human processor prefers to maximize the set of properties that are assignable to each item X as X is processed, thereby increasing O(n-line) P(roperty) to U(ltimate) P(roperty) ratios. The maximization difference between competing orders and structures will be a function of the number of properties that are misassigned or unassigned to X in a structure/sequence S, compared with the number in an alternative. (Hawkins 2004, p. 51)

MaOP predicts a number of asymmetries, many involving *wh*-phenomena. Crucially, fillers tend to precede gaps in *wh*-questions and relative clauses, as well as other filler-gap constructions. According to Hawkins, MaOP explains these asymmetries along lines proposed by Fodor (1983): When parsing a filler such as a *wh*-phrase in a non-argument position, the hearer is primed to search for a co-referential gap. By contrast, a gap is inaudible and can easily go undetected by the hearer — especially if it precedes its filler in linear order. Crucially, during on-line sentence processing, more properties (categorical, selectional, etc.) are immediately assignable to an overt *wh*-phrase than to a gap.

This appears to explain why the displacement of *wh*-words, in addition to being non-universal, is asymmetric. In almost all languages, *wh*-phrases move to the left and not to the right, i.e. to clause-initial position (as first noted in Bach 1971, p. 160).[7] Hawkins argues that this universal asymmetry can be explained by Fodor's (1983) principle *Fillers First*, which, according to Hawkins, is subsumed under MaOP:

---

[7] At least two exceptional cases have been cited where *wh*-phrases move obligatorily to the right-peripheral position of the clause: Khasi (Mon-Khmer, Austro-Asiatic, SVO), cited by Ultan (1978b), and Tangale (Chadic, Afro-Asiatic, SVO), cited by Kenstowicz (1987).

(17) *Fillers First*
The human processor prefers to process fillers before their co-indexed subcategorizers or gaps. (Hawkins 2004, p. 204)

Hawkins (2004) claims that filler-gap dependencies are generally difficult to process. Why then are they attested at all? On one view (cf. Cheng 1991), in direct *wh*-questions, the matrix C carries interrogative force, and the relevant feature on this C is associated with the *wh*-phrase to form a content question. This association is accomplished in English-type languages by moving the *wh*-phrase into the Spec of the interrogative C, thus satisfying the so-called *Wh-Criterion* (May 1985) at the expense of diminished processing ease. Languages with Q-particles available to satisfy the interrogative feature on C can avoid the processing difficulties that come with filler-gap dependencies by leaving the *wh*-word *in situ*. These observations suggest an interaction between the formal mechanisms of *wh*-fronting and its functional motivation — which brings us to Miyagawa's (2001) proposal, to be discussed in section 6. First, a note on the role of word order in the *wh*-question typology.

5 *Wh*-Questions and Word Order in Typology

Basic word order is relevant to the *wh*-question typology, since there exists a much-discussed correlation between basic verb position and the probability of syntactic *wh*-fronting. According to Dryer (1991), approximately 40% of the world's languages exhibit wh-fronting, while Bruening (2004), based on Dryer's typological database of over 500 languages (described at http://wings.buffalo.edu/soc-sci/linguistics/people/faculty/dryer/dryer/database), claims that between 60 to 70 percent of the world's languages employ question particles, whether with *wh-in-situ* or *wh*-movement. Dryer's data show that VO languages tend strongly to have overt *wh*-movement, while OV languages tend to be *in-situ* languages. As for the word order breakdown, according to Dryer, while 71% of verb-final languages are *in-situ* languages, 42% of SVO languages lack *wh*-movement, while only 16% of verb-initial languages lack *wh*-movement, as shown in Table 1, which also shows the correlation between *wh-in-situ* and final Q-particles:

Table 1
**Proportion of languages with either *wh*-in-situ or final question particles,
by word order type (Dryer 1991)**

	V-final	SVO	V-initial
wh-in-situ	71%	42%	16%
final Q-particles	73%	30%	13%

Greenberg (1963) and Ultan (1978a) provide partially overlapping data on the correlation between *wh*-fronting and basic verb position that serves as a corollary to Table 1. Hawkins (2004) combines these data as in (18), along with similar data provided by Dryer (1991) from Dryer's own genetically and areally controlled sample, presented in terms of genera:

(18) *Wh-fronting and basic verb position*
       Greenberg/Ultan (Hawkins 1999, p. 274)   Dryer (1991)
       V-initial: 17/20 lgs = 85%   23/29 genera = 79%
       SVO: 25/34 lgs = 73.5%   21/52 genera = 40%
       SOV: 7/33 lgs = 21%   26/82 genera = 32%

These typological generalizations are worth considering in terms of the division of labor between UG and performance preferences as expressed by Hawkins' efficiency principles. If one follows Chomsky (1995) in assuming that Merge is less costly for Economy of Derivation than Move, it makes sense to consider Merge of a Q-particle to a clause as a primary strategy for *wh*-question formation cross-linguistically, and *wh*-movement as a 'last resort' when a Q-particle is not available to satisfy the relevant feature in C. As Hawkins (2004) argues in detail, MiD explains why languages tend to prefer Q-particle Merge to *wh*-movement the further their basic verb positions are to the right, since the greater the complexity of the FGD formed by *wh*-movement, the more difficult that FGD will be to process. I believe this lends support to the QSDH, which states that UG determines the possible strategies available for *wh*-question formation, while performance preferences — here, MiD in particular — determine the cross-linguistic distribution of the available strategies. In the next section, I discuss a recent proposal from the literature for how the *wh*-question strategies made available by UG can be precisely formalized in grammatical theory.

6 Miyagawa's (2001) Proposal

Miyagawa (2001) offers a formal account of *wh*-question phenomena which appears useful for typology when cast in the light of the PGCH. Miyagawa follows Hagstrom (1998), who proposes that the Q-particle in Japanese originates within the same constituent as the *wh*-phrase. Under Miyagawa's account, in a Japanese *wh*-question, the Q-particle is raised to C, being attracted by the EPP feature on C. If correct, this analysis unifies Japanese and English *wh*-questions in the sense that they both exhibit overt movement: either to C (as in Japanese) or to [Spec, C] (as in English), either way serving to satisfy the Q-feature on C. This proposal appears to reflect Cheng's (1991) idea that a *wh*-question must be grammatically typed as such, either by *wh*-movement or by a Q-particle located in C (whether by Move or Merge).

According to Miyagawa, in English, C is associated with both the Q- and *wh*-features. Chomsky (2000) suggests that head-to-head movement can satisfy the EPP-feature on the target head. Miyagawa (2001) assumes that the Q-feature is universally on C, though not the *wh*-feature. Under Miyagawa's analysis, in English both the Q-feature and the *wh*-feature occur on the *wh*-phrase and are morphologically inseparable, thus requiring the entire *wh*-phrase to pied-pipe along with the Q-feature to satisfy the EPP on C. In Japanese, by contrast, the two features are morphologically separable and distributed accordingly: when the Q-feature associated with the Q-particle raises to C, the *wh*-feature remains in situ along with the *wh*-phrase, as in (9), repeated here:

(9) Taroo-ga nani-o   kat-ta   no?   (Japanese)
    Taro-NOM what-ACC buy-PAST Q
    'What did Taro buy?'
    (Miyagawa 2001, p. 311)

In further support of his account of *wh-in-situ* in Japanese, Miyagawa, employing Japanese data involving both negation and quantifier phrases, argues that the *wh*-feature in Japanese is on T, not C, as illustrated in (19):

(19) Dare-to$_i$ zen'in-ga   t$_i$ asoba-nakat-ta   no?   (Japanese)
      who-with$_i$ all-NOM   t$_i$ play-NEG-PAST   Q
      'With whom, all did not play?'
      not > all, (all > not)
      (Miyagawa 2001, p. 318)

In (19), the oblique *wh*-phrase *dare-to* 'with whom' moves to T to satisfy both the *wh*-feature and the EPP on T, allowing the subject quantifier *zen'in-ga* 'all' to remain in [Spec, v] and thus be interpreted with narrow scope in relation to negation. Miyagawa's proposal is compatible with Cheng's Clausal Typing Hypothesis, to a degree. In English, one *wh*-phrase must move, either to clause-type the sentence as a question (under Cheng's analysis) or to satisfy the EPP on C (under Miyagawa's analysis). In Japanese, the Q-particle raises (or is Merged) for the same purpose (in both analyses).[8] Thus the attested facts of the *wh*-question typology can be boiled down to morphology, supporting the first clause of the QSDH (repeated below), if Miyagawa's proposal is adopted. But what about the typological distribution of these strategies? Here apparently is where a strict appeal to grammar fails, and performance must be appealed to instead for explanatory adequacy. This leads to the explanatory advantage of the PGCH for typology, supporting the second clause of the QSDH.

(3) *Question Strategy Determination Hypothesis (QSDH)*
     The strategy choices available to a language for typing a sentence as a question
     are determined by UG, while the typological distribution of the available strategies
     is determined by the conventionalization of performance preferences.

7   Defending the QSDH

In this section, I will defend the QSDH by first showing how a grammar-based proposal — namely, Miyagawa's feature-driven proposal discussed in section 6 — can explain the range of attested facts in the *wh*-question typology. I will then show how performance preferences as formalized in the PGCH and its accompanying principles can account for the rarity of Vata-type languages. In earlier sections, I have presented

---

[8] Cheng (1991) mentions the possibility that the Q-particle may originate somewhere below C, but assumes that it is base-generated in C for ease of exposition. My discussion in section 5 of Q-particle Merge in relation to processing complexity suggests the potential depth of the question whether Q-particle Merge is External or Internal, which is beyond the scope of this paper.

elements of the defense that follows; here I will gather these elements to make this defense explicit.

7.1 Grammar in the *Wh*-Question Typology

Miyagawa's proposal to separate the UG elements that determine available *wh*-question strategies into distinct morphosyntactic *wh*- and Q-features appears to straightforwardly account for the existence of the three *wh*-question types discussed in this paper: obligatory *wh*-movement (English-type), *wh-in-situ* with a clause-peripheral Q-particle (Japanese-type), and obligatory *wh*-movement co-occurring with a Q-particle (the admittedly rare Vata-type). I have already discussed how this proposal accounts for the first two types in section 6. As for the Vata-type, I will avoid further detailed analysis of the grammar of Vata in this paper, and instead offer a simple suggestion: It seems fairly straightforward to argue that, under Miyagawa's proposal, the Vata Q-particle *là* can be merged with the clause to satisfy the interrogative feature on C, while the *wh*-word can also raise to [Spec, C], if we assume that Vata is similar to English in having both *wh*- and Q-features on C. Thus, feature mismatch should not be a problem in this analysis, only the apparent violation of Economy of Derivation due to the redundant use of two clause-typing strategies.

As stated earlier, *wh*-movement in Vata cannot be analyzed as scrambling, since it is obligatory. I will simply suggest here that there is some requirement in the grammar of Vata that takes priority over Economy of Derivation, therefore either allowing or forcing *wh*-movement along with Q-particle Merge. I will leave the formal development of this suggestion for future research.

7.2 Performance in the *Wh*-Question Typology

Now I arrive at the question why Vata-type languages, shown above to be among the class of 'possible languages', are typologically rare. First, I review how the PGCH predicts the more common *wh*-question types.

As discussed in section 5, the UG-based assumption that Merge is less costly than Move leads to the suggestion that Merge of a Q-particle to a clause should be a primary strategy for *wh*-question formation cross-linguistically, and that *wh*-movement must be a 'last resort' when a Q-particle is not available to type the sentence as a *wh*-question. In

turn, MiD explains in processing terms why Q-particle Merge tends to be preferred to *wh*-movement in V-final languages, since the greater the complexity of the FGD formed by *wh*-movement, the more difficult that FGD will be to process. The closer the basic verb position of a given language is to the left periphery, the more that language will tolerate wh-movement as an alternative to Q-particle Merge.[9]

Given that Vata is an underlyingly SOV language, MiD explains its rarity with respect to *wh*-movement, as discussed above. As for the co-occurring Q-particle, it may be the case that Vata-type questions also violate MiF, since the final Q-particles in (4) and (8) are redundant to the parser. Thus, performance principles can explain the cross-linguistic rarity of Vata-type languages, as well as the typological distribution of the *wh*-question strategies made available by UG.

7.3 Another *Wh*-Movement Problem in Vata

Before concluding, I will briefly present another set of Vata data that suggests something interesting in the grammar of Vata related to *wh*-movement. In addition to its exceptions to CTH, Vata presents an interesting problem with respect to Keenan & Comrie's (1977) Accessibility Hierarchy, which posits that subjects tend to be easier to extract than non-subjects cross-linguistically, as well as Hawkins' (2004) closely-related Resumptive Pronoun Hierarchy Prediction (RPHP), from which it follows that resumptive pronouns should tend to occur more frequently than gaps in extraction sites as one goes down this hierarchy:

(20) *Resumptive Pronoun Hierarchy Prediction (RPHP)*
      If a resumptive pronoun is grammatical in position P on a complexity hierarchy H, then resumptive pronouns will be grammatical in all lower and more complex positions that can be relativized at all. (Hawkins 2004, p. 186)

Consider the following subject-object asymmetry in Vata, involving resumptive pronouns. When a subject is *wh*-moved, a resumptive pronoun must occur in subject position, as shown in (21). When non-subjects are moved, the occurrence of resumptive pronouns is excluded, as shown in (22):

---

[9] The question why morphologically free Q-particles are not a universal feature in all languages is beyond the scope of this paper.

(21) àlÓ *(Ò) lē   saká lá   (Vata)
     who he-R eat rice Q
     'who is eating rice?'

(22) yĪ   Kòfí lē (*mÍ) lá
     what Kofi eat (*it-R) Q
     'what is Kofi eating?'
     (Koopman 1984, p. 37)

I grant two facts here: one, the data in (21) and (22) involve *wh*-questions, not relativization (although these are closely related); and two, the RPHP is a typological prediction, not an absolute universal. Nevertheless, the data in (21) and (22), coupled with the co-occurrence of *wh*-movement and Q-particles in the same language, suggests an area for future research involving *wh*-movement constructions in Vata.

8 Summary and Conclusions

I have argued in this paper that the strategy choices available for a language to type sentences as *wh*-questions are determined by Universal Grammar, while the typological distribution of the available strategies is determined by the conventionalization of performance preferences. I have further argued that Miyagawa's (2001) proposal to divide *wh*-question morphology into separate and universal *wh*- and Q-features can account for the different forms of *wh*-questions in English-type, Japanese-type and Vata-type languages. In addition, I have attempted to show how Hawkins' (2004) Performance-Grammar Correspondence Hypothesis and its predictions about the typological distribution of *wh*-question forms can account for the rarity of Vata-type *wh*-questions, which employ both *wh*-movement and Q-particles. In other words, UG can account for the existence of Vata-type *wh*-question forms, while the conventionalization of performance preferences can account for why such forms are not more abundant: the redundancy of this strategy, as well as its violation of the efficiency principles Minimize Domains and Minimize Forms, makes it less preferable to grammars than the alternative strategies of either *wh*-movement or Q-particle Merge to establish interrogative force in a sentence.

A final note: I have presented the Vata data in this paper for comparative purposes, not to support any possible counterclaims against Cheng's (1991) Clausal Typing Hypothesis or Hawkins' (2004) Performance-Grammar Correspondence Hypothesis. Nevertheless, if one accepts that such data are accurately attested, they should be accounted for somehow, and I have argued here that Miyagawa's (2001) UG-based proposal suggests a possible account. The precise form of that account, which would require a more thorough investigation of Vata grammar than appropriate for this paper, is left for future research.

References

Bach, Emmon. 1971. Questions. *Linguistic Inquiry* 2: 153–166.

Bruening, Benjamin. 2004. *Wh*-indefinites, question particles, and *wh*-in-situ: There is no relation. ms., University of Delaware.

Cheng, Lisa L.-S. 1991. On the typology of *wh*-questions. Doctoral dissertation, MIT. Published 1997, Garland, New York.

Chomsky, Noam. 1991. Some notes on economy of derivation and representation. *Principles and parameters in comparative grammar*. Ed. R. Freidin. Cambridge: MIT Press. 417–454.

Chomsky, Noam. 1995. *The minimalist program*. Cambridge: MIT Press.

Chomsky, Noam. 2000. Minimalist inquiries: The framework. *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*. Ed. R. Martin, D. Michaels, and J. Uriagereka. Cambridge: MIT Press. 89–155.

Chomsky, Noam and Howard Lasnik. 1977. Filters and control. *Linguistic Inquiry* 8: 425–504.

Dryer, Matthew S. 1991. SVO languages and the OV:VO typology. *Journal of Linguistics* 27: 443–482.

Fodor, Janet D. 1983. Phrase structure parsing and the island constraints. *Linguistics and Philosophy* 6: 163–223.

Greenberg, Joseph H. 1963. Some universals of language with special reference to the order of meaningful elements. *Universals of language*. Ed. J. Greenberg. Cambridge: MIT Press. 73–113.

Greenberg, Joseph H., Charles A. Ferguson, and Edith A. Moravcsik, eds. 1978. *Universals of human language, vol. 4: Syntax*. Stanford: Stanford University Press.

Hagstrom, Paul. 1998. Decomposing questions. Doctoral dissertation, MIT.

Hawkins, John A. 1999. Processing complexity and filler-gap dependencies. *Language* 75: 244–285.

Hawkins, John A. 2004. *Efficiency and complexity in grammars*. Oxford: Oxford University Press.

Keenan, Edward L. and Bernard Comrie. 1977. Noun phrase accessibility and Universal Grammar. *Linguistic Inquiry* 8: 63–99.

Kenstowicz, Michael. 1987. The phonology and syntax of *wh*-expressions in Tangale. *Phonology Yearbook* 4: 229–241.

Koopman, Hilda. 1984. *The syntax of verbs: From verb movement rules in the Kru languages to Universal Grammar*. Dordrecht: Foris.

Kuroda, Sige-Yuki. 1965. Generative grammatical studies in the Japanese language. Ph.D. dissertation, MIT.

May, Robert. 1985. *Logical Form: Its structure and derivation*. Cambridge: MIT Press.

Miyagawa, Shigeru. 2001. The EPP, scrambling, and *wh*-in-situ. *Ken Hale: A life in language*. Ed M. Kenstowicz. Cambridge: MIT Press. 293–338.

Newmeyer, Frederick J. in press. *Possible and probable languages: A generative perspective on linguistic typology*. To be published by Oxford University Press.

Ultan, Russell. 1978a. Toward a typology of substantival possession. *Universals of human language, vol. 4: Syntax*. Ed. Greenberg, Joseph H., Charles A. Ferguson, and Edith A. Moravcsik. Stanford: Stanford University Press. 11–50.

Ultan, Russell. 1978b. Some general characteristics of interrogative systems. *Universals of human language, vol. 4: Syntax*. Ed. Greenberg, Joseph H., Charles A. Ferguson, and Edith A. Moravcsik. Stanford: Stanford University Press. 211–248.

A Note on the Phonetic Correlates of Stress
in Yakima Sahaptin[*]

Sharon Hargus   Virginia Beavert

sharon@u.washington.edu

1   Introduction

In this article we two aspects of what previous researchers on Sahaptin have referred to as "stress" in that language. First, we are interested in establishing the phonetic correlates of stress in YS. For practical reasons, much of the descriptive work on Sahaptin (e.g. for lexicography, Beavert and Hargus in preparation-a) has been accomplished via impressionistic transcriptions prepared by non-native speakers. However expedient, this process is risky (although rarely questioned in documentary linguistics), as non-native speakers' judgments of stress placement may be at odds with those of native speakers (see Hargus 2005 for more discussion). Therefore, one goal of this article is to see whether non-native speaker impressions of the location of primary stress can be confirmed instrumentally. If not, the basis for much of the prior and current description of Sahaptin must be re-examined. Secondly, we are also interested in whether or not YS should be categorized as a "pitch-accent" (or simply "accentual") language.

We focus here on the Yakima dialect of Sahaptin (YS). Yakima Sahaptin is highly endangered, with five remaining speakers, all elderly. All data in this study come from one speaker, the second author, who is a native speaker of the Yakima dialect, one of the northwest cluster of Sahaptin dialects (Rigsby and Rude 1996). The Yakima dialect is the only northwest Sahaptin dialect which is still spoken.

Following presentation of segment inventories in §1, we summarize various qualitative aspects of Sahaptin stress (§3), presenting pitch tracks for selected YS words. Next (§4) we present a quantitative study of stress in YS. In (§5) we discuss characteristics of stress, tone and pitch accent systems, and describe how YS fits within this typology. In §6 we outline a number of areas for possible future investigation.

2 Segment Inventory

The Sahaptin consonant inventory is given in (1):

(1)	p p'	t t'	tɬ tɬ'	ts ts'	tʃ tʃ'	k k'	kʷ kʷ'	q q'	qʷ qʷ'	ʔ
			ɬ	s	ʃ	x	xʷ	χ	χʷ	h
	m	n	l							
	w				j					

The Sahaptin vowel inventory is given in (2):

(2)	i ii	ɨ	u uu
	a[1] aa		

There are both short vowel and long vowel 'diphthongs' in Sahaptin. The short vowel diphthongs consist of [iw uj aw aj]. The long vowel diphthongs consist of [iiw uuj aaw aaj]. The term diphthong is put in quotes because it will be seen in 3.4 that the short vowel diphthongs do not pattern with long vowels, so diphthong is something of a misnomer.

3 Qualitative Observations about Sahaptin Stress

3.1 Unpredictability of Stress

Rigsby and Rude 1996 noted that 'primary stress...is distinctive and...occurs on one syllable of every word' (p. 671). They noted that [ámapa][2] 'husband' (obj.) and [amápa]

---

[1] We use the IPA to transcribe Sahaptin, except that we transcribe long vowels as two vowels rather than with the colon, we transcribe stress/accent with an acute accent over the vowel, and we use [a] rather than [ɑ] for the mid-to-low central unrounded vowel. The vowel transcribed [a] is actually closer in quality to [ə] than to [ɑ]. Instrumentally obtained vowel charts for Sahaptin stressed and unstressed vowel qualities are presented in Hargus 2001.

'island' are a minimal pair for stress in the Umatilla dialect. In YS there are similar (near-) minimal pairs for stress; e.g. [wjánawi]-³ 'arrive', [anáwi]- 'be hungry', [kʷ'ajawí] 'mountain lion'.

In an earlier study (Hargus and Beavert 2001), we noted that despite such contrasts, there were nonetheless statistical preferences for stress placement in roots. There is a greater-than-random attraction of stress to heavy syllables (–VV or -VC), a preference for trochaic stress when syllable weight is not a factor (e.g. initial stress in CVCV roots), and a preference for right-directionality (penultimate stress in CVCVCV roots).

Note that "predictability" here means predictability of the location of the accent within an accented morpheme. Whether or not a morpheme has an accent is a separate issue, discussed next.

3.2 Contrast Possibilities within Morpheme Classes

Jacobs 1931:118-119 noted that certain prefixes ('anterior root elements') and suffixes 'invariably obtain word accent', presenting examples such as those in (3).

(3)   [pá]- inverse vs. [pa]- 3PL.NOM
       [páwat'ana] 'he struck at him'
       [pawát'ana] 'they struck'

Such examples reveal that there is a distinction between stressed and unstressed affixes in Sahaptin. In our current lexical files (Beavert and Hargus in preparation-a), more than half (57%, or 54 of 95) have no underlying accent.

In contrast to affixes, there do not appear to be many unaccented roots in Sahaptin. We define root here as a morpheme which either undergoes the type of affixation characteristic of nouns, verbs or adjectives in Sahaptin, or else is not clearly an affix to a lexical or functional category. All known unaccented roots are listed in (4):

---

[2] It is traditional in Sahaptin linguistics not to transcribe word-initial glottal stop, which is predictable on the surface. We will adopt this practice in this article, as the presence vs. absence of word initial [ʔ] is not crucial for present purposes.

[3] Verb roots must surface with an affix in order to form a well-formed word. Hence the hyphen indicates that the verbal root is a bound morpheme, although it is somewhat arbitrary to place the hyphen after the verb root. The verbal affix which can be added may be either a prefix (such as ʔi- 3S.NOM or a suffix, such as –k IMP.SG). (We use the affix glossing conventions of Rigsby and Rude 1996.)

(4) Unaccented roots

   a. conjunctions: ku 'and', uu 'or', kutja 'but, however'

   b. wa 'be'

   c. evidentials:[4] akut/jakut 'supposedly', χaʃ 'I wonder', χat 'in the world, on earth'

Note that the stressed root [kú]- 'do' is thus a minimal pair with [ku] 'and' for presence/absence of stress.

   Morphemes are normally simply specified for stress or not. In addition, there are a small number of suffixes which appear to requires stress on the preceding syllable:

(5) a. -ˊtʼa ~ -átʼa 'want'

   b. -ˊɬam AGT

The pre-stressing suffix in (5)b. appears to be unproductive, relative to the very productive -ɬá AGT, and is so far only attested in the lexical items in (6):

(6) atɬʼawíɬam 'beggar'; cf. atɬʼáwi- 'ask, beg for, request'

   paχwíɬam 'thief'; cf. páχwi- 'steal'

The vowel-initial form of the suffix in (5)a. is used after consonant-final roots and after monosyllabic short vowel roots. The consonant-initial, stress-shifting form is used after other vowel-final roots or vowel-final affixes. Compare the forms in (7):

(7) Pre-stressed –ˊtʼa

   tʃíi- 'drink'

   tʃíi-tʼa- 'want to drink'

   tʃii-tá-tʼa 'want to go drink'

3.3 Culminativity

   It is implicit in Jacobs' description and examples that there is only one main stress per word. Rigsby and Rude 1996 also agree that 'primary stress… occurs on one syllable

---

[4]   The evidential morphemes in
(4)c. appear to occupy second position in the sentence, a position shared by another class of unstressed morphemes, the better-known second position pronominal clitics of Sahaptin (=naʃ/Vʃ 1SG; =nam/Vm 2SG; =maʃ 1SG.2SG; =mataʃ 1PL.2, 1SG.2PL; =na, =nataʃ 1DU.INCL; =taʃ 1DU/PL.EXCL; =natk/namtk 1PL.INCL., =pat 3PL.INVERSE).

of every word'. Predicting which syllable surfaces with stress when a word contains multiple stressed affixes, Jacobs 1931:119 wrote that prefixes and anterior root elements are stressed 'except where...the suffix receives the accent...In the verb...the root is stressed if nothing else is stressed.' In other words, suffixes attract stress over prefixes, and prefixes attract stress over roots. Exemplification of this rule can be found in Hargus and Beavert 2002a, Hargus and Beavert 2002b and throughout this article.

3.4 Stress Realized as High Pitch

Jacobs 1931:117 noted that 'stress and high tone are one phenomenon in northern Sahaptin; they are very strongly marked in northwest Sahaptin...light monosyllabic words are invariably stressed and have high tone...in all dialects the syllable that has the stress takes high or falling---that is, high to normal---tone. Short vowels have high tone'.

In our inspection of pitch tracks of words containing short vowels, we concur with Jacobs that short vowels have high and essentially non-falling pitch. A good example of this pattern is shown in (8). The mean pitch of the vowel in this word is 211 Hz. For reasons of space, in this section we compare syllables in only one position within the word, namely word-finally.[5]

(8)   mjú 'brother-in-law (man's wife's brother)'

---
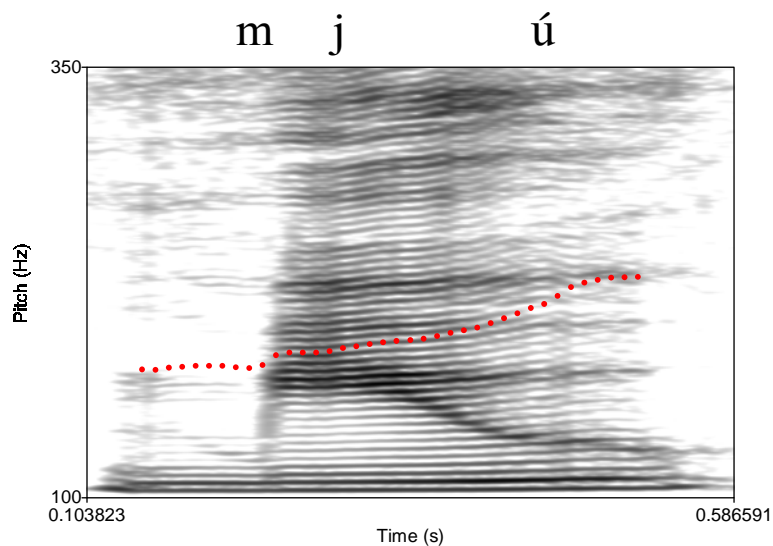
[5] The pitch tracks in this section were generated with Praat 4.3.27. The following 'advanced' pitch settings were used: Voicing threshold = 0.45, Octave cost = .04, Octave jump cost = 16.0, Voiced/unvoiced cost = 0.6.

If there is a fall in such CCV words, it is very short and occurs only on the final few milliseconds of the word, and is quite different from the long pitch fall that starts in the middle of a long vowel (as can be seen below in (12) and other graphs). For example, in (9), the final pitch fall occurs over the final 4 ms of the word. The mean pitch over the final vowel in this word is 226 Hz.

(9)   [ntʃʼí] 'big'

n   t   ʃ   ʼ   í
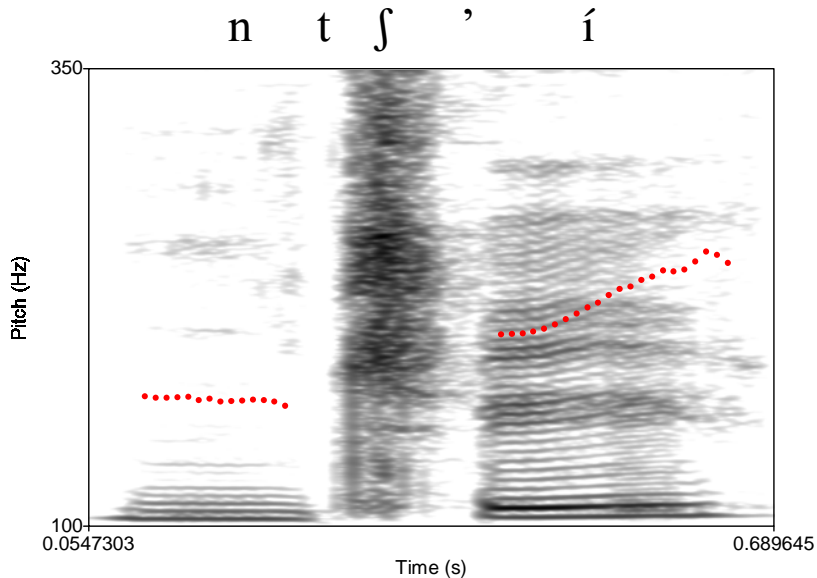
Stress in syllables closed with sonorant consonants [n m l] also appears to be realized with high, non-falling pitch. This point was also essentially noted by Jacobs 1931:117, who writes that 'One moraed or light monosyllabic words are invariably stressed and have high tone. A tonal glide back to normal is rarely heard, either in the monosyllable, or in the succeeding mora or word. Thus, wá, náxc, t'sá, mún, … q'áp.' (Jacobs 1931 is transcribing 'tone beside primary stress' in these and other words in this section of his grammar.) (10) contains an example of this high, non-falling pitch pattern. The average pitch over the final rhyme [un] is 218 Hz.

(10) [pamún] 'sometimes'

p   a   m   ú   n
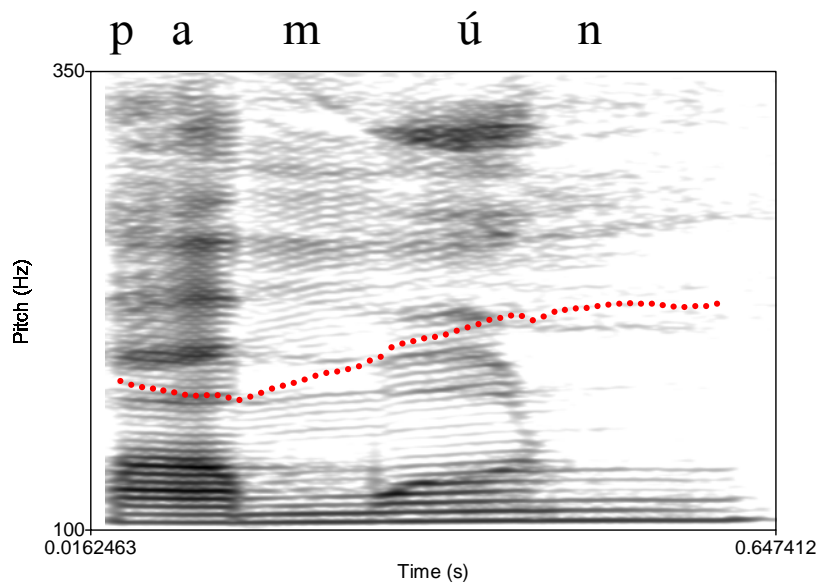
If there is a pitch fall in sonorant-consonant-closed short-vowel words, it is confined to the final few milliseconds, as with word-final short vowel words. An example with a final short fall is given in (11). In this word, the average pitch over the final rhyme is 228 Hz, and the pitch fall occurs over the final 3 ms.

(11) [patúl] 'junk'

p a   t   ú   l
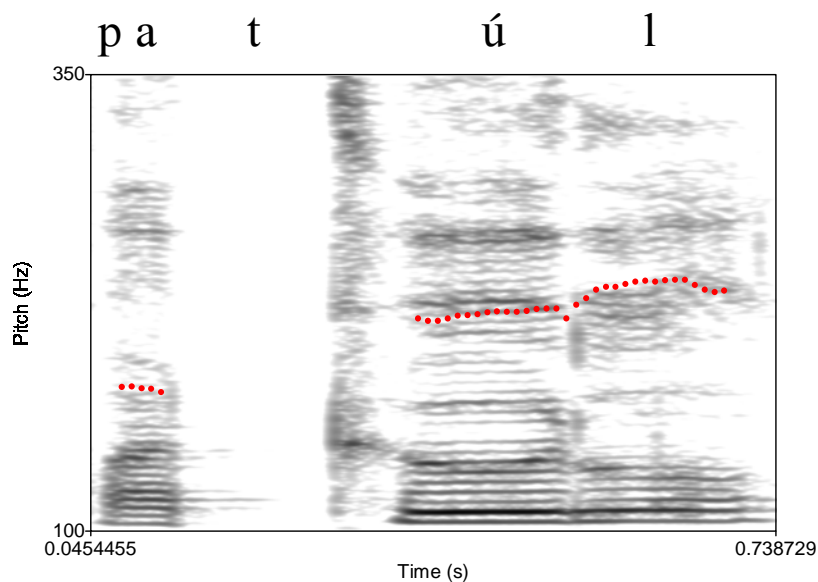
Jacobs 1931:117 noted that 'long vowels or diphthongs in accented syllables have falling tone, high to normal. Two moraed or heavy monosyllabic words invariably take

stress with falling — high to normal — tone ... ấu, tíːn, wấu, cpʼấu, hûi.' Inspection of pitch tracks again shows that Jacobs was right on this point, at least as far as the long vowels are concerned. We find that stressed long vowels in open syllables are pronounced with a fall in pitch which begins around the midpoint of the vowel. In the example given in (12), the average pitch over the final rhyme is 207 Hz:

(12) [tʃʼiwáa] 'bow-legged'
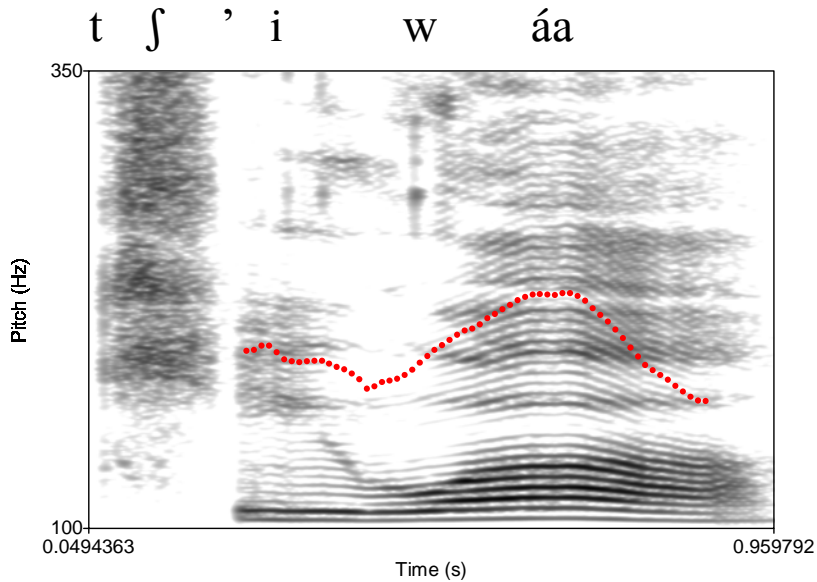
Jacobs' ambiguous phrase 'long vowels or diphthongs' apparently meant 'long vowels or short diphthongs', judging from the examples provided (above (10)) in illustration of the supposed falling pitch on both types of rhymes. On this point we differ from Jacobs, in that we find that short diphthongs are characteristically pronounced without the pitch fall found with long vowels. An example can be seen in (13). The average pitch over the final rhyme is 199 Hz.

(13) [ɨmnɨmwáj] 'playful, mischievous'

ɨ m n ɨ m w á j
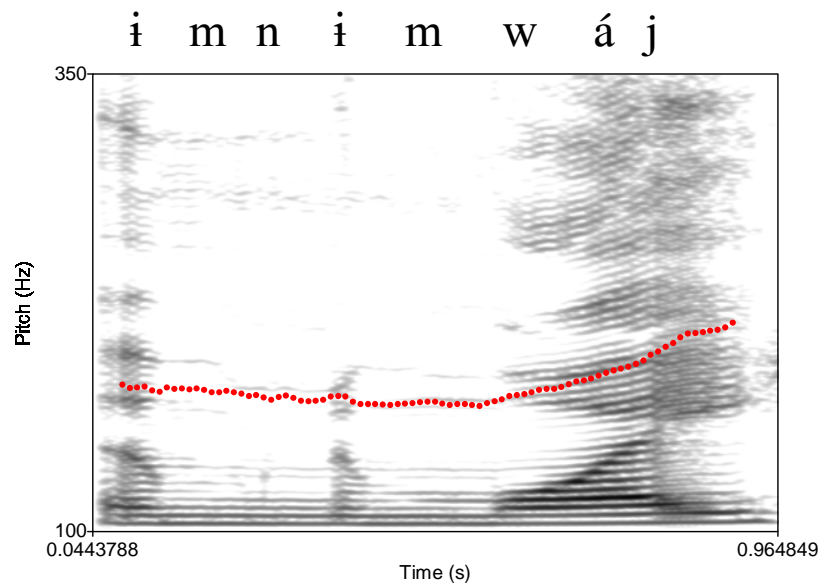
If there is a pitch fall in words with final short-vowel diphthongs, it is confined to the final few milliseconds of the rhyme, as with short vowels in open or sonorant-consonant closed syllables. An example can be seen in (14), where the pitch falls over the final 7 ms. only. The average pitch on the final rhyme is 243 Hz.

(14) [hananúj] 'bothersome'

h a n a n ú j
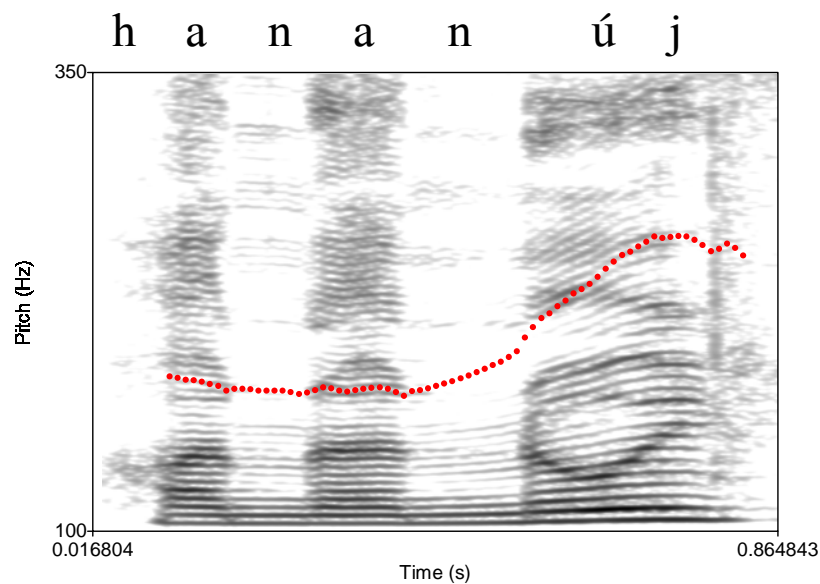
In contrast to short diphthongs, with long diphthongs we find the same pitch fall as with long vowels. An example is shown in (15). The average pitch on the final rhyme is 193 Hz.

(15) [palaláaj] 'lots'

p a l a l á a j
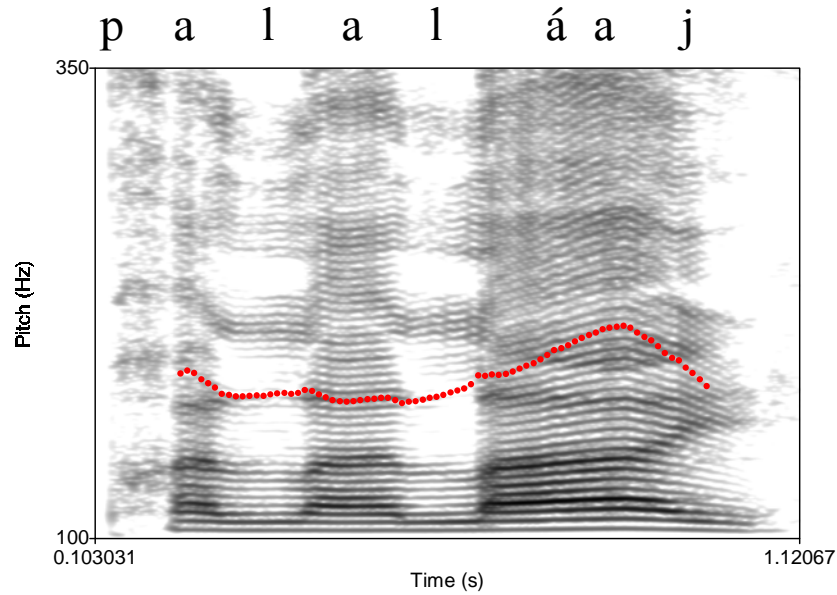
In short vs. long vowels in word-final syllables closed with an obstruent here too we find that the pitch peak of a short vowel is aligned either with the end of the vowel or very close to its end. A representative pitch track on a short-vowel obstruent-closed syllable is shown in (16). The average pitch on the final rhyme is 204 Hz:

(16) [ajajáʃ] 'stupid person, idiot'

a   j   a   j   á   ʃ
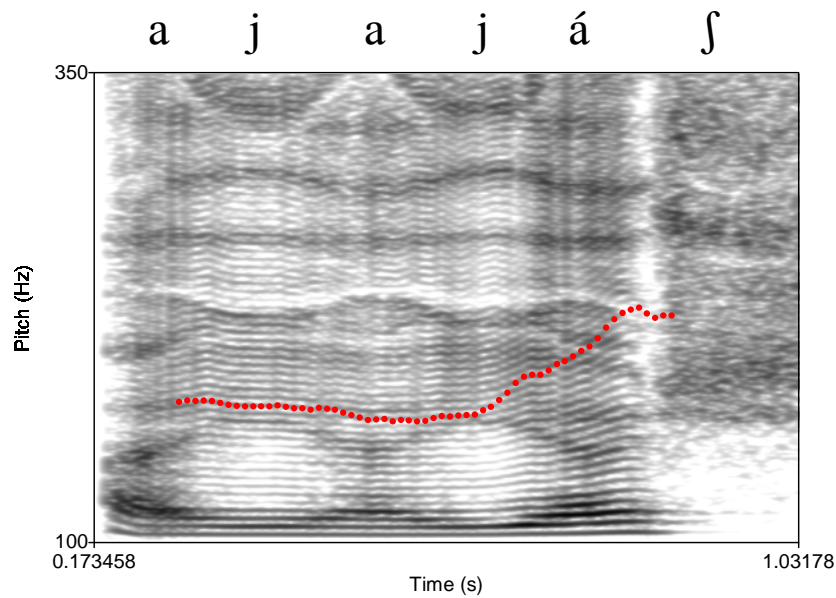
Contrast the pitch pattern seen in (17), which also contains a long vowel closed with an obstruent. The average pitch on the final rhyme is 211 Hz.

(17) [hawláak] 'empty space'

h a   w   l   áa   k
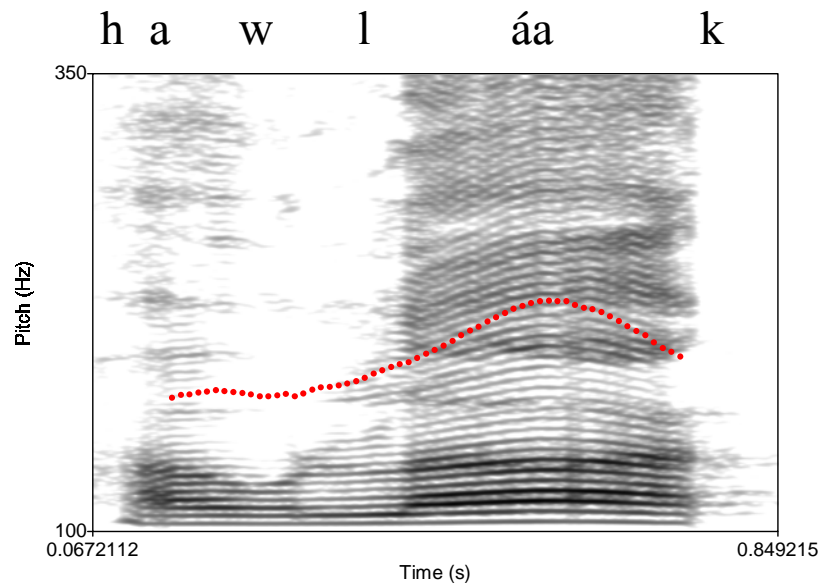
3.5 Stress Insertion

In addition to selecting the strongest stress or moving stress to the preceding syllable, the phonology associated with Sahaptin stress also appears to insert accent in

unaccented words. The unaccented free morphemes in (4) are not normally said in isolation. However, in the course of preparing Beavert and Hargus in preparation-b, a dictionary with sound files accompanying every lexical entry, we recorded the normally unaccented conjunctions [ku] 'and', [uu] 'or', and [kutja] 'but', as well as the variably accented [wa] 'be' (see Hargus and Beavert 2006 for discussion of this morpheme). We found that these morphemes have a word initial high pitch (or falling pitch in the case of [uu]) pitch when pronounced in isolation. (18)-(19) present pitch tracks for [uu] 'or' in isolation and in the middle of a phrase.

(18) [uu] 'or' (pronounced in isolation as [úu])

úu
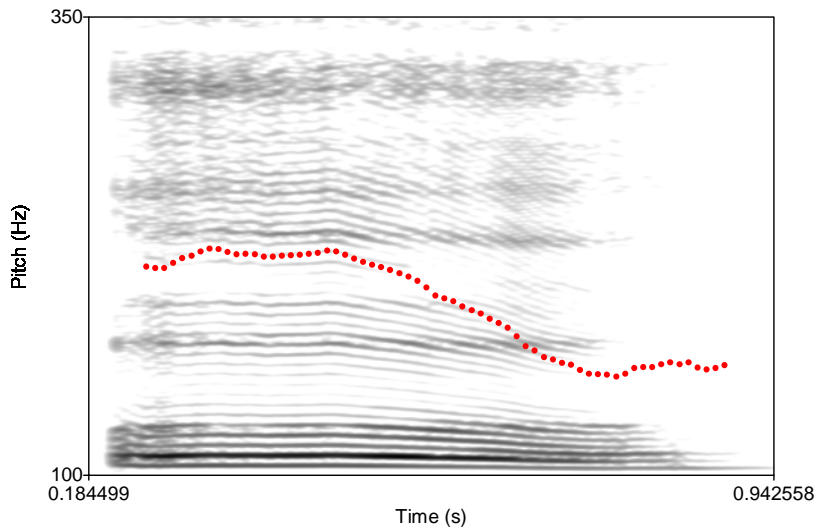
(19) [uu] 'or' in context, showing intrinsic lack of accent. The phrase is [ʃukʷáataj míʃ áw iwáta majkmáal łkʷí uu majkmáal sts'át] 'to find out whether there would be longer day or longer night now'
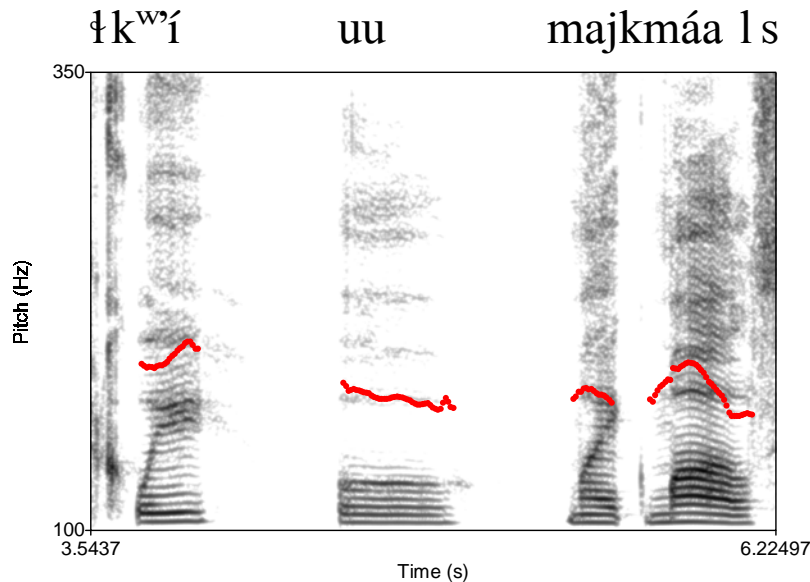
This alternation between unstressed and stressed forms of conjunctions suggests that there is insertion of accent in some accentual domain, since stressed morphemes have not been observed to lose stress in any phrasal contexts. Insertion of stress is consistent with our previous account (Hargus and Beavert 2006) of the accentual properties of [wa] 'be', which is also predictably stressed or not in surface forms. There we suggested that [wa] becomes accented when a stress lapse of more than two syllables would otherwise occur.

3.6 Summary

Previous descriptions of Sahaptin have stress have noted that there is one primary stress per word. When multiple morphemes with underlying stress coincide in the same word, stress is realized on the outermost suffix, then the outermost prefix, and then on the root.

The primary phonetic correlate of stress was previously described as high or falling pitch. Our pitch tracks of stressed word-final syllables in open syllables or in syllables closed with various types of consonants confirm this description. We have seen that

short stressed vowels have a high (non-falling) pitch contour whereas long stressed vowels fall in pitch.

4   A quantitative Study of the Phonetic Correlates of Primary Stress

The pitch tracks presented in 3.4-0s partially confirm our (and previous researchers') auditory impression that stress is realized as high pitch in Sahaptin, at least in the restricted word-final environment considered there. However, as noted by e.g. Fry 1955, stress may be realized on vowel nuclei as longer duration, higher pitch, and/or increased amplitude. Our research question in (20) therefore remains unanswered by the information presented in the previous section:

(20) What are the phonetic correlates of primary stress in Sahaptin?

4.1   Methods

An acoustic study of the correlates of stress in a controlled sample of Yakima Sahaptin words was undertaken.

4.1.1   Word List

In this study, we compared the underlined vowels in the words in (21).

(21) Word list

primary stress location	*word-initial morpheme*	
second syllable	[pa]- third person plural nominative	[pap'ɨχʃa] 'they remember'
		[papnúʃa] 'they're sleeping'
initial syllable	[pápa]- reciprocal	[pápap'ɨχʃa] 'they remember each other'
	[pá]- inverse	[páptɬ'aʃa] 'he's baptizing her'

The words in (21) control for (a) vowel quality in the initial syllable, (b) surrounding consonantism, and (c) syllable closure. Only four words were examined in this study. However, the vowels of interest in these words occur in balanced numbers of open (pap'íχʃa, pápap'iχʃa) and closed syllables (papnúʃa, páptɬaʃa), an important consideration since syllable closure can affect vowel duration (Maddieson 1985). The word list could have been expanded to include other words with [pa]- third person plural nominative, [pápa]- reciprocal, [pá]- inverse, but it was felt that additional words with these prefixes would not increase the generalizability of the study. The word list was not expanded to include other words, say those with initial [tat], because there are no other affixes which occur in the same position in the word and which differ only in stress like these affixes do.

4.1.2 Recording

Each word was pronounced in isolation five times by the second author in her home in Toppenish, Washington. The four words were recorded in random order (in blocks of five repetitions each). Recordings were made on an analog tape recorder (Sony WM-D6C) with an external microphone (Sony ECM-929LT). The recording was later digitized at a sampling rate of 10,000 samples/second using SoundEdit.

4.1.3 Measurements

Measurements were performed with Multi-Speech 2.5.1. In a first pass through the data, tags were inserted in three places: before the stop burst (of the [p] preceding the vowel of interest), at the onset of the vowel, and at the offset of the vowel. Vowel onset was defined as the onset of the relatively high amplitude portion of the vowel, and vowel offset as the point in the vowel where amplitude drops sharply. Vowel onset and offset did not always coincide with the onset or offset of periodicity. Also in the first pass through the data, voiced period marks inserted by Multi-Speech were inspected for accuracy and corrected by hand, if necessary. (22) shows a waveform and spectrogram for one repetition of [pápap'iχʃa] 'they remember each other'. Note that the bulk of the vowel occurs between the second and third tags.

(22) Waveform and spectrogram for [pápap'ɨχʃa] 'they remember each other'
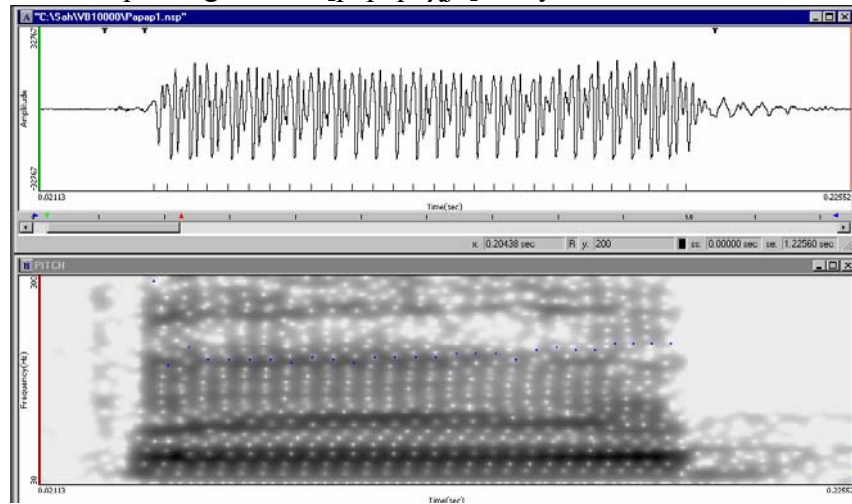
In a second pass through the data, the location of the tags were examined for consistency throughout the data set. Then duration, pitch and energy were measured as described below.

Two duration measures were obtained from the first vowel in each word, from the burst of the preceding stop to vowel onset, and from vowel onset to vowel offset. Pitch and energy were calculated over a 30 ms. average of the vowel midpoint for each of the first and second vowels. Pitch was measured from the voiced period marks inserted by Multi-Speech.

4.1.4 Normalization

Measuring the pitch and energy of the first and second vowels of each word allowed word-normalized pitch and energy measures to be calculated by subtracting the pitch of the vowel of the second syllable pitch from that of the first syllable. Normalized energy was calculated in an analogous way.

Duration was not normalized in this manner because the consonant offset following the first vowel — [p'], [pn], [p], [ptɬ'] — could not be controlled for, nor could the quality difference seen in the second syllable. Since both of these factors could affect the duration of the vowel of the second syllable, any conclusions about stress based on second syllable vowel duration could not be regarded as untainted by factors unrelated to stress.

The vowel quality differences of the second syllable are also a potential confound for the normalization method for energy. Consider the initial unstressed sequences [a] ... [ɨ́], [a] ... [ú] vs. the initial stressed sequences [á] ... [a], [á] ... [a]. High vowels contain lower intrinsic amplitude than non-high vowels (Ladefoged 1992). The high vowels of the second syllables in the first case should make it easier to obtain a statistically significant result for energy, since there is already an energy difference between the two types of sequences independent of stress. The energy difference between the vowels in the first case should be smaller than the energy difference between the vowels in the second case.

4.1.5 Statistics

Inferential and descriptive statistics were obtained using StatView 5.0.1. The inferential statistical tool used in this study was factorial analysis of variance (ANOVA).

4.2 Results
4.2.1 Duration

The results for each of the duration measures, burst-vowel onset and vowel onset-offset, are first presented separately. Then the result of adding the measures together is presented.
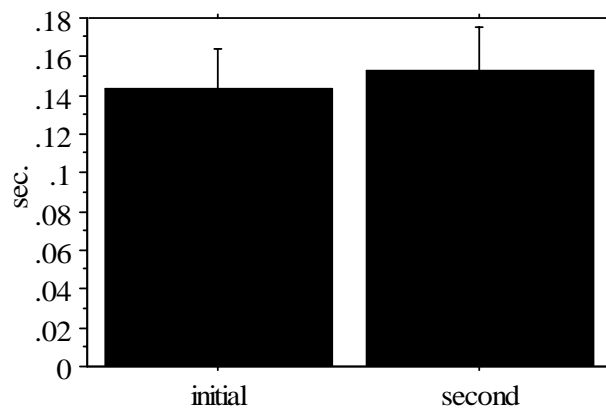
(23) shows the average duration of the burst-vowel onset measure for the word-initial vowels in pápapʼɨ̀χʃa, páptɬʼaʃa (initial stress) vs. papʼɨ́χʃa, papnúʃa (second syllable stress). Error bars in this and other graphs represent one standard derivation. The average duration of the burst-vowel onset measure for the vowels predicted to be stressed was 10 ms. (s = 1.9), and that of the vowels predicted to be unstressed 17 ms. (s = 5.5). The difference in burst-vowel onset duration for stressed and unstressed vowels was statistically significant ($F[1,19] = 13.632$, $p = .00015$), although the opposite of what would be predicted if the vowels in (21) are correctly transcribed for stress (and if duration is a phonetic correlate of stress).

(23) Burst-vowel onset measure for stressed vowels ("initial") vs. unstressed vowels ("second")
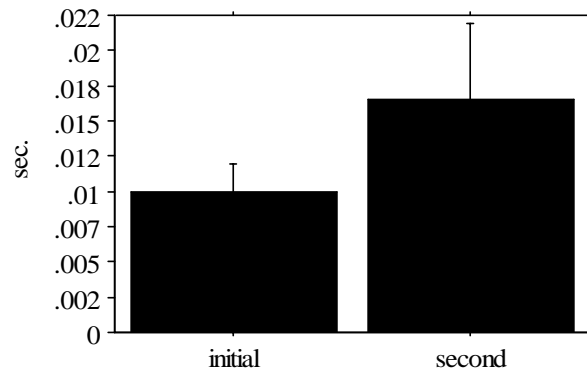

(24) shows the average duration of the vowel onset-offset measure for the first syllables in pápapʼɨχʃa, páptɬʼaʃa (initial stress) vs. papʼɨχʃa, papnúʃa (second syllable stress). The average onset-offset duration of the vowels predicted to be stressed was 143 ms. (s = 20.3) and that of the vowels predicted to be unstressed 153 ms. (s = 22.6). The difference was not statistically significant.

(24) Vowel onset-offset measure for stressed vowels ("initial") vs. unstressed vowels ("second")

Finally, (25) shows the average duration of the sum of these two measures for the word-initial vowels in pápapʼɨχʃa, páptɬʼaʃa (initial stress) vs. papʼɨχʃa, papnúʃa (second syllable stress).

**(25)** Sum of duration measures for initial vowels in words with initial vs. second syllable primary stress
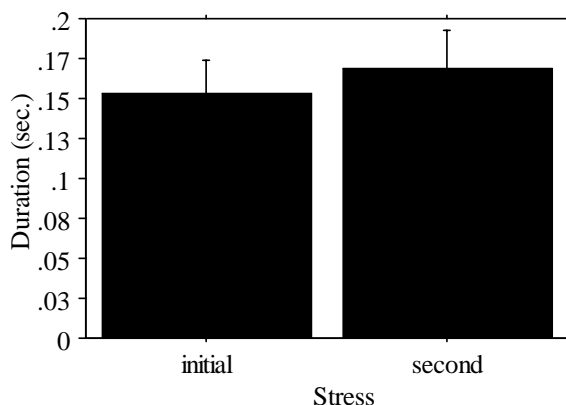
The sum of the duration measures for the first vowels in the words with initial syllable stress was 154 milliseconds (s = 20.7). The sum of the duration measures for the first vowels in the words with second syllable stress was 170 milliseconds (s = 23.9). This difference was not statistically significant.

4.2.2 Pitch and Energy

(26) plots the average pitch of the vowels of the first (pitch1) vs. second (pitch2) syllables for words with initial vs. second syllable stress. As can be seen from (27), there is a fall in pitch from the first (192 Hz) to the second (173 Hz) syllables in pápapʼɨχʃa, páptɬʼaʃa but a rise in pitch from the first (172 Hz) to the second syllables (224 Hz) in papʼɨχʃa, papnúʃa.

(26) Pitch contours of words with initial vs. second syllable primary stress
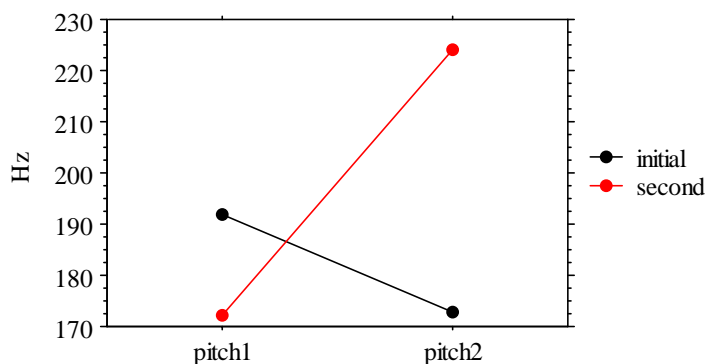
(27) plots the average energy of the vowels of the first (energy1) vs. second (energy2) syllables for words with initial vs. second syllable stress. There is a fall in energy from the first (74.3 dB) to the second syllable (70.3 dB) in words with stress on the first syllable, whereas words with second syllable stress have a rise in energy from the first (67.2 dB) to the second syllable (71.0 dB).

(27) Energy contours of words with initial vs. second syllable primary stress
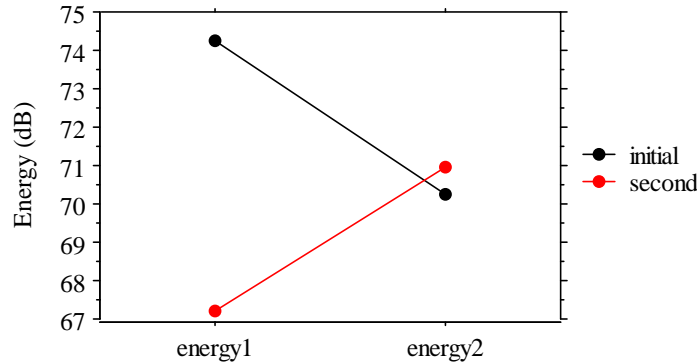
Initial syllable stressed words showed an average fall in pitch of 19 Hz (s = 12.2) and fall in energy of 3.9 dB (s = 3.63), whereas second syllable stressed words showed an average rise in pitch of 52 Hz (s = 10.5) and a rise in energy of 3.8 dB (s = 3.76). The difference in normalized pitch and energy for initial vs. second syllable stressed words was significant for both measures (pitch, F[1,19] = 199.404, p < .0001; energy, F[1,19] = 23.224, p = .0001). As discussed in 4.1.4, there are intrinsic differences in energy between the two comparison sequences which stem from vowel quality differences in the second syllables. Given the nature of the current results, this potential concern can be ignored. The vowel quality difference between the vowels with second syllable stress should cause the energy difference to be smaller than it would otherwise be if the vowels of the second syllable were [a] rather than high vowels. Therefore, we conclude that the statistically significant result for normalized energy would only have been greater if the vowels of the second syllables had been perfectly matched for vowel quality.

4.3 Discussion

Our finding of pitch as a phonetic correlate of stress at the left edge of the word is in accord with our qualitative finding that stress at the right edge of the word is marked by

higher pitch. Our finding that energy is also a phonetic correlate (at least in syllables at the left of the word) was not predicted but not unexpected from a cross-linguistic point of view. As mentioned above, in a survey of the phonetic correlates of stress, Fry 1955 notes that higher pitch, longer duration or greater energy are all possible phonetic correlates of stress. Lehiste 1970:144 has noted that is not surprising for both pitch and energy to mark stress:

There is solid evidence…that various dependence relationships exist between [fundamental frequency and intensity]…[I]ncreases in subglottal pressure produce an increase in the rate of vibration of the vocal folds, unless there is some compensatory adjustment in their tension.

Two out of the three duration measures were not significant, and the one duration measure that was statistically significant could not be explained by stress, as the longer duration occurred before the vowels which were predicted to be unstressed. We therefore conclude that duration is not a phonetic correlate of stress. This is perhaps not surprising given that Sahaptin uses duration to contrast long and short vowels (e.g. /tun/ 'what', /tuun/ 'which'). Hayes 1995 predicted that languages with vowel length distinctions would not in fact employ duration as a phonetic correlate of stress. On the other hand, this is not the case in Aleut (Rozelle 1997, Taff, Rozelle, Cho, Ladefoged, Dirks and Wegelin 2001), which has a vowel length contrast like Sahaptin.

5   Is Sahaptin a Pitch-Accent Language?

In this article we have referred to the "stress" system of Sahaptin, using the term of previous researchers (Jacobs, Rigsby, Rude) (although Jacobs also refers to "accent" in Sahaptin). However, we have seen that pitch, along with energy, is a phonetic correlate of stress in Sahaptin. The question then arises as to whether Sahaptin might best be classified as a 'pitch-accent' language, as opposed to a 'stress(-accent)' language.

Defining 'pitch-accent language' is a question which has vexed many generative phonologists. Part of the difficulty is that this term has changed in meaning over time from 'tone language' to refer to a language that uses tone (pitch) in some prosodic way (see Beckman 1986 for a review of the uses of the term 'accent'). For example, in the first part of the 20th century, Sapir 1925 used the term in reference to the prosodic system of Sarcee, as did Hoijer 1943 with respect to languages of the Apachean

subfamily (Athabaskan). These are linguistic varieties which are now recognized as having canonical tone.

In this section, we first review the characteristics of pitch-accent, also simply known as 'accent', and then see whether this term is applicable to Sahaptin.

5.1 Pitch-Accent vs. Tone

Various researchers (McCawley 1978, Beckman 1986, Zhivov 1978, Odden 1999) have suggested that the differences between tone and pitch-accent systems are primarily functional and phonological in nature rather than phonetic.

McCawley 1978 describes the prosodic system of standard Japanese (usually considered a model pitch-accent language), other varieties of Japanese ((a) the central Honshuu dialects and 'most of Shikoku', (b) the dialects of western and southern Kyuushuu), which have more complex variations on the standard Japanese prosodic system, such as allowing pitch contrasts on the initial syllable), three Bantu languages (Ganda, Tonga, and Kikuyu, which differ in degree of tonal specification/predictability), and Mandarin Chinese (a canonical tone language). He concludes that pitch-accent and tone languages are two ends of a continuum, and that 'there is no reason for squeezing the diversity of phonological systems discussed here into a simple dichotomy' (p. 128).[6]

Zhivov 1978 and Beckman 1986 follow McCawley 1978 in suggesting that the different phonological rules found in accent vs. tonal systems might be a distinguishing property of the two types of languages. Zhivov 1978:97 is primarily concerned with differentiating 'restricted tone' languages (term originally due to Voorhoeve 1973) from pitch accent languages, and notes that 'tone sandhi rules can exist only in [restricted tone] languages (as well as in genuine tone languages) and cannot exist in [pitch accent] languages.' According to Beckman 1986: 'unlike tonal alternations...accentual alternations...are best described as the suppression of an accent in order to preserve the culminative principle of one primary accent per lexical unit'. Odden 1999 considers the history of analysis of Bantu pitch-accent systems (e.g. Tonga), and concludes that

---

[6] This point is essentially echoed by Hulst and Smith 1988:ix: "What we will suggest here is that a simple division among the Non-stress cases into two categories (Pitch Accent Language – Tone Language) represents a gross over-simplification of the facts. Rather, it seems to be the case that languages employing non-intonational pitch distinctions make up a continuum that from a theoretical point of view should be approached in terms of a set of parameters that seem to define systems as being more typically "Tone Languages" or more typically "Pitch Accent Languages"."

such languages could just as well be analyzed as having restricted tone. However, he stops short of saying that pitch-accent has no theoretical status as a typological prosodic category.

5.2 Pitch Accent vs. Stress

Less attention has been paid to distinguishing stress and accent languages than has been paid to the distinction between stress and tone languages. Beckman 1986 suggests that the distinction is primarily phonetic, and provides production and perception evidence from Japanese vs. English for her 'stress-accent hypothesis'; namely, that 'stress accent differs phonetically from non-stress accent in that it uses to a greater extent material other than pitch.' In her production study, she found that the phonetic correlate of Japanese accent was pitch only, whereas the phonetic correlates of English 'stress-accent' were pitch, duration and amplitude. In her perception study, she found that monolingual American English speakers' judgments of stress in stimuli derived from English words were based on pitch, total amplitude, spectral qualities, and duration, although there was some variability among the American English speakers. In contrast, Japanese speakers judging initial vs. final accent in Japanese-derived stimuli consistently judged only pitch as a reliable cue to accent placement; duration, total amplitude and spectral qualities were "completely ineffective" cues (p. 184).

Beckman 1986:1 ff. proposes that stress and accent are two closely related notions, which she defines as follows. '"accent" means a system of syntagmatic contrasts used to construct prosodic patterns which divide an utterance into a succession of shorter phrases and to specify relationships among these patterns which organize them into larger phrasal groupings.' Accent serves an 'organizational function', in that 'in any given utterance, more prominent portions alternate and contrast syntagmatically with less prominent portions, creating a series of accentual phrases that are delimited by or centered around the prominent portions.' In contrast, "stress" means a phonologically delimitable type of accent in which the pitch shape of the accentual pattern cannot be specified in the lexicon but rather is chosen for a specific utterance from an inventory of shapes provided by the intonation system.'

Adding to the confusion in the use of these terms, 'accent' is also used in the intonation literature to refer to a 'prominence-lending pitch movement … a rise, fall, or

combination of the two' which is added to a 'word or a word group…in focus' (Sluijter 1995:2).

5.3 Yakima Sahaptin as a Pitch-Accent Language

The best agreed upon characteristics of pitch-accent in comparison with canonical stress on the one hand and tone on the other are summarized in (28), which is based on the references cited in 5.1-5.2 and also draws on work by Alderete 1999, Gomez-Imbert 2001, Hyman 2001, and Yip 2002. The last three characteristics in (28) are from Beckman 1986, from her explicit comparison of stress and pitch-accent (tone is not compared with respect to these traits). It can be seen that there is overlap in the characteristics of all three types of prosodic systems, as noted in the literature:

(28)	stress	pitch-accent	tone
predictability	stress may or may not be predictable	location of accent may or may not be predictable	location of tone may or may not be predictable
surface contrasts on monosyllables	no	accented vs. unaccented	more than binary contrasts on monosyllables possible
underlying contrasts on morphemes	unstressed vs. stressed	accented vs. unaccented	no tone vs. tone(s) (of particular quality(s))
possible phonological phenomena	one accent per domain alternating stress resolve stress clash	one accent per domain accent shifts	tone sandhi tone spread OCP, dissimilations
pitch contour	stress has various possible phonetic correlates, including pitch	pitch predictable from underlying accent	pitch predictable from underlying tone

distinctive load[7]	?	relatively small number of minimal pairs; relatively predictable 'accentual pattern'	relatively large number of minimal pairs; morphemes may consist of tone alone
probable historical origins	?	'rephonologization of tone' (Clements and Goldsmith 1984) or 'grammaticalization of intonation' (Hyman 1977)	reanalysis of originally segmental contrast
attitude by native speakers	?	suprasegmentalization makes sense to native speakers	may be puzzled by linguist's suprasegmentalization analysis of tone

It can also be seen from (28) that pitch-accent appears to be a special case of stress, with a syndrome of characteristics defining a pitch-accent language as such: phonetic realization of stress as high pitch, division of morphemes into accented vs. unaccented, unpredictable location of accent within a morpheme, and the marking of culminativity within a certain domain the phonological phenomenon universally found in such languages. As reviewed in §3, YS appears to have all of the predicted characteristics of a pitch-accent language, including realization of accent with a predictable pitch contour, in this case the high pitch which is often attracted to stress and which stress often attracts (de Lacy 2002).

6 Future Work

Our study of the phonetic characteristics of accent in YS raises a number of areas for possible future investigation.

---

[7] defined as 'the relative amount of work they do as distinctive features' (Beckman 1986:36)

6.1 Compounds and Phrases

The predictability of accent in compounds and phrases is not entirely clear at this time. However, it appears that normally both phrases and compounds surface with a pitch peak on each word of the phrase/compound.

Some lexical items contrast with homophonous phrases in that the lexicalized form contains one accent whereas the fully compositional form contains two accents. One such pair is given in (29)-(30):
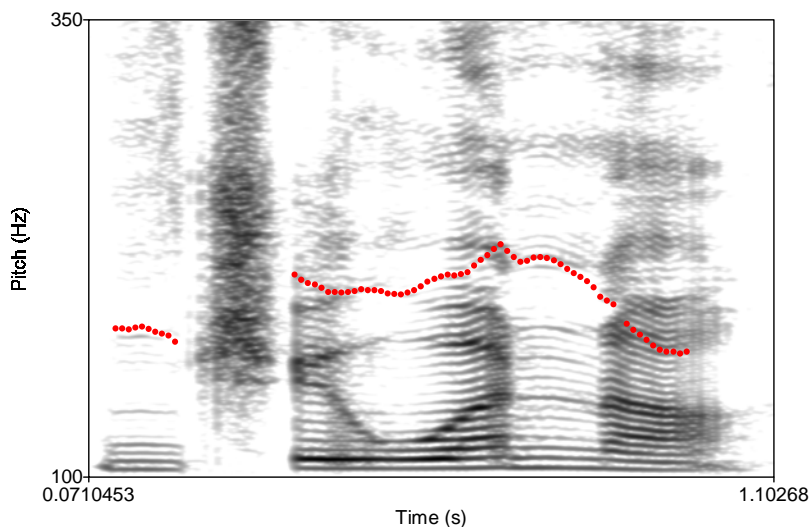
(29)   [ntʃʼí wána] 'big river'
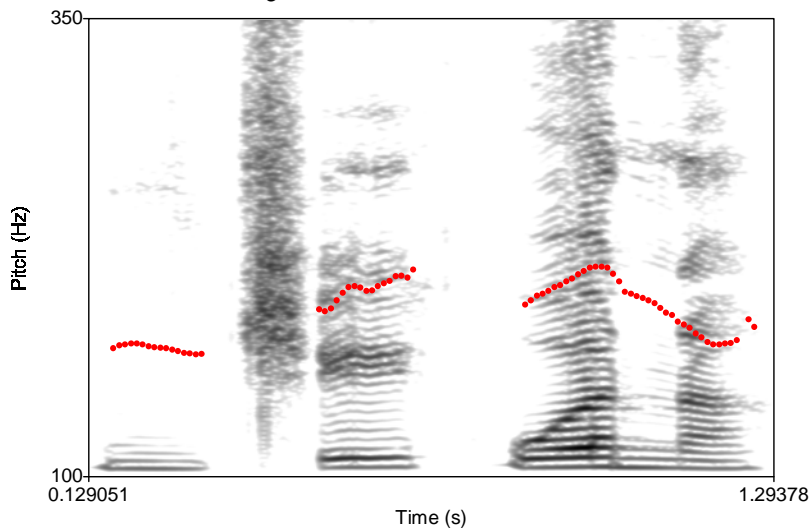
(30)   [ntʃʼi wána] 'Columbia R.' (lit. 'big river')

Such contrasts are reminiscent of another distinction that we have noted (Hargus and Beavert 2002a) in reduplicative compounds: single vs. double applications of [ɨ] epenthesis within certain sonorant initial clusters; e.g. [wɨχáwχa] 'feet, legs' (collective pl.) vs. [wɨχá wɨχa] 'feet, legs' (distributive pl.). We have analyzed this as one Prosodic Word vs. two Prosodic Words.

6.2 Intonation

Beckman 1986:5 notes that 'it is impossible to give an adequate description of the production and perception of accent patterns in English without describing at the same time the phonetic and phonological structures of intonation.' Some very basic questions about the relation between stress and intonation in Sahaptin remain to be answered. Is there a system of intonation pitch contours in addition to the word accent system? If so, are there different boundary intonational contours associated with different types of sentences (declarative, yes/no question, *wh*-question, imperative, etc.)? Do certain words or certain syntactic structures within a sentence attract the highest intonation peaks? If there is downdrift within a sentence, does pitch reset itself with every sentence, or perhaps with every clause (in the case of complex sentences)?

The YS words which were the subject of the acoustic study in §4 were recorded in isolation. This method had the advantage of not introducing a focus-driven pitch on such words which might override the inherent stress. However, recording and analyzing words within two types of carrier phrases, one with focus on the target word and one with focus on some word in the carrier phrase, along the lines of the study carried out by Sluijter 1995 for Dutch,[8] could shed further light on the acoustic characteristics of stress/accent in YS.

---

[8] In one experiment, Slujter recorded target words contrasting in stress in two types of carrier phrases:

	initial stress	final stress
accent on the target	Wil je KAnon zeggen (en niet liedje)	Wil je kaNON zeggen (en niet geweer)
	'Will you canon say (rather than song)'	'Will you cannon say (rather than rifle)'
no accent on the target	Wil je **ka**non ZEGgen (en niet opschrijven)	Wil je ka**non** ZEGgen (en niet opschrijven)
	'Will you canon say (rather than write down)'	'Will you cannon say (rather than write down)'

See Sluijter 1995:44 ff. for more information.

6.3 Secondary Stresses

According to Rude 1988, '[secondary stress] is predictable in bisyllabic reduplication, e.g. kʼùsikʼúsi 'dog', and is regular in the imperfective past -*shana* [ʃana] and [habitual past] –*xana* [χana]. It is not at this time clear whether the secondary stress is always predictable.'[9] Rigsby and Rude 1996 note that 'nondistinctive secondary and lesser stresses occur phonetically … but are not discussed here.' Jacobs 1931:117 indicates that presence/absence of secondary stress may differentiate Sahaptin dialects. He writes that in the northwest Sahaptin dialects (which include YS) 'ordinary words have only one syllable accented and no secondary stress … whereas in the Umatilla reservation dialects that may be two, three or four accented syllables to a word.'

Hargus and Beavert 2002a found support for secondary stress in YS from reduplication and from certain asymmetries between prefixes and suffixes. Hargus and Beavert 2002b identified a phenomenon which we called Destressed High Vowel Deletion. Unstressed /i u ɨ/ delete unless an illicit cluster would be formed, with the further restriction that /i u/ must be adjacent to a homorganic glide in order to delete. The examples in (31) illustrate the deletion of [ɨ]. The first example contains the third person singular nominative prefix [i]-, an unstressed prefix, whereas the second contains [á]- absolutive, a stressed prefix. Both are prefixed to a root with underlying stress on the first syllable.

(31) /pɨ́tja/- 'spear'

  /i-pɨ́tja/   [májtsqi ipɨ́tja] 'she speared it this morning'

  /á-pɨ́tja/   [májtsqiiʃ áptja] 'I speared it this morning'

In the second example, deletion of [ɨ] has taken place because the surface stress falls on the prefix, and with the removal of the stress from the vowel [ɨ], the vowel too disappears. In Hargus and Beavert 2002a we reported that whereas deletion of destressed [ɨ] regularly takes place when stress shifts to a prefix, this was not the case with suffixes:

---

[9] Note that in YS, 'dog' is [kʼusɨ́kʼusi]; cf. [kʼúsi] 'horse'.

(32)	/pɨ́tja-ɬá/ 'spear'-agentive	[pɨtjaɬá]	'spearer'
	cf. [ptjaw]		'mink'
	/amɨt-mí/ 'grey squirrel'-genitive	[amɨtmí tánawit]	'the grey squirrel's den'
	cf. [pamt]		'nephew'
	/pajúwi-t-pamá/ 'be sick'-gerundive-'for, concerning'	[pajuwitpamá]	'hospital'
	cf. [ɨwájwiʃ]		'collar, necklace'

We suggested that the presence of the destressed high vowels in the examples in (32) could be accounted for if they carried a secondary stress. Since unstressed high vowels would basically be deleted in analogous contexts, these data suggest that the difference between secondarily stressed and unstressed vowels may involve duration (i.e. some vs. no duration).

The phonetic correlates of the possible secondary stresses noted by Rude 1988, Rigsby and Rude 1996, and Hargus and Beavert 2002a remain to be investigated. Although the experiment reported in §4 showed that duration is not a phonetic correlate of primary stress, the phenomenon mentioned above suggests that duration may be relevant for secondary stress. In future experimental work on Sahaptin stress, it would be good to test the phonetic correlates of secondary stress relative to both primary stressed and unstressed syllables. However, in practice it may be impossible to construct examples which control for all confounding factors. Note that if secondary stresses can be shown to exist in YS, this would set YS apart from canonical pitch-accent languages, where secondary stresses are not reported.

References

Alderete, John. 1999. *Morphologically Governed Accent in Optimality Theory,* Department of Linguistics, University of Massachusetts, Amherst: Ph.D.

Beavert, Virginia and Hargus, Sharon. in preparation-a. Ichishkíin Sɨ́nwit. Yakima Sahaptin Bilingual Dictionary: Sahaptin-English and English-Sahaptin. 512. Seattle: University of Washington

Beavert, Virginia and Hargus, Sharon. in preparation-b. *Yakima (Yakama) Sahaptin Bilingual Dictionary: Sahaptin-English and English-Sahaptin*. Seattle: University of Washington Press.

Beckman, Mary E. 1986. *Stress and Non-stress Accent*. Dordrecht: Foris.

Clements, G.N. and Goldsmith, John. 1984. Autosegmental Studies in Bantu Tone: Introduction. *Autosegmental Studies in Bantu Tone,* ed. by John Goldsmith. Dordrecht: Foris.

de Lacy, Paul. 2002. The Interaction of Tone and Stress in Optimality Theory. *Phonology,* 19.1-32.

Fry, D.B. 1955. Duration and intensity as physical correlates of linguistic stress. *Journal of the Acoustical Society of America,* 27.765-68.

Gomez-Imbert, Elsa. 2001. More on the Tone versus Pitch Accent Typology: Evidence from Barasana and Other Eastern Tukanoan Languages. *Proceedings of the Symposium "Cross-Linguistic Studies of Tonal Phenomena, Tonogenesis, Japanese Accentology, and Other Topics,* ed. by Shigeki Kaji, 369-412. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa, Tokyo University of foreign Studies.

Hargus, Sharon. 2001. Quality-sensitive stress reconsidered. *University of Washington Working Papers in Linguistics,* 20.25-56.

Hargus, Sharon. 2005. Prosody in two Athabaskan languages of northern British Columbia. *Athabaskan Prosody,* ed. by Keren Rice, 393-423. Amsterdam: John Benjamins.

Hargus, Sharon and Beavert, Virginia. 2001. Initial clusters and minimality in Yakima Sahaptin. *University of Washington Working Papers in Linguistics,* 20.1-24.

Hargus, Sharon and Beavert, Virginia. 2002a. Predictable vs. underlying vocalism in Yakima Sahaptin. *International Journal of American Linguistics,* 68.316-40.

Hargus, Sharon and Beavert, Virginia. 2002b. Yakima Sahaptin clusters and epenthetic [ɨ]. *Anthropological Linguistics,* 44.1-47.

Hargus, Sharon and Beavert, Virginia. 2006. Word-initial clusters and minimality in Yakima Sahaptin. *Phonology,* 23.21-58.

Hayes, Bruce. 1995. *Metrical Stress Theory: Case Studies and Principles*. Chicago: University of Chicago Press.

Hoijer, Harry. 1943. Pitch Accent in the Apachean Languages. *Language,* 19.38-41.

Hulst, Harry van der and Smith, Norval. 1988. The Variety of Pitch Accent Systems: Introduction. *Autosegmental Studies on Pitch Accent,* ed. by Norval Smith, ix-xxiv. Dordrecht: Foris.

Hyman, Larry. 1977. On the nature of linguistic stress. *Studies in Stress and Accent,* ed. by Larry Hyman.

Hyman, Larry. 2001. Tone systems. *Language Typology and Language Universals,* ed. by Wolfgang Raible, 1367-80. Berlin: Mouton.

Ladefoged, Peter. 1992. *A Course in Phonetics*. New York: Harcourt Brace Jovanovich.

Lehiste, Ilse. 1970. *Suprasegmentals*. Cambridge: MIT Press.

Maddieson, Ian. 1985. Phonetic Cues to Syllabification. *Phonetic Linguistics,* ed. by Vicki Fromkin, 203-21. San Diego: Academic Press.

McCawley, James D. 1978. What is a tone language? *Tone: A Linguistic Survey,* ed. by Victoria Fromkin, 113-31. New York: Academic Press.

Odden, David. 1999. Typological issues in tone and stress in Bantu. *Cross-Linguistic Studies of Tonal Phenomena: Tonogenesis, Typology, and Related Topics,* ed. by Shigeki Kaji, 187-215. Tokyo: Institute for the Study of Languages and Cultures of Africa and Asia, Tokyo University of Foreign Studies.

Rigsby, Bruce and Rude, Noel. 1996. Sketch of Sahaptin, a Sahaptian Language. *Languages,* ed. by Ives Goddard, 666-92. Washington DC: Smithsonian Institution.

Rozelle, Lorna. 1997. The Effect of Stress on Vowel Length in Aleut. *University of California Working Papers in Phonetics,* 95.91-101.

Rude, Noel. 1988. A report from the field: Department of Linguistics, University of Oregon

Sapir, Edward. 1925. Pitch Accent in Sarcee, an Athabaskan Language. *Journal de la Société des Américanistes de Paris,* 17.185-205.

Sluijter, Agatha Martha Cornelia. 1995. *Phonetic Correlates of Stress and Accent*. The Hague: Holland Academic Graphics.

Taff, Alice, Rozelle, Lorna, Cho, Taehong, Ladefoged, Peter, Dirks, Moses and Wegelin, Jacob. 2001. Phonetic Structures of Aleut. *Journal of Phonetics,* 29.231-71.

Voorhoeve, Jan. 1973. Safwa as a restricted tone system. *Studies in African Linguistics 4,* 1.1-22.

Yip, Moira. 2002. *Tone*. Cambridge: Cambridge University Press.

Zhivov, V. 1978. Restricted Tone Systems versus Pitch Accent Systems. *Estonian Papers in Phonetics*.99-101.