Sequence Structure

D. Terence Langendoen

Department of Linguistics, University of Arizona

This paper examines the extent to which the structure of the expressions of a language, viewed, following Harris (1946), as a set of morphemes and morpheme sequences, can be determined by the subsequence relations they bear to one another.* I define a sequence structure as a set of morphemes and sequences of those morphemes over which the subsequence relation (a partial order) is defined. Assuming that a sequence is sequentially ambiguous if and only if it is the conjunction of two or more pairs of maximal subsequences, I show that sequential ambiguity can be made to model structural ambiguity as it is ordinarily understood, although in some cases adjustments have to be made in the kinds of sequences that are considered part of the language. I also show that sequence structures can model structural relations for which movement transformations have been thought to be necessary. Finally, I discuss the problem of determining the interpretations of sequences in a sequence structure.

1. Introduction

Harris (1946) anticipated current interest in the avoidance of unnecessary constructs in linguistic theory when he wrote:

> [T]here is an advantage in avoiding [constructs such as 'morphological levels'] if we can achieve the same results by direct manipulation of the observable morphemes. The method described in this paper will require no elements other than morphemes and sequences of morphemes, and no operation other than substitution, repeated time and again.

Assuming with Harris that a language consists of morphemes and sequences of morphemes, I investigate the structures of morpheme sequences in a language as determined by their relation to their subsequences which also belong to that language.

2. Fundamental notions

In this section, I define and illustrate the notions of subsequence, maximal (proper) subsequence, sequence structure, and conjunction of sequences.

2.1. Subsequences

Let L be a set of sequences of morphemes.[1] Then $s \in L$ has as its parts any subsequence $r \in L$, where the subsequence relation $\geq_S$ is defined as in (1). $\geq_S$ is reflexive, antisymmetric, and transitive; i.e., it is a partial order.

(1) For all $r, s \in L$, ***r* is a subsequence of *s* in *L*** ($s \geq_S r$) if and only if there is an $n > 0$ such that there are sequences $r_1, \ldots, r_n \in L$, and $x_0, \ldots, x_n$ such that $r = r_1 \ldots r_n$ and $s = x_0 r_1 x_1 \ldots r_n x_n$.

In this paper, I make three additional assumptions. First, I assume that L contains every morpheme that appears in any sequence in L. Without this assumption, one could consider L to be a set of words, possibly polymorphemic, which does not contain some of the morphemes (i.e. the affixes and bound roots) that appear in those words. Second, I assume the stronger version of the subsequence relation $\geq_Q$ in (2).

(2) For all $r, s \in L$, ***r* is a strict subsequence of *s* in *L*** ($s \geq_Q r$) if and only if there is an $n > 0$ such that there are sequences $r_1, \ldots, r_n, x_1, \ldots, x_{n\text{-}1} \in L$, and x_0, x_n either $\in L$ or empty, such that $r = r_1 \ldots r_n$ and $s = x_0 r_1 x_1 \ldots r_n x_n$.

By requiring that $x_1, \ldots, x_{n\text{-}1}$ also belong to L, $\geq_Q$ is not transitive over every set of morpheme sequences. For example, let $L_0 = \{a, b, c, d, ad, bd, cd, abd, acd, bad, bcd, cad, cbd, abcd, acbd, bacd, bcad, cabd, cbad\}$. Then $abcd \geq_Q acd$, and $acd \geq_Q ad$, but $abcd \not\geq_Q ad$, since $bc \notin L_0$. Hence $\geq_Q$ is not transitive and therefore not a partial order on L_0. My third assumption is that the sets of morpheme sequences of linguistic significance include only those for which $\geq_Q$ is a partial order. Henceforth by “subsequence” I mean “strict subsequence”.

2.2. Sequence structures

Let L_1 = {*cure*, *able*, *ity*, *curable*, *curability*}, where *cure*, *able*, and *ity* are morphemes. Each morpheme has only itself as a subsequence, whereas *curable* has *cure* and *able*, as **proper** subsequences, and *curability* has *curable*, (and hence also *cure* and *able*) and *ity*, as proper subsequences. A proper subsequence *r* of a sequence *s* is **maximal** if *s* has no other proper subsequence of which *r* is a proper subsequence. For example, *curable* is a maximal proper subsequence of *curability*, but *cure* and *able* are not.[2] and it induces over L_1 the inverse hierarchical structure Q_1 diagrammed in Figure 1, in which reflexive arcs and arcs derivable from transitivity are omitted. I call such a structure a **sequence structure**.

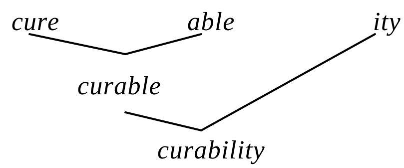

Figure 1. Sequence structure Q_1

2.3. Conjunction of sequences

Since $\geq_Q$ is a partial order, the greatest lower bound of any two sequences in a sequence structure, if it exists, is their **conjunction**. For example, the conjunction of *cure* with *able* in Q_1 is *curable*, and the conjunction of *curable* with *ity* is *curability*. In these examples, the conjunction operator is equivalent to concatenation (in a certain order). However for other examples, it is not; for example, the conjunction of *curable* with *able* is *curable*, of *cure* with *ity* is *curability*, and of *able* with *ity* is also *curability*.

The sequence structure Q_1 is that of L_1 as a whole, not just of the root morpheme sequence *curability*. To see this more clearly, let us add to L_1 the morphemes *prove*, *sane*, and *mouse*, and the morpheme sequences *sanity*, *provable* and *provability*. The sequence structure Q_2 of the resulting set L_2 is shown in Figure 2. In Q_2, the conjunctions of the elements in Q_1 remain the same, except that of *able* with *ity*, which have no conjunction.[3] Both *curability* and *provability* are candidates (both have *able* and *ity* as subsequences), but neither is a subsequence of the other (i. e. there is no **greatest** lower bound for *able* and *ity* in L_2). In addition, *sane* and *able* have no conjunction in Q_2, there being no sequence which has both *sane* and *able* as subsequences.[4] Q_2 is a multiple (or multiply rooted) inverse hierarchy.

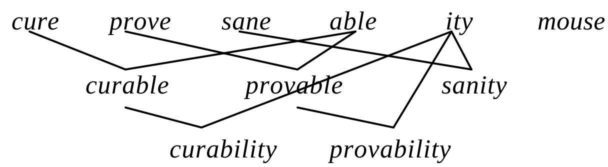

Figure 2. Sequence structure Q_2

Sequence structures can also involve classes of morphemes. An example is Q_3 in Figure 3, which is based on L_3 = {*V*, *A*, *N*, *A**V*, *N**A*, *V A**V*, *A N**A*, *V A**V N**A*}, derived from L_2 by setting *V* = {*cure*, *prove*}, *A* = {*sane*}, *N* = {*mouse*}, *A**V* = {*able*}, and *N**A* ={*ity*}. Sequences with the same potential for entering into longer sequences can be established by cancellation in the usual manner, so for example the sequences *A* and *V A**V* are combinatorially equivalent, both combining with *N**A*, etc.

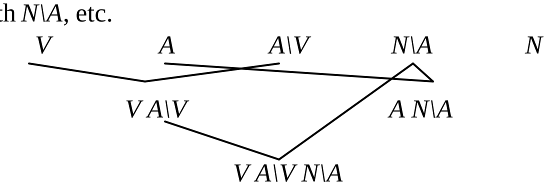

Figure 3. Sequence structure Q_3

2.4. *Mergers of sequences*

The relation $\geq_Q$ can be generalized to allow for any number of antecedent sequences. For this purpose, I define a **merger** of two sequences as in (3).

(3) The sequence *u* is a merger of sequences *s*, *t* in *L* if $u \geq_Q s$, $u \geq_Q t$, and every morpheme in *u* also appears in either *s* or *t*.

Then $s_1, \ldots, s_n \geq_Q r$ if *r* is a subsequence of a merger of $s_1, \ldots, s_n$. It follows that if *u* is a merger of *s* with *t*, then *u* is also the conjunction of *s* with *t* provided that there is no other sequence *v* in *L* which is a merger of *s* with *t*.

2.5. *Subsequences versus substrings*

The sequence structures Q_1 through Q_3 are compatible with a stronger mereological relation than subsequence: the substring relation $\geq_R$ defined in (4).

(4) For all $r, s \in L$, *r* is a **substring** of *s* ($s \geq_R r$) if and only if there are sequences x_0, x_1 such that $s = x_0 r x_1$.

However, let L_4 = {*the*, *only*, *point*, *the point*, *the only point*}. In L_4, *the point* occurs as a subsequence of *the only point*, but not as a substring. Assuming that *the point* is correctly analyzed as a part of *the only point*, as in the sequence structure Q_4 in Figure 4, then the weaker subsequence relation is the desired mereological relation for linguistic analysis. In Q_4, *the only point* is the conjunction of *the point* with *only*.

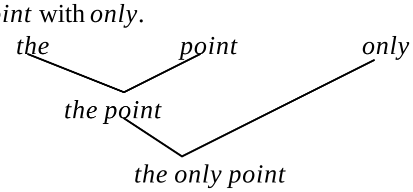

Figure 4. Sequence structure Q_4

3. Sequential ambiguity

In this section, I define a notion of sequential ambiguity which can be made to match that of structural ambiguity as defined for constituent structures. A sequence *s* is **sequentially unambiguous** in *Q* if it is the conjunction of at most one pair {*t*, *u*} of maximal subsequences in *Q*; otherwise it is **sequentially ambiguous**. Every sequence in Q_1 through Q_4 is sequentially unambiguous. Similarly, every sequence in the sequence structure Q_5 of L_5 = {*very*, *nice*, *result*, *very nice*, *nice result*, *very nice result*} in Figure 5 is sequentially unambiguous, including *very nice result*, since it is the unique conjunction of its maximal subsequences *very nice* and *nice result*. The fact that *very nice result* is also the conjunction of *very* with *nice result* and of *very nice* with *result* is irrelevant, since these pairs are not maximal subsequences of *very nice result*.

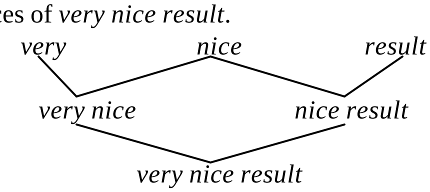

Figure 5. Sequence structure Q_5

On the other hand, the sequence *old clothing store* in the sequence structure Q_{6a} of L_{6a} = {*old*, *clothing*, *store*, *old clothing*, *clothing store*, *old store*, *old clothing store*} in Figure 6 is sequentially ambiguous, since it is the conjunction of three different pairs of its maximal subsequences *old clothing*, *old store* and *clothing store*.

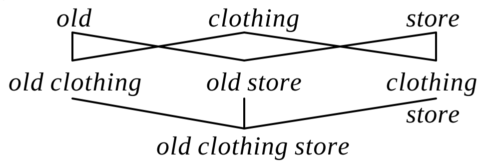

Figure 6. Sequence structure Q_{6a}

Viewing the sequence *old clothing store* as the conjunction of the maximal subsequence *old clothing* with *clothing store* corresponds to the constituent structure [[*old clothing*] *store*], whose interpretation is that the clothing in the store is old, but not necessarily the store. Viewing it as the conjunction of *old store* with *clothing store* corresponds to the structure [*old* [*clothing store*]] whose interpretation is that the store is old, but not necessarily the clothing in it. However the conjunction of *old clothing* with *old store* does not correspond to any interpretation of *old clothing store* in English. If it did, it would be to one in which both the clothing and the store (not necessarily a clothing store) are old. To account for the fact that *old clothing store* is only two ways, and not three ways sequentially ambiguous in English, we observe that the members of L_{6a} are part of a language that also includes the sequences *store clothing* and *old store clothing*, with the sequence structure Q_{6b} in Figure 7. In Q_{6b}, *old clothing* and *old store* have two different mergers, *old clothing store* and *old store clothing* (the latter shown with dashed lines), so that neither one is the conjunction of *old clothing* with *old store*. The sequence *old clothing store* is therefore the conjunction of only the two pairs of maximal subsequences, *old clothing* and *clothing store*, and *old store* and *clothing store*, as desired.[5]

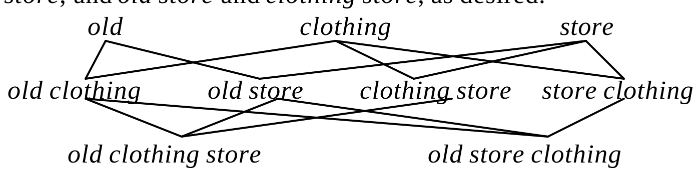

Figure 7. Sequence structure Q_{6b}

This correlation between sequential ambiguity and structural ambiguity extends to more complex cases, such as that of Q_{6d} of L_{6d}, which includes all of the sequences in L_{6b} together with those in L_{6c} = {*fine*, *fine clothing*, *fine store*, *fine clothing store*, *fine store clothing*, *fine old clothing*, *fine old store*, *old fine clothing*, *old fine store*, *fine old clothing store*, *fine old store clothing*, *old fine* *clothing store*, *old fine store clothing*}. Figure 8 shows that in Q_{6d} the sequence *fine old clothing store* is four-ways sequentially ambiguous, being the conjunction of the following four pairs of maximal subsequences: (1) *fine old clothing* and *fine clothing store* (corresponding to the constituent structure [[*fine* [*old clothing*]] *store*]); (2) *fine old clothing* and *old clothing store* (corresponding to the constituent structure [[[*fine old*] *clothing*] *store*]); (3) *fine old store* and *fine clothing store* (corresponding to the constituent structure [*fine* [*old* [*clothing store*]]]); and (4) *fine old store* and *old clothing store* (corresponding to the constituent structure [[*fine old*] [*clothing store*]]). It is not the conjunction of *fine clothing store* with *old clothing store*, since *old fine clothing store* is also a merger of those maximal subsequences (shown with dotted lines). Nor is it the conjunction of *fine old clothing* with *fine old store*, since *fine old store clothing* is also a merger of those maximal subsequences (shown with dashed lines). On interpretations (1) and (2), the clothing is fine and old, but not necessarily the store; on interpretations (3) and (4), the store is fine and old, but not necessarily the clothing. The sequence lacks an interpretation corresponding to the constituent structure [*fine* [[*old clothing*] *store*]], in which the store but not necessarily the clothing is fine, and in which the clothing but not necessarily the store is old. But this is correct, since that interpretation is only possible if there is a clear juncture between *fine* and *old*, and again between *clothing* and *store*, i.e. if the sequence contains at least one additional "juncture" morpheme.

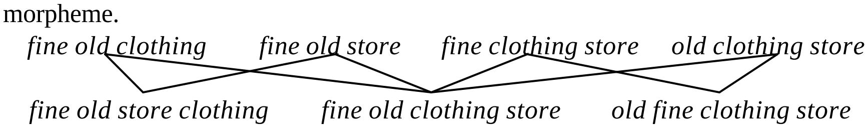

Figure 8. Part of sequence structure Q_{6d}

3.1. The need for nonstandard morpheme sequences

In order for subsequence ambiguity to match structural ambiguity in some cases, certain morpheme sequences that are not standardly assumed to be part of a language must be posited. For example, the structural ambiguity of the sequence *old men and women* in English is not matched by sequential ambiguity in the sequence structure Q_{7a} of L_{7a} = {*and*, *men*, *women*, *old*, *and men*, *and women*, *old men*, *old women*, *and old men*, *and old women*, *men and women*, *women and men*, *old men and women*, *old women and men*, *men and old women*, *women and old men*}. Figure 9 shows that *old men and women* is sequentially unambiguous in

Q_{7a}; it is the conjunction of the maximal subsequences *old men* and *men and women* only.

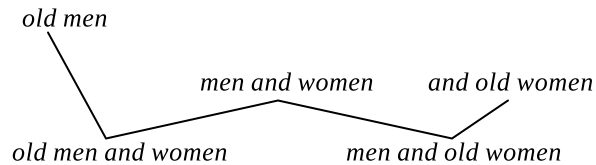

Figure 9. Part of Q_{7a} showing lack of sequential ambiguity of *old men and women*

To get the result that *old men and women* is sequentially ambiguous, we must replace in L_{7a} the sequences {*and men*, *and women*, *and old men*, *and old women*} that have been standard in most recent analyses of English with the nonstandard sequences {*men and*, *women and*, *old men and*, *old women and*}. Figure 10 shows that in the resulting sequence structure Q_{7b}, *old men and women* is sequentially ambiguous, being the conjunction of the maximal subsequence *old men and* with *men and women* (the interpretation being that of the constituent structure [[old men] and women], in which the men are old, but not necessarily the women), and of *old men and* with *old women* (the interpretation being that of the constituent structure [old [men and women]], in which both the men and the women are old). However, *old men and women* is not the conjunction of *old women* with *men and women*, since both *old men and women* and *men and old women* are mergers of those maximal subsequences (the latter shown with dashed lines).

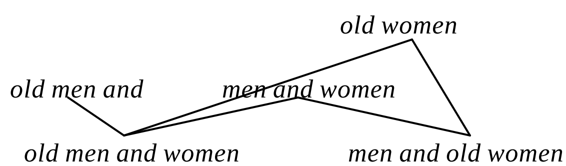

Figure 10. Part of Q_{7b} showing sequential ambiguity of *old men and women*

3.2. Sequential vs. morphemic (lexical) ambiguity

Sequential ambiguity must be distinguished from morphemic (lexical) ambiguity just as structural ambiguity must. In the sequence structure Q_8 in Figure 11, based on L_8 = {*un*¹, *un*², *ed*, *pack*, *un*²*pack*, *packed*, *un*¹*packed*, *un*²*packed*}, neither *un*¹*packed* nor *un*²*packed* is sequentially ambiguous. The former is the conjunction of the maximal subsequences *un*¹ and *packed*; the latter of *un*²*pack* and *packed*.

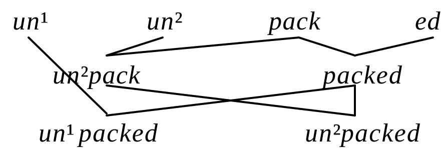

Figure 11. Sequence structure Q_8

The ambiguity of expressions like *Irish grammarian* in Q_9 of L_9 = {*Irish*^{1}, *Irish*^{2}, *grammar*, *ian*} in Figure 12 is similarly explained. On the interpretation ‘grammarian who is Irish’, it is the conjunction of the maximal subsequences *Irish*^{1} (referring to a person) and *grammarian*; whereas on the “bracketing paradox” interpretation ‘student of Irish grammar’, it is the conjunction of the maximal subsequences *Irish*^{2} *grammar* and *grammarian* (*Irish*^{2} referring to a language).

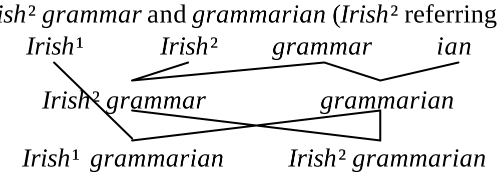

Figure 12. Sequence structure Q_9

3.3. Repetition of morphemes

A morpheme can occur more than once in a morpheme sequence, as in *I think I drink* in L_{10a} = {*I*, *think*, *drink*, *I think*, *I drink*, *I think I drink*}, in which the morpheme *I* occurs twice. In order to represent the sequence structure of sets containing such sequences, the two occurrences of the morpheme must be distinguished, say by indices, with the convention that subsequent occurrences have higher indices. For example, L_{10a} may be replaced by L_{10b} = {I_1, I_2, *think*, *drink*, I_1 *think*, I_2 *think*, I_1 *drink*, I_2 *drink*, I_1 *think* I_2 *drink*}, which has the sequence structure Q_{10b} in Figure 13. The latter is the conjunction of the maximal subsequences I_1 *think* and I_2 *drink* only, and so is sequentially unambiguous.

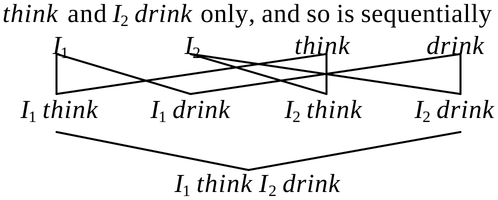

Figure 13. Sequence structure Q_{10b}

4. Modeling transformational relations using sequence structures

Sequence structures can also model mereological relations for which movement transformations have been thought necessary.

4.1. Apparent "local movement": Subject-Auxiliary inversion in English

For example the sequence *can tan* is assumed to be part of *can Fran tan* in English, but is not considered a constituent of it. Given that it is a constituent of *Fran can tan*, the sentence *can Fran tan* is generally considered to be derived from it by the "movement" of the auxiliary verb *can* around the subject noun *Fran*. However *can tan* is a subsequence of *can Fran tan* just as much as it is of *Fran can tan*. No movement of *can* around *Fran* is necessary to identify *can tan* as part of *can Fran tan* in the appropriate sequence structure. Specifically in Q_{11}, the sequence structure of L_{11} = {*Fran*, *can*, *tan*, *Fran can*, *can Fran*, *can tan*, *Fran can tan*, *can Fran tan*} in Figure 14, *can Fran tan* is the conjunction of the maximal subsequences *can Fran* and *can tan*, while *Fran can tan* is the conjunction of the maximal subsequences *Fran can* and *can tan*.

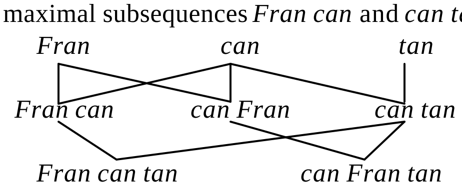

Figure 14. Sequence structure Q_{11}

4.2. Apparent "unbounded movement": Wh-*movement in English*

The "movement" of *can* in *can Fran tan* does not reorder the subsequences of *can tan*, so that *can tan* is, as has already been observed, a subsequence of *can Fran tan*. However the "movement" of *what* (an instance of *wh*-movement) in *what will Phil mill* does reorder the subsequences of *mill what*, which allegedly occurs as part of *what will Phil mill*. As a result, *mill what* is not a subsequence of *what will Phil mill*, so that subsequence structure does not provide a model of *wh*-movement in English as it is usually analyzed in generative grammar.

The model of *wh*-movement that emerges from sequence structure analysis is one in which the "moved" *wh*-element appears at the left edge of successively longer sequences, including eventually the verb of which it is an argument (if it is one), as in the sequence structure Q_{12} of L_{12} = {*Phil*, *will*, *mill*, *what*, *Phil will*, *what will*, *will Phil*, *will mill*, *mill what*, *Phil will mill*, *will Phil mill*, *will mill what*, *what will Phil*, *Phil will mill what*, *what Phil will mill*, *what will Phil mill*} in Figure 15. In Q_{12}, the shortest sequence containing "moved" *what* is *what will*. The next longer sequences *what Phil will* and *what will Phil* are the conjunctions of *what will* with *Phil will* and *will Phil* respectively. Finally the sequences *what Phil will mill* (a subsequence of *I know what Phil will mill*, in an extension of L_{12}) and *what will Phil mill* are the conjunctions of *what Phil will* with *Phil will mill* and of *what will Phil* with *will Phil mill*, respectively.[6]

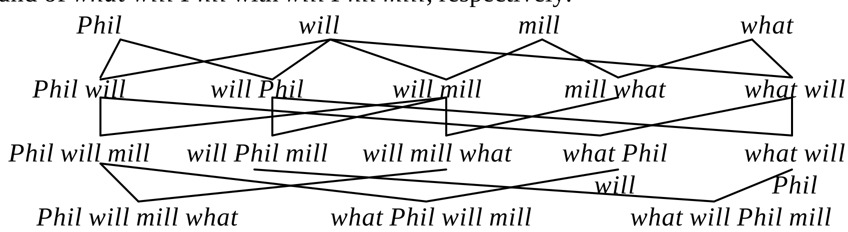

Figure 15. Sequence structure Q_{12}

In cases in which a *wh*-element is "extracted" from a complement, that element will be part of a sequence expressing exactly that complement, so that the grammatical relation of the "extracted" element to other elements in the complement can be determined in a sequence structure. For example, the sequence *who may Fay say Phil will thrill* is analyzed in the appropriate sequence structure *Q* as the conjunction of the maximal subsequences *who may Fay say* and *who Phil will thrill*, so that the grammatical relation of *who* with *thrill* in *who may Fay say Phil will thrill* can be determined. In addition, the fact that certain elementary sequences containing two *wh*-elements do not belong to the sequence structure directly accounts for the absence of longer such sequences, without appeal to "constraints" on movement. For example, the absence of the elementary sequence **who what will drill* in *Q* accounts for the absence of **who may Fay say what will drill*, since the latter would have to be analyzed as the conjunction of the maximal subsequence *who may Fay say* with *who what will drill* in *Q*. On the other hand, if the sequence *what who will drill* belongs to *Q*, then we may expect that the sequence *what may Fay say who will drill* will also belong to *Q*, since the latter is the conjunction of *what may Fay say* with *what who will drill*.

5. Determining the interpretations of sequence structures

The interpretation of a sequence *s* is fully determined by its place in a sequence structure *Q*. If *s* is a single morpheme *m*, then its interpretation is simply that which is assigned to *m*. If *s* is a sequence of two or more morphemes, then either it is (1) not the conjunction of any pair of maximal subsequences $\{t, u\}$, (2) the conjunction of exactly one such pair, or (3) the conjunction of more than one such pair. Case (1) arises when for every pair $\{t, u\}$ of maximal subsequences of *s*, there is another sequence *r* in *Q* which is a merger of $\{t, u\}$. In this case, the interpretation of *s* is a function of the interpretation of each pair of its maximal subsequences. The simplest subcase of case (1) arises when $s = tu$ and $r = ut$, as in Q_{6d}, where *s* = *clothing store*. In Q_{6d}, the sequences of which *s* and *r* are maximal subsequences do not fall under case (1); i.e. in Q_{6d}, the property of failing to be the conjunction of maximal subsequences is not inherited by the sequences of which the sequences with that property are maximal subsequences. (Both *old clothing store* and *old store clothing* are conjunctions of pairs of maximal subsequences in Q_{6d}.) The property is only inherited under very specific circumstances, as in the hypothetical sequence structure Q_{13} in Figure 16. In Q_{13}, neither *ab* and *ba*, nor *bc* and *cb* are conjunctions of their maximal subsequences *a* and *b*, and *b* and *c* respectively. In addition, neither *bac* nor *bca* are conjunctions of their maximal subsequences *ba* and *bc*; i.e., both *bac* and *bca* inherit the property. On the other hand, *abc*, *acb*, *cab*, and *cba* are conjunctions of their respective maximal subsequences; i.e., they don't inherit the property.

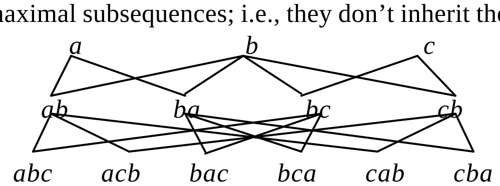

Figure 16. Sequence structure Q_{13}

In addition, I am not aware of any natural language examples in which a sequence is the merger of more than one pair of maximal subsequences, and that sequence is the conjunction of none of those pairs. If that restriction indeed holds, then all instances of case (1) are sequentially unambiguous.

Case (2) has two subcases: (a) when a sequence has exactly one pair of maximal subsequences, and (b) when it has more than one such pair, but it is the conjunction of the members of exactly one of them. The sequence *very nice result* in Q_5 is an instance of case (2a); no instances of case (2b) are given here.

Case (3) also has two subcases: (a) when a sequence is the conjunction of every pair of its maximal subsequences (there being at least three such pairs), and (b) when it is the conjunction of more than one pair of its maximal subsequences, but not of every pair. Case (3a) does not arise in any of the examples we have considered, and I believe does not arise at all in natural languages. The sequences *old clothing store* in Q_{6d} and *old men and women* in Q_{8b} are instances of case (3b). In this case the sequence has as many interpretations as the number of pairs of maximal subsequences of which it is the conjunction; in each case the interpretation is a function of the interpretations of those maximal subsequences.

The degree of sequential ambiguity of any case (3b) sequence is determined simply by the number of pairs of maximal subsequences of which it is the conjunction. The sequential ambiguity of those maximal subsequences does not enter into the calculation. For example, on one of its interpretations, the sequence *fine old clothing store* in Q_{7d} is the conjunction of the pair of maximal subsequences *fine old clothing* and *old clothing store*. The fact that the latter is itself sequentially ambiguous does not matter. In other words, given an analysis of a sequence into a pair of maximal subsequences of which it is the conjunction, each member of that pair makes a univocal contribution to the interpretation of that sequence, even if it is itself sequentially ambiguous.

6. Concluding remarks

For the past several years, I have been exploring the applicability to linguistic analysis of Koslow's (1992) notion of an "implication structure," which consists solely of a set *S* and an implication relation (a partial ordering) defined over *S*. It had occurred to me that the substring relation induces an implication structure over a set of strings, but that structure turns out not have quite the right properties for general linguistic analysis. At first, when I considered the passage in Harris (1946) quoted at the beginning of this essay, I did not realize that the subsequence relation is distinct from the substring relation. It was when I realized that they are distinct, and that the subsequence relation is the weaker of the two, that I discerned that the structure that the subsequence relation induces over a set of sequences (or strings, it does not matter what you call them) does have properties suitable for linguistic analysis. I hope to explore the matter further in future work.

References

Harris, Zellig S. 1942. Morpheme alternants in linguistic analysis. *Language* 18 169-180.

Harris, Zellig S. 1946. From morpheme to utterance. *Language* 22. 161-183.

Koslow, Arnold 1992. *A Structuralist Theory of Logic*. Cambridge: Cambridge University Press.

Notes

[*] Prepublication version of paper published in *The Legacy of Zellig Harris: Language and Information into the 21st Century*, vol. 2: *Computability of Language and Computer Applications*, ed. by Bruce Nevin and Stephen M. Johnson, 61-75. Amsterdam: John Benjamins, 2002.

[1] For our purposes, a single morpheme counts as a morpheme sequence. However, *L* does not contain the "empty" sequence λ, though it can contain morphemes with "zero allomorphs".

[2] Henceforth I drop the term "proper" in describing proper maximal subsequences.

[3] However, if the sequence *ability* is added to L_2, then the conjunction of *able* with *ity* is *ability*.

[4] For justification of this view of conjunction, see Koslow (1992).

[5] In Q_{6b} the conjunction of *clothing* with *store* is also undefined, since both *clothing store* and *store clothing* are mergers of *clothing* with *store*. However, both *clothing store* and *store clothing* are interpretable sequences in Q_{6b}, for the simple reason that there is no conjunction at all of their maximal subsequences *clothing* and *store*. It is only if a sequence has an analysis as the conjunction of maximal subsequences that nonconjunctive mergers of maximal subsequences are ignored.

[6] Again certain nonstandard sequences must be postulated in order for the analysis to work.