COORDINATE GRAMMAR

D. Terence Langendoen

University of Arizona
and
National Science Foundation

Chomsky (1959a) presented an algorithm for constructing a finite transducer that is strongly equivalent to a Chomsky-normal-form context-free grammar for all sentences generated by that grammar with up to any specified finite degree of center embedding. This article presents a new solution using a variety of COORDINATE GRAMMAR to assign nonembedding (paratactic) structures strongly equivalent to those assigned by an embedding grammar, which can in turn be directly computed by a finite transducer. It proposes that the bound on center embedding is really a consequence of a bound on alternation between right and left embedding, called here ZIGZAG EMBEDDING. Coordinate grammars can also be used to assign nonembedding structures equivalent to those with up to any specified finite degree of coordinate embedding (the occurrence of a coordinate structure as a member of a coordinate structure of the same type). It concludes that coordinate grammars or the finite transducers strongly equivalent to them are psychologically real, and that the existence of a finite bound on the degree of zigzag and coordinate embedding is a consequence of the increasing size and complexity of such grammars or transducers as the bound increases.*

1. INTRODUCTION: WEAK EQUIVALENCE BETWEEN FINITE-STATE AUTOMATA AND NONCENTER-EMBEDDING CONTEXT-FREE GRAMMARS.[1] One of the best-known results in formal language theory is that if there is a NONCENTER-EMBEDDING (NCE) CONTEXT-FREE GRAMMAR (CFG) G that generates a language L, then there is a FINITE-STATE AUTOMATON (FA) or regular grammar Q that accepts L; that is, Q is 'weakly equivalent' to G (Chomsky 1959b, 1963:394).[2] Moreover, if every grammar G that generates L is CENTER-EMBEDDING (CE), then for every nonnegative integer i, there is an FA Q^i that generates L^i, the subset of L whose members manifest up to degree i of CE, that is, CE° $\leq i$. Such languages are called STRICTLY CONTEXT FREE and each L^i and Q^i is called

* I began work on the problem discussed here in my undergraduate thesis under Noam Chomsky's direction (Langendoen 1961). I did not solve the problem at the time, and I have returned to it several times over the course of my career. Some of the ideas in this article were presented in my January 1999 LSA presidential address 'Constraints on subordination', and in talks at California State University, Fresno in 2003, the 2005 CUNY Human Sentence Processing conference at the University of Arizona, and the 2007 Maryland Mayfest. This material is based in part on work that was supported while I was serving at the National Science Foundation. Any opinions and conclusions are those of the author and do not necessarily reflect the views of the National Science Foundation.

[1] For ease of reference, I present here a list of acronyms used in this article: CE: center embedding, CE°: degree of center embedding (similarly for CoE°, LE°, LRE°, RE°, RLE°, ZE°, ΨCE°, ΨZE°), CFG: context-free grammar, CNF: Chomsky-normal form, CoE: coordinate embedding, ENFLC: extended normal-form linked coordinate, ENFSC: extended normal-form strictly coordinate, FA: finite automaton, FT: finite transducer, LC: left coordinating, LE: left embedding, LRE: left-then-right embedding, NCE: noncenter embedding, NFC: normal-form coordinate, NFLC: normal-form linked coordinate, NFSC: normal-form strictly coordinate, RC: right coordinating, RE: right embedding, RLE: right-then-left embedding, ZE: zigzag embedding, ΨCE: pseudo-center embedding, ΨZE: pseudo-zigzag embedding.

[2] Center embedding is also known as 'self embedding'. A CFG is center embedding (CE) if it supports subderivations of the form $A \Rightarrow \varphi\ A\ \psi$, where φ and ψ are non-null. It is right embedding (RE) if it supports subderivations of the form $A \Rightarrow \varphi\ A$, where φ is non-null. It is left embedding (LE) if it supports subderivations of the form $A \Rightarrow A\ \psi$, where ψ is non-null. Right embedding and left embedding are also known as 'right recursive' and 'left recursive' respectively.

a finite-state (or regular) 'approximation' to L and G respectively (Nederhof 2000). The fact that otherwise well-formed structures with CE° greater than some small n are unacceptable in any language has been interpreted in at least two different ways. The prevailing view, following Miller and Chomsky (1963), is that such structures are grammatical, but unacceptable because of linguistically irrelevant limitations on working memory. Another view, expressed by Krauwer and des Tombe (1979), is that such structures are ungrammatical, because the bound on CE° is built into the theory of grammar. One of the results of this article is that such a bound is not an arbitrary stipulation in such a theory.

The rest of this article is organized as follows. I first discuss and illustrate Chomsky's original notion of strong equivalence between finite transducers and NCE CFGs and point out an inadequacy with this notion. I then develop the heart of this article, introducing a new notion of strong equivalence, and providing a series of procedures that lead to the construction of finite-state transducers that are strongly equivalent to CE CFGs up to any desired bound on the degree of 'zigzag' embedding, which is a more refined notion of structural complexity than center embedding. I go on to show why degree of zigzag embedding is a more effective measure of structural complexity than center embedding, and then argue that finite-state transducers of the sort constructed by the procedures developed earlier provide better models of human knowledge of language than computationally more powerful devices such as context-free grammars. Finally, I return to the notion of coordinate embedding defined earlier and show that the bound on coordinate embedding can be explained in the same way that the bound on zigzag (or center) embedding can.

2. Strong equivalence between finite-state transducers and cfgs. Less well known than the result in §1, but linguistically more significant, is the fact that given an NCE chomsky-normal-form (CNF) CFG G, there is a finite-state transducer (FT) T that is 'strongly equivalent' to G in the following sense.[3] Let Φ be an effective one-to-one mapping of the structural descriptions (parses, in the form of trees or labeled bracketings) of L onto strings in the output vocabulary of T. Then T is 'strongly equivalent' to G if and only if whenever G generates x with structural description y, then T generates the pair $<x, \Phi(y)>$. Furthermore there is an effective procedure (algorithm) Ψ for constructing the weakly equivalent FA Q from G, on which T is based (Chomsky 1959a, 1963:396, definition 9 and theorem 34). Chomsky's characterization of this type of strong equivalence is diagrammed in 1.

(1) $\mathrm{G} \xrightarrow{\Psi} \mathrm{Q} \xrightarrow{\Phi} \mathrm{T}$

Strong equivalence between an NCE CNF CFG G and an FT T according to Chomsky (1963); Q is an FA weakly equivalent to G.

The FT T that Chomsky's algorithm constructs partially traverses depth first and in order the structural descriptions (labeled trees or bracketing structures) that G associates with each x in L, and outputs the sequence of states it goes through in accepting x. Accordingly, if y is a structural description of x with respect to G, then T maps x onto $\Phi(y)$, where $\Phi(y)$ is one of the sequences of states that T goes through in accepting x. The inverse mapping Φ^{-1}, which 'recovers' y from $\Phi(y)$, cannot in general be

[3] The productions of a CNF CFG are of the form $A \rightarrow B\ C$ and $A \rightarrow a$, where A, B, C are nonterminal (category) symbols, and a is a terminal symbol. Chomsky called such grammars 'normal'; the term 'Chomsky normal form' later came to be used to distinguish his normal form definition from others that were proposed around the same time.

carried out by an FT, however, since the set Σ of structural descriptions of a language L generated by an NCE CFG G is not a regular language, unless L is finite. That is, Chomsky's notion of strong equivalence between an NCE CFG and an FT cannot be strengthened to require that the latter generate the set of pairs $<x, y>$ directly, where y is a structural description of x with respect to G.

An example will illustrate. Let G_1 be the NCE CNF CFG with the productions $\Pi(G_1)$ in 2 that generates the regular language L_1 in 3 with the strictly context-free set of structural descriptions Σ_1 in 4. Applying the mapping Ψ to G_1 results in the weakly equivalent FA Q_1 with the transitions $\Pi(Q_1)$ in 5.[4]

(2) $\Pi(G_1)$
  a. S → A B
  b. B → C S
  c. A → they
  d. B → fled
  e. C → said

(3) L_1 = they (said they)* fled

(4) $\Sigma_1 = [_S\ [_A$ they$]$ $([_B\ [_C$ said$]\ [_S\ [_A$ they$])^n\ [_B$ fled$](]])^n]$

(5) $\Pi(Q_1)$

a.	$[S]_1$	→	ε	$[S\ A]_1$
b.	$[S]_2$	→	ε	f
c.	$[S\ A]_1$	→	they	$[S\ A]_2$
d.	$[S\ A]_2$	→	ε	$[S\ B]_1$
e.	$[S\ B]_1$	→	fled	$[S\ B]_2$
f.	$[S\ B]_1$	→	ε	$[S\ B\ C]_1$
g.	$[S\ B]_2$	→	ε	$[S]_2$
h.	$[S\ B\ C]_1$	→	said	$[S\ B\ C]_2$
i.	$[S\ B\ C]_2$	→	ε	$[S]_1$

Equipping Q_1 with an output tape on which to write the sequence of states it goes through in generating L_1 (together with the word w that is output at each transition of the form $[\sigma]_1 \rightarrow w\ [\sigma]_2$) converts it into an FT T_1 that is strongly equivalent to G_1 in the sense of Chomsky's definition. However, since Σ_1 is a strictly context-free language, it is not possible to compose T_1 with another FT U_1 that maps those sequences onto the members of Σ_1 (i.e. so that $T_1(U_1)$ maps every x onto its structural description y with respect to G_1). For example, the output of $\Pi(Q_1)$ in accepting the string in 6 is the sequence in 7, which does not keep track of the number of closing brackets at the end of the structural description of 6 with respect to G_1.

(6) they said they said they fled

(7) $<[S]_1$, $[S\ A]_1$, they, $[S\ A]_2$, $[S\ B]_1$, $[S\ B\ C]_1$, said, $[S\ B\ C]_2$, $[S]_1$, $[S\ A]_1$, they, $[S\ A]_2$, $[S\ B]_1$, $[S\ B\ C]_1$, said, $[S\ B\ C]_2$, $[S]_1$, $[S\ A]_1$, they, $[S\ A]_2$, $[S\ B]_1$, fled, $[S\ B]_2$, $[S]_2$, $f>$

3. Coordinate CFGs. As has just been shown, the drawback of Chomsky's definition of strong equivalence between an NCE CNF CFG G and an FT T is that the output of Φ cannot directly represent the recursive subordinating (hypotactic) structures of the expressions that G generates. To remedy this, I interpose between G and T a CFG J of a type that assigns only coordinating (paratactic) structural descriptions of a sort

[4] The FA notation is that of Chomsky 1959a, 1963, except that spaces have been added between symbols in the state names. The start symbol of Q_1 is $[S]_1$. The symbol ε is the empty string. The symbol f is the 'final' state.

that can be directly computed by *T*. *J* may be called a NORMAL-FORM COORDINATE (NFC) CFG. It is defined like a CNF CFG, but with the differences described in 8.[5]

(8) Properties of an NFC CFG *J* that distinguish it from a CNF CFG

a. The nonterminal vocabulary (set of categories) N of J is partitioned into two subvocabularies: the nonrecursive categories N_n and the coordinate categories N_c. The latter may be empty.

b. The productions involving only members of N_n and of V, the terminal vocabulary, are of the same form as those of a CNF CFG. At least one start symbol of J must belong to N_n. For no A in N_n, however, are there strings φ, ψ such that $A \Rightarrow \varphi\ A\ \psi$ with respect to J.

c. For any D in N_n there may be categories $D_b{}^1$, $D_d{}^2$, $D_b{}^\rho$, $D_d{}^\lambda$, D_a, D_b, D_c, D_d, D_e, and D_f in N_c and the productions in i–iv and v–viii.

i. $\mathrm{D_b}^1 \rightarrow \mathrm{D_a}\ \ \mathrm{D_b}$
ii. $\mathrm{D_b}^1 \rightarrow \mathrm{D_c}\ \ \mathrm{D_b}^\rho$
iii. $\mathrm{D_b}^\rho \rightarrow \mathrm{D_a}\ \ \mathrm{D_b}$
iv. $\mathrm{D_b}^\rho \rightarrow \mathrm{D_c}\ \ \mathrm{D_b}^\rho$
v. $\mathrm{D_d}^2 \rightarrow \mathrm{D_d}\ \ \mathrm{D_e}$
vi. $\mathrm{D_d}^2 \rightarrow \mathrm{D_d}^\lambda\ \mathrm{D_f}$
vii. $\mathrm{D_d}^\lambda \rightarrow \mathrm{D_d}\ \ \mathrm{D_e}$
viii. $\mathrm{D_d}^\lambda \rightarrow \mathrm{D_d}^\lambda\ \mathrm{D_f}$

d. The categories $D_b{}^1$ and $D_d{}^2$ may be additional start symbols of J, or they may replace the category D in the right-hand side of copies of productions of J.

e. Structural descriptions of the expressions generated by J are constructed in the usual way from their derivations, except that the categories introduced by 8c.iii and 8c.iv are daughters of $D_b{}^1$, not of $D_b{}^\rho$, and the categories introduced by 8c.vii and 8c.viii are daughters of $D_d{}^2$, not of $D_d{}^\lambda$. As a result, neither $D_b{}^\rho$ nor $D_d{}^\lambda$, the only potentially embedding (subordinating or hypotactic) categories in J, are constituents of any expression generated by J. Derivations involving $D_b{}^\rho$ 'grow' to the right and the resulting structures may be called RIGHT COORDINATING (RC), and those involving $D_d{}^\lambda$ do so to the left and may be called LEFT COORDINATING (LC). The distinction between RC and LC may not be obvious upon inspection of those structures.

f. In a normal-form STRICTLY COORDINATE (NFSC) CFG, if $D_x \in \{D_a, D_b, D_c, D_d, D_e, D_f\}$, then either $D_x = D$ or there are productions of the form $D_x \rightarrow K\ D$ or $D_x \rightarrow D\ K$, where K is a coordinator.[6]

[5] For some time it was maintained that coordinate structure could not be assigned by a CFG. Chomsky and Miller (1963:298) stated:

> This difficulty [of assigning structure] changes from a serious complication to an inadequacy in principle when we consider the case of true coordination . . . [example omitted]. In order to generate such strings, a constituent-structure grammar must either impose some arbitrary structure (e.g., using a right recursive rule), in which case an incorrect structural description is generated, or it must contain an infinite number of rules. Clearly, in the case of true coordination, by the very meaning of this term, no internal structure should be assigned at all within the sequence of coordinate items.

They also anticipated and rejected the use of rule schemata, later developed in detail to handle the problem (Langendoen 1976, Gazdar et al. 1985:Ch. 8). An early version of the solution proposed here that involves the suppression of structure and does not require rule schemata is presented in Langendoen 1979.

[6] I do not consider here the problem of coordination of 'mixed' categories.

An example of an LC NFSC CFG is the grammar $J_2{}^0$, whose start symbols are S and $S_c{}^2$ and whose productions are listed in 9. $J_2{}^0$ generates the language L_2 in 10, and associates with it the set $\Sigma_2{}^0$ of structural descriptions in 11. Since $J_2{}^0$ is an NFSC CFG, it does not associate any structure containing an LC or RC member with the strings it generates; that is, all the members of $\Sigma_2{}^0$ manifest DEGREE OF COORDINATE EMBEDDING (CoE°) = 0. For example, it does not associate the CoE° = 1 structure in 13 with the string in 12. How to extend $J_2{}^0$ to associate structures with CoE° > 0 is taken up below in §8.

(9) $\Pi(J_2{}^0)$
- a. $\mathrm{S_c}^2 \rightarrow \mathrm{S} \quad \mathrm{S_c}$
- b. $\mathrm{S_c}^2 \rightarrow \mathrm{S_c}^\lambda\ \mathrm{S_c}$
- c. $\mathrm{S_c}^\lambda \rightarrow \mathrm{S} \quad \mathrm{S_c}$
- d. $\mathrm{S_c}^\lambda \rightarrow \mathrm{S_c}^\lambda\ \mathrm{S_c}$
- e. $\mathrm{S_c} \rightarrow \mathrm{C} \quad \mathrm{S}$
- f. $\mathrm{C} \rightarrow$ and
- g. $\mathrm{S} \rightarrow$ {black, green, red, white}

(10) L_2 = (black | green | red | white) (and (black | green | red | white))* = {black, . . . , black and white, . . . , black and white and red, . . . , black and white and red and green, . . . }

(11) $\Sigma_2{}^0$ = [$_\mathrm{S}$ (black | green | red | white)] | [$_{\mathrm{S_c}^2}$ [$_\mathrm{S}$ (black | green | red | white)] ([$_{\mathrm{S_c}}$ [$_\mathrm{C}$ and] [$_\mathrm{S}$ (black | green | red | white)]])$^+$] = {[$_\mathrm{S}$ black], . . . , [$_{\mathrm{S_c}^2}$ [$_\mathrm{S}$ black] [$_{\mathrm{S_c}}$ [$_\mathrm{C}$ and] [$_\mathrm{S}$ white]]], . . . , [$_{\mathrm{S_c}^2}$ [$_\mathrm{S}$ black] [$_{\mathrm{S_c}}$ [$_\mathrm{C}$ and] [$_\mathrm{S}$ white]] [$_{\mathrm{S_c}}$ [$_\mathrm{C}$ and] [$_\mathrm{S}$ red]]], . . . , [$_{\mathrm{S_c}^2}$ [$_\mathrm{S}$ black] [$_{\mathrm{S_c}}$ [$_\mathrm{C}$ and] [$_\mathrm{S}$ white]] [$_{\mathrm{S_c}}$ [$_\mathrm{C}$ and] [$_\mathrm{S}$ red]] [$_{\mathrm{S_c}}$ [$_\mathrm{C}$ and] [$_\mathrm{S}$ green]]]}

(12) black and white and red and green

(13) A CoE° = 1 structure that $J_2{}^0$ does not associate with the string 12:
[$_{\mathrm{S_c}^2}$ [$_{\mathrm{S_c}^2}$ [$_\mathrm{S}$ black] [$_{\mathrm{S_c}}$ [$_\mathrm{C}$ and] [$_\mathrm{S}$ white]]] [$_{\mathrm{S_c}}$ [$_\mathrm{C}$ and] [$_{\mathrm{S_c}^2}$ [$_\mathrm{S}$ red]] [$_{\mathrm{S_c}}$ [$_\mathrm{C}$ and] [$_\mathrm{S}$ green]]]]]

Given an NFC CFG J, one can construct an FT T that pairs every member of $L(J)$ directly with its structural descriptions in $\Sigma(J)$; thus T is directly strongly equivalent to J. The procedure Φ' for constructing T from J is an extension of Chomsky's procedure Φ.[7] The transitions of the FT $T_2{}^0$ obtained from $J_2{}^0$ by means of the procedure Φ' are listed in 14; the start symbols of $T_2{}^0$ are $[S]_1$ and $[S_c{}^2]_1$.[8]

(14) $\Pi(T_2{}^0)$

a.	$[\mathrm{S}]_1$	$\rightarrow$ black $[\mathrm{S}]_2$	[$_\mathrm{S}$ black
b.	$[\mathrm{S}]_1$	$\rightarrow$ green $[\mathrm{S}]_2$	[$_\mathrm{S}$ green
c.	$[\mathrm{S}]_1$	$\rightarrow$ red $[\mathrm{S}]_2$	[$_\mathrm{S}$ red
d.	$[\mathrm{S}]_1$	$\rightarrow$ white $[\mathrm{S}]_2$	[$_\mathrm{S}$ white

[7] The procedure Φ' incorporates a slightly modified version of Chomsky's procedure Ψ for constructing the weakly equivalent FA Q; see n. 8.

[8] The transitions in $\Pi(T_2)$ may be read as follows. From the state in the left side of the transition, go to the new state on the right side, reading the symbol that appears to the left of the new state on the input tape, and printing the string of symbols that appears to the right of the new state on the output tape. The repetition of the category S in the states in 14m–p is disallowed in Chomsky's original procedure Ψ, but is a legitimate extension, since there is no possibility of more than two occurrences of the same category in any state name. Alternatively, the second occurrence of S can be replaced by $S_c{}^\rho$ (or some other distinct symbol) in the right-hand state name in production 14q, and everywhere it occurs in state names in productions 14r–v.

e.	$[S]_2$	$\rightarrow$	ε	f	]
f.	$[S_c^2]_1$	$\rightarrow$	ε	$[S_c^2\ S]_1$	$[_{S_{c^2}}$
g.	$[S_c^2]_2$	$\rightarrow$	ε	f	]
h.	$[S_c^2\ S]_1$	$\rightarrow$	black	$[S_c^2\ S]_2$	$[_S$ black
i.	$[S_c^2\ S]_1$	$\rightarrow$	green	$[S_c^2\ S]_2$	$[_S$ green
j.	$[S_c^2\ S]_1$	$\rightarrow$	red	$[S_c^2\ S]_2$	$[_S$ red
k.	$[S_c^2\ S]_1$	$\rightarrow$	white	$[S_c^2\ S]_2$	$[_S$ white
l.	$[S_c^2\ S]_2$	$\rightarrow$	ε	$[S_c^2\ S\ S_c]_1$	]
m.	$[S_c^2\ S\ S_c]_1$	$\rightarrow$	ε	$[S_c^2\ S\ S_c\ C]_1$	$[_{S_c}$
n.	$[S_c^2\ S\ S_c]_2$	$\rightarrow$	ε	$[S_c^2\ S\ S_c]_1$	]
o.	$[S_c^2\ S\ S_c]_2$	$\rightarrow$	ε	$[S_c^2]_2$	]
p.	$[S_c^2\ S\ S_c\ C]_1$	$\rightarrow$	and	$[S_c^2\ S\ S_c\ C]_2$	$[_C$ and
q.	$[S_c^2\ S\ S_c\ C]_2$	$\rightarrow$	ε	$[S_c^2\ S\ S_c\ S]_1$	]
r.	$[S_c^2\ S\ S_c\ S]_1$	$\rightarrow$	black	$[S_c^2\ S\ S_c\ S]_2$	$[_S$ black
s.	$[S_c^2\ S\ S_c\ S]_1$	$\rightarrow$	green	$[S_c^2\ S\ S_c\ S]_2$	$[_S$ green
t.	$[S_c^2\ S\ S_c\ S]_1$	$\rightarrow$	red	$[S_c^2\ S\ S_c\ S]_2$	$[_S$ red
u.	$[S_c^2\ S\ S_c\ S]_1$	$\rightarrow$	white	$[S_c^2\ S\ S_c\ S]_2$	$[_S$ white
v.	$[S_c^2\ S\ S_c\ S]_2$	$\rightarrow$	ε	$[S_c^2\ S\ S_c]_2$	]

All that remains to establish strong equivalence between an NCE CNF CFG *G* and an FT *T* is to provide a procedure Ω for transforming any such *G* into a strongly equivalent NFC CFG *J*, as diagrammed in 15.

(15) $\mathrm{G} \xrightarrow{\Omega} \mathrm{J} \xrightarrow{\Phi'} \mathrm{T}$

Strong equivalence between an NCE CNF CFG *G* and an FT *T*; *J* is a strongly equivalent NFC CFG

4. Construction of an nfc cfg strongly equivalent to an nce cnf cfg. The procedure Ω can be thought of as recursively flattening RE and LE structures from top down, by making each embedded constituent a sister of the constituent that contains it, and leaving a 'trace'.[9] The resulting structures are like the ones proposed in Langendoen 1975 that result from the application of 'readjustment rules', but with the addition of traces. Langendoen 2003 shows how such structures can be obtained by a certain kind of internal merge (i.e. 'move') operation. My concern here, however, is to show how such structures are assigned by a variety of NFC CFG that may be called a normal-form linked coordinate (NFLC) CFG that allows for the appearance of 'vacuous movement' of embedded structures so as to render them coordinate. The NFC CFG *J* that Ω constructs from an NCE CNF *G* is strongly equivalent to *G* inasmuch as the structural descriptions that *J* assigns to a string *s* can be converted to those that *G* assigns to the corresponding string *s′* by 'reconstructing' the vacuously moved constituents into the positions of their traces.

Let *G* be an NCE CNF CFG. A strongly equivalent NFLC CFG *J* is obtained from *G* by first constructing an intermediate CFG *H* that eliminates RE in favor of RC (steps 1–5), and then *J* by eliminating LE in favor of LC (steps 6–10), as follows.

[9] The structural descriptions that *G* assigns can be uniquely recovered from those that *J* assigns by an inverse mapping Φ'^{-1} that 'reconstructs' each formerly embedded constituent into the position of its trace, relabels certain categories, and removes the outermost pair of brackets. As a result *J* is strongly equivalent to *G*. The fact that the procedure Φ'^{-1} cannot in general be carried out by an FT is immaterial if the linguistically 'real' grammar is *J*, rather than *G*; see §7.

STEP 1. If G is RE, set H identical to G; otherwise go to step 6.

STEP 2. Since G is RE, there is a category A in G such that $A \Rightarrow x\ A$ with respect to G. Consequently, either there is a category X such that the production in 16 is a member of $\Pi(G)$, or there are categories $X_1, \ldots, X_m, A_1, \ldots, A_m$ ($m \geq 1$) such that the productions in 17 are members of $\Pi(G)$.

(16) Possible member of $\Pi(G)$ if G is RE
$\mathrm{A} \rightarrow \mathrm{X\ A}$

(17) Alternate possible members of $\Pi(G)$ if G is RE
$\mathrm{A} \rightarrow \mathrm{X_1\ A_1}$
. . .
$\mathrm{A_m} \rightarrow \mathrm{A_m\ A}$

For each such category A, add the categories A^A, A^{A^A}, τ^A, and τ^{A^A} to the set of nonrecursive categories N_n of H. The category A^A is interpreted as 'A missing the A to its right', A^{A^A} as 'A missing the A to its right that's missing the A to its right', τ^A as the 'trace of A to its right' (i.e. the position in which the A to its right 'originated'), and τ^{A^A} as the 'trace of A missing the A to its right' (i.e. the position in which the A to its right that's missing an A originated). Add A^1 and A^ρ to the set of coordinate categories N_c of H. Also for each category A_i ($i \leq m$), if any, add the categories A_i^A and $A_i^{A^A}$ to N_n. Remove from $\Pi(H)$ the members of $\Pi(G)$ in 18 that reintroduce A into the derivation.

(18) Members of $\Pi(G)$ that are not members of $\Pi(H)$
$\mathrm{A} \rightarrow \mathrm{X\ \ A}$
$\mathrm{A_m} \rightarrow \mathrm{X_m\ A}$

STEP 3. Add to $\Pi(H)$ the productions in 19. Note that H is not an NFSC CFG since the productions 19a–d are not of the form specified in 8f. Instead they conform to the requirement that either $D_x = D$, or D_x is a 'slash' category derived from D in which an occurrence of D is 'missing'.[10]

(19) New members of $\Pi(H)$ for each RE category A in G
a. $\mathrm{A^1} \rightarrow \mathrm{A^A\ \ A}$
b. $\mathrm{A^1} \rightarrow \mathrm{A^{A^A}\ A^\rho}$
c. $\mathrm{A^\rho} \rightarrow \mathrm{A^A\ \ A}$
d. $\mathrm{A^\rho} \rightarrow \mathrm{A^{A^A}\ A^\rho}$
e. $\tau^{\mathrm{A}} \rightarrow \varepsilon$
f. $\tau^{\mathrm{A^A}} \rightarrow \varepsilon$

STEP 4. If the production in 16 is in $\Pi(G)$, then add the productions in 20 to $\Pi(H)$. But if the productions in 17 are in $\Pi(G)$, then add the productions in 21 to $\Pi(H)$.

(20) New members of $\Pi(H)$ if the production in 16 is in $\Pi(G)$
$\mathrm{A^A} \rightarrow \mathrm{X}\ \tau^{\mathrm{A}}$
$\mathrm{A^{A^A}} \rightarrow \mathrm{X}\ \tau^{\mathrm{A^A}}$

(21) New members of $\Pi(H)$ if the productions in 17 are in $\Pi(G)$
$\mathrm{A^A} \rightarrow \mathrm{X_1\ \ A_1{}^A}$
$\mathrm{A^{A^A}} \rightarrow \mathrm{X_1\ \ A_1{}^{A^A}}$
. . .
$\mathrm{A_m{}^A} \rightarrow \mathrm{X_m}\ \tau^{\mathrm{A}}$
$\mathrm{A_m{}^{A^A}} \rightarrow \mathrm{X_m}\ \tau^{\mathrm{A^A}}$

[10] The τ^A that occurs within A^A can be thought of as 'linked' to the A to its immediate right; similarly the τ^{A^A} that occurs within A^{A^A} can be thought of as linked to the A^A to its immediate right. Ultimately, the structures that J creates can be considered to be linked lists.

STEP 5. If A is a start symbol of H, then so is A^1. Otherwise, for every member of $\Pi(G)$ that introduces A on the right side, add the identical production to $\Pi(H)$, replacing A with A^1 as in 22.

(22) New member of $\Pi(H)$ for each production in $\Pi(G)$ of the form $D \rightarrow \varphi\, A\, \psi$
$\mathrm{D} \rightarrow \varphi\ \mathrm{A}^1\ \psi$

STEP 6. Set J identical to H.

STEP 7. If G is LE, then there is a category B in G such that $B \Rightarrow B\, y$ with respect to G. Consequently, either there is a category Y such that the production in 23 is a member of $\Pi(G)$, or there are categories $Y_1, \ldots, Y_n, B_1, \ldots, B_n$ ($n \geq 1$) such that the productions in 24 are members of $\Pi(G)$.

(23) Possible member of $\Pi(G)$ if G is LE
$\mathrm{B} \rightarrow \mathrm{B}\ \mathrm{Y}$

(24) Alternative possible members of $\Pi(G)$ if G is LE
$\mathrm{B} \rightarrow \mathrm{B}_1\ \mathrm{Y}_1$
. . .
$\mathrm{B}_n \rightarrow \mathrm{B}\ \ \mathrm{Y}_n$

For each such category B, add the categories ${}^{B}B$, ${}^{BB}B$, ${}^{B}\tau$, and ${}^{BB}\tau$ to the set of nonrecursive categories N_n of J, and B^2 and B^λ to the set of coordinate categories N_c of J. Also, for each category B_i ($i \leq n$), if any, add the new categories ${}^{B}B_i$ and ${}^{BB}B_i$. Remove from $\Pi(J)$ the members of $\Pi(G)$ in 25 that reintroduce B into the derivation.

(25) Members of $\Pi(G)$ that are not members of $\Pi(J)$
$\mathrm{B} \rightarrow \mathrm{B}\ \mathrm{Y}$
$\mathrm{B}_n \rightarrow \mathrm{B}\ \mathrm{Y}_n$

STEP 8. Add to $\Pi(J)$ the productions in 26.

(26) New members of $\Pi(J)$ for each LE category B in G
a. $\mathrm{B}^2 \rightarrow \mathrm{B}\ \ {}^{\mathrm{B}}\mathrm{B}$
b. $\mathrm{B}^2 \rightarrow \mathrm{B}^\lambda\ {}^{\mathrm{BB}}\mathrm{B}$
c. $\mathrm{B}^\lambda \rightarrow \mathrm{B}\ \ {}^{\mathrm{B}}\mathrm{B}$
d. $\mathrm{B}^\lambda \rightarrow \mathrm{B}^\lambda\ {}^{\mathrm{BB}}\mathrm{B}$
e. ${}^{\mathrm{B}}\tau \rightarrow \varepsilon$
f. ${}^{\mathrm{BB}}\tau \rightarrow \varepsilon$

STEP 9. If the production in 23 is in $\Pi(G)$, then add to $\Pi(J)$ the productions in 27. But if the productions in 24 are in $\Pi(G)$, then add to $\Pi(J)$ the productions in 28.

(27) New members of $\Pi(J)$ if the production in 23 is in $\Pi(G)$
${}^{\mathrm{B}}\mathrm{B} \rightarrow {}^{\mathrm{B}}\tau\ \ \mathrm{Y}$
${}^{\mathrm{BB}}\mathrm{B} \rightarrow {}^{\mathrm{BB}}\tau\ \mathrm{Y}$

(28) New members of $\Pi(J)$ if the productions in 24 are in $\Pi(G)$
${}^{\mathrm{B}}\mathrm{B} \rightarrow {}^{\mathrm{B}}\mathrm{B}_1\ \ \mathrm{Y}_1$
${}^{\mathrm{BB}}\mathrm{B} \rightarrow {}^{\mathrm{BB}}\mathrm{B}_1\ \mathrm{X}_1$
. . .
${}^{\mathrm{B}}\mathrm{B}_n \rightarrow {}^{\mathrm{B}}\tau\ \ \ \mathrm{X}_n$
${}^{\mathrm{BB}}\mathrm{B}_n \rightarrow {}^{\mathrm{BB}}\tau\ \ \mathrm{X}_n$

STEP 10. If B is a start symbol of J, then so is B^2. Otherwise for every member of $\Pi(H)$ that introduces B on the right side, add the identical production to $\Pi(J)$ replacing B with B^2 as in 29.

(29) New member of $\Pi(J)$ for each production in $\Pi(H)$ of the form $D \rightarrow \varphi\ B\ \psi$
$\mathrm{D} \rightarrow \varphi\ \mathrm{B}^2\ \psi$

This completes the statement of the algorithm Ω for constructing an NFLC CFG J that is strongly equivalent to an NCE CNF CFG G.

4.1. Construction of an RC NFLC CFG J_1 from an RE CNF CFG G_1. Given the RE CNF CFG G_1 in 2 that generates the regular language L_1 in 3, Ω constructs the RC NFLC CFG J_1 whose productions are listed in 30, along with the step numbers of Ω that were used to create them. Among the structures that J_1 assigns to the expressions it generates are those in 31, where τ^S represents $[_{\tau^S}\ \varepsilon]$ and τ^{S^S} represents $[_{\tau^{S^S}}\ \varepsilon]$. These are equivalent to the structures that G_1 assigns to the expressions it generates, inasmuch as the information contained in the embedding structures assigned by G_1 is preserved by the antecedent-trace relations in the coordinate structures assigned by J_1.

(30)

	$\Pi(J_1)$				STEP
a.	S	$\rightarrow$	A	B	1
b.	S^1	$\rightarrow$	S^{S}	S	3, 5
c.	S^1	$\rightarrow$	$\mathrm{S}^{\mathrm{S}^{\mathrm{S}}}$	S^{ρ}	3, 5
d.	S^{ρ}	$\rightarrow$	S^{S}	S	3
e.	S^{ρ}	$\rightarrow$	$\mathrm{S}^{\mathrm{S}^{\mathrm{S}}}$	S^{ρ}	3
f.	S^{S}	$\rightarrow$	A	B^{S}	4
g.	$\mathrm{S}^{\mathrm{S}^{\mathrm{S}}}$	$\rightarrow$	A	$\mathrm{B}^{\mathrm{S}^{\mathrm{S}}}$	4
h.	B^{S}	$\rightarrow$	C	τ^{S}	4
i.	$\mathrm{B}^{\mathrm{S}^{\mathrm{S}}}$	$\rightarrow$	C	$\tau^{\mathrm{S}^{\mathrm{S}}}$	4
j.	A	$\rightarrow$	they		1
k.	B	$\rightarrow$	fled		1
l.	C	$\rightarrow$	said		1
m.	τ^{S}	$\rightarrow$	ε		3
n.	$\tau^{\mathrm{S}^{\mathrm{S}}}$	$\rightarrow$	ε		3

(31) Some structures of sentences generated by J_1

a. $[_{\mathrm{S}}\ [_{\mathrm{A}}\ \text{they}]\ [_{\mathrm{B}}\ \text{fled}]]$

b. $[_{\mathrm{S}^1}\ [_{\mathrm{S}^{\mathrm{S}}}\ [_{\mathrm{A}}\ \text{they}]\ [_{\mathrm{B}^{\mathrm{S}}}\ [_{\mathrm{C}}\ \text{said}]\ \tau^{\mathrm{S}}]]\ [_{\mathrm{S}}\ [_{\mathrm{A}}\ \text{they}]\ [_{\mathrm{B}}\ \text{fled}]]]$

c. $[_{\mathrm{S}^1}\ [_{\mathrm{S}^{\mathrm{S}^{\mathrm{S}}}}\ [_{\mathrm{A}}\ \text{they}]\ [_{\mathrm{B}^{\mathrm{S}^{\mathrm{S}}}}\ [_{\mathrm{C}}\ \text{said}]\ \tau^{\mathrm{S}^{\mathrm{S}}}]]\ [_{\mathrm{S}^{\mathrm{S}}}\ [_{\mathrm{A}}\ \text{they}]\ [_{\mathrm{B}^{\mathrm{S}}}\ [_{\mathrm{C}}\ \text{said}]\ \tau^{\mathrm{S}}]]\ [_{\mathrm{S}}\ [_{\mathrm{A}}\ \text{they}]\ [_{\mathrm{B}}\ \text{fled}]]]$

d. $[_{\mathrm{S}^1}\ [_{\mathrm{S}^{\mathrm{S}^{\mathrm{S}}}}\ [_{\mathrm{A}}\ \text{they}]\ [_{\mathrm{B}^{\mathrm{S}^{\mathrm{S}}}}\ [_{\mathrm{C}}\ \text{said}]\ \tau^{\mathrm{S}^{\mathrm{S}}}]]\ [_{\mathrm{S}^{\mathrm{S}^{\mathrm{S}}}}\ [_{\mathrm{A}}\ \text{they}]\ [_{\mathrm{B}^{\mathrm{S}^{\mathrm{S}}}}\ [_{\mathrm{C}}\ \text{said}]\ \tau^{\mathrm{S}^{\mathrm{S}}}]]\ [_{\mathrm{S}^{\mathrm{S}}}\ [_{\mathrm{A}}\ \text{they}]\ [_{\mathrm{B}^{\mathrm{S}}}\ [_{\mathrm{C}}\ \text{said}]\ \tau^{\mathrm{S}}]]\ [_{\mathrm{S}}\ [_{\mathrm{A}}\ \text{they}]\ [_{\mathrm{B}}\ \text{fled}]]]$

4.2. Construction of an LC NFLC CFG J_3 from the LE CNF CFG G_3. In 32, I list the productions of an LE CNF CFG G_3 that generates the regular language L_3 in 33.[11] Since G_3 is not RE, steps 1–5 of Ω are skipped. Step 6 initially sets the productions of J_3 to those in 32. In 33, I list the final set of its productions, along with the steps used to create them. Among the structures that J_3 assigns to the expressions it generates are those in 35. These are equivalent to the structures that G_3 assigns to the expressions it generates, inasmuch as the information contained in the embedding structures assigned by G_3 is preserved by the antecedent-trace relations in the coordinate structures assigned by J_3.

[11] Except for *they fled*, the members of L_3 are not grammatical in English but closely resemble grammatical English sentences. For example, the sentence *they fled amused them* corresponds to, and may be understood as, ‘it amused them that they fled’.

(32) $\Pi(G_3)$

a. S $\rightarrow$ A B
b. S $\rightarrow$ S D
c. D $\rightarrow$ E F
d. A $\rightarrow$ they
e. B $\rightarrow$ fled
f. E $\rightarrow$ amused
g. F $\rightarrow$ them

(33) they fled (amused them)*

(34) $\Pi(J_3)$

					STEP
a.	S	$\rightarrow$	A	B	6
b.	S^2	$\rightarrow$	S	${}^{S}S$	8, 10
c.	S^2	$\rightarrow$	S^{λ}	${}^{SS}S$	8, 10
d.	S^{λ}	$\rightarrow$	S	${}^{S}S$	8
e.	S^{λ}	$\rightarrow$	S^{λ}	${}^{SS}S$	8
f.	${}^{S}S$	$\rightarrow$	${}^{S}\tau$	D	9
g.	${}^{SS}S$	$\rightarrow$	${}^{SS}\tau$	D	9
h.	D	$\rightarrow$	E	F	9
i.	A	$\rightarrow$	they		6
j.	B	$\rightarrow$	fled		6
k.	E	$\rightarrow$	amused		6
l.	F	$\rightarrow$	them		6
m.	${}^{S}\tau$	$\rightarrow$	ε		8
n.	${}^{SS}\tau$	$\rightarrow$	ε		8

(35) Some structures of sentences generated by J_3

a. $[_{S}\ [_{A}$ they$]\ [_{B}$ fled$]]$
b. $[_{S^2}\ [_{S}\ [_{A}$ they$]\ [_{B}$ fled$]]\ [_{{}^{S}S}\ {}^{S}\tau\ [_{D}\ [_{E}$ amused$]\ [_{F}$ them$]]]]$
c. $[_{S^2}\ [_{S}\ [_{A}$ they$]\ [_{B}$ fled$]]\ [_{{}^{S}S}\ {}^{S}\tau\ [_{D}\ [_{E}$ amused$]\ [_{F}$ them$]]]\ [_{{}^{SS}S}\ {}^{SS}\tau\ [_{D}\ [_{E}$ amused$]\ [_{F}$ them$]]]]$
d. $[_{S^2}\ [_{S}\ [_{A}$ they$]\ [_{B}$ fled$]]\ [_{{}^{S}S}\ {}^{S}\tau\ [_{D}\ [_{E}$ amused$]\ [_{F}$ them$]]]\ [_{{}^{SS}S}\ {}^{SS}\tau\ [_{D}\ [_{E}$ amused$]\ [_{F}$ them$]]]\ [_{{}^{SS}S}\ {}^{SS}\tau\ [_{D}\ [_{E}$ amused$]\ [_{F}$ them$]]]]$

4.3. Construction of a strongly equivalent ft from an nflc cfg. In 36 I list the transitions of T_1 that the procedure Φ' constructs from J_1 in 30; both J_1 and T_1 are strongly equivalent to G_1 in 2. The construction of J_3 from T_3 is similar; both J_3 and T_3 are strongly equivalent to G_3 in 32.

(36) $\Pi(T_1)$

a.	$[S]_1$	$\rightarrow$	ε	$[S\ A]_1$	$[_{S}$
b.	$[S]_2$	$\rightarrow$	ε	f	$]$
c.	$[S^1]_1$	$\rightarrow$	ε	$[S^1\ S^S]_1$	$[_{S^1}$
d.	$[S^1]_1$	$\rightarrow$	ε	$[S^1\ S^{S^S}]_1$	$[_{S^1}$
e.	$[S^1]_2$	$\rightarrow$	ε	f	$]$
f.	$[S\ A]_1$	$\rightarrow$	they	$[S\ A]_2$	$[_{A}$ they
g.	$[S\ A]_2$	$\rightarrow$	ε	$[S\ B]_1$	$]$
h.	$[S\ B]_1$	$\rightarrow$	fled	$[S\ B]_2$	$[_{B}$ fled
i.	$[S\ B]_2$	$\rightarrow$	ε	$[S]_2$	$]$
j.	$[S^1\ S]_1$	$\rightarrow$	ε	$[S^1\ S\ A]_1$	$[_{S}$
k.	$[S^1\ S]_2$	$\rightarrow$	ε	$[S^1]_2$	$]$
l.	$[S^1\ S^S]_1$	$\rightarrow$	ε	$[S^1\ S^S\ A]_1$	$[_{S^S}$
m.	$[S^1\ S^S]_2$	$\rightarrow$	ε	$[S^1\ S]_1$	$]$

n.	$[S^1\ S^{S^S}]_1$	$\rightarrow$	ε	$[S^1\ S^{S^S}\ A]_1$	$[_{S^{S^S}}$
o.	$[S^1\ S^{S^S}]_2$	$\rightarrow$	ε	$[S^1\ S^S]_1$	$]$
p.	$[S^1\ S^{S^S}]_2$	$\rightarrow$	ε	$[S^1\ S^{S^S}]_1$	$]$
q.	$[S^1\ S\ A]_1$	$\rightarrow$	they	$[S^1\ S\ A]_2$	$[_A$ they
r.	$[S^1\ S\ A]_2$	$\rightarrow$	ε	$[S^1\ S\ B]_1$	$]$
s.	$[S^1\ S\ B]_1$	$\rightarrow$	fled	$[S^1\ S\ B]_2$	$[_B$ fled
t.	$[S^1\ S\ B]_2$	$\rightarrow$	ε	$[S^1\ S]_2$	$]$
u.	$[S^1\ S^S\ A]_1$	$\rightarrow$	they	$[S^1\ S^S\ A]_2$	$[_A$ they
v.	$[S^1\ S^S\ A]_2$	$\rightarrow$	ε	$[S^1\ S^S\ B^S]_1$	$]$
w.	$[S^1\ S^S\ B^S]_1$	$\rightarrow$	ε	$[S^1\ S^S\ B^S\ C]_1$	$[_{B^S}$
x.	$[S^1\ S^S\ B^S]_2$	$\rightarrow$	ε	$[S^1\ S^S]_2$	$]$
y.	$[S^1\ S^{S^S}\ A]_1$	$\rightarrow$	they	$[S^1\ S^{S^S}\ A]_2$	$[_A$ they
z.	$[S^1\ S^{S^S}\ A]_2$	$\rightarrow$	ε	$[S^1\ S^{S^S}\ B^{S^S}]_1$	$]$
aa.	$[S^1\ S^{S^S}\ B^{S^S}]_1$	$\rightarrow$	ε	$[S^1\ S^{S^S}\ B^{S^S}\ C]_1$	$[_{B^{S^S}}$
bb.	$[S^1\ S^{S^S}\ B^{S^S}]_2$	$\rightarrow$	ε	$[S^1\ S^{S^S}]_2$	$]$
cc.	$[S^1\ S^S\ B^S\ C]_1$	$\rightarrow$	said	$[S^1\ S^S\ B^S\ C]_2$	$[_C$ said
dd.	$[S^1\ S^S\ B^S\ C]_2$	$\rightarrow$	ε	$[S^1\ S^S\ B^S\ \tau^S]_1$	$]$
ee.	$[S^1\ S^S\ B^S\ \tau^S]_1$	$\rightarrow$	ε	$[S^1\ S^S\ B^S\ \tau^S]_2$	$[_{\tau^S}$
ff.	$[S^1\ S^S\ B^S\ \tau^S]_2$	$\rightarrow$	ε	$[S^1\ S^S\ B^S]_2$	$]$
gg.	$[S^1\ S^{S^S}\ B^{S^S}\ C]_1$	$\rightarrow$	said	$[S^1\ S^{S^S}\ B^{S^S}\ C]_2$	$[_C$ said
hh.	$[S^1\ S^{S^S}\ B^{S^S}\ C]_2$	$\rightarrow$	ε	$[S^1\ S^{S^S}\ B^{S^S}\ \tau^{S^S}]_1$	$]$
ii.	$[S^1\ S^{S^S}\ B^{S^S}\ \tau^{S^S}]_1$	$\rightarrow$	ε	$[S^1\ S^{S^S}\ B^{S^S}\ \tau^{S^S}]_2$	$[_{\tau^{S^S}}$
jj.	$[S^1\ S^{S^S}\ B^{S^S}\ \tau^{S^S}]_2$	$\rightarrow$	ε	$[S^1\ S^{S^S}\ B^{S^S}]_2$	$]$

5. Construction of a strongly equivalent 'extended' NFLC for a CE CNF CFG up to a finitely bounded degree of 'zigzag embedding'. CE arises in a CNF CFG G as a result of the combination of LE and RE of the same category. For example, if the productions of the RE CNF CFG G_1 in 2 are merged with those of the LE CNF CFG G_3 in 32, the result is the CE CNF CFG G_4 whose productions are in 37 and in which the category S is both LE and RE. G_4 associates the structural descriptions in 38 (with all but the S-brackets omitted, for clarity) to the string 38, each with degree of CE (CE°) = 1.

(37) $\Pi(G_4) = \Pi(G_1) \cup \Pi(G_3)$
  a. S $\rightarrow$ A B
  b. S $\rightarrow$ S D
  c. B $\rightarrow$ C S
  d. D $\rightarrow$ E G
  e. A $\rightarrow$ they
  f. B $\rightarrow$ fled
  g. C $\rightarrow$ said
  h. E $\rightarrow$ amused
  i. G $\rightarrow$ them

(38) they said they fled amused them

(39) Structures with CE° = 1 that G_4 associates with the string 38
  a. $[_S\ [_S$ they said $[_S$ they fled]] amused them]
    'it amused them that they said that they fled'
  b. $[_S$ they said $[_S\ [_S$ they fled] amused them]]
    'they said that it amused them that they fled'

Applying the procedure Ω to G_4, the RC and LC NFLC CFG $J_4{}^0$ is obtained, with the start symbols S, S^1, and S^2, and the productions in 40.

(40) $\Pi(J_4{}^0)$

a. $S \rightarrow A \quad B$
b. $S^1 \rightarrow S^S \quad S$
c. $S^1 \rightarrow S^{SS} \quad S^\rho$
d. $S^2 \rightarrow S \quad {}^{S}S$
e. $S^2 \rightarrow S^\lambda \quad {}^{SS}S$
f. $S^\rho \rightarrow S^S \quad S$
g. $S^\rho \rightarrow S^{SS} \quad S^\rho$
h. $S^\lambda \rightarrow S \quad {}^{S}S$
i. $S^\lambda \rightarrow S^\lambda \quad {}^{SS}S$
j. $S^S \rightarrow A \quad B^S$
k. $S^{SS} \rightarrow A \quad B^{SS}$
l. $B^S \rightarrow C \quad \tau^S$
m. $B^{SS} \rightarrow C \quad \tau^{SS}$
n. ${}^{S}S \rightarrow {}^{S}\tau \quad D$
o. ${}^{SS}S \rightarrow {}^{SS}\tau \quad D$
p. $D \rightarrow E \quad F$
q. $A \rightarrow$ they
r. $B \rightarrow$ fled
s. $C \rightarrow$ said
t. $E \rightarrow$ amused
u. $F \rightarrow$ them
v. $\tau^S \rightarrow \varepsilon$
w. $\tau^{SS} \rightarrow \varepsilon$
x. ${}^{S}\tau \rightarrow \varepsilon$
y. ${}^{SS}\tau \rightarrow \varepsilon$

However, $J_4{}^0$ is not even weakly, much less strongly, equivalent to G_4, since it fails to generate 38, or any other sentence generated by G_4 with CE° > 0; consequently, the FT $T_4{}^0$ that the procedure Φ' constructs from it is also not strongly equivalent to G_4. Indeed, NFLC CFGs as defined so far can only generate strings with structures that correspond to those with CE° $= 0$ when generated by a CE CNF CFG. In the next section, I introduce the notion of zigzag embedding (ZE), which provides a more refined measure of structural complexity than does CE, and show how the procedure Ω can be extended to Ω^+ to construct NFLC CFGs that generate strings with up to any desired degree k of ZE (i.e. with ZE° $\leq k$).

5.1. Enabling nflc cfgs to handle bounded degrees of ze. In 38a, CE results from a combination of left-then-right embedding (LRE), and in 38b it does so from a combination of right-then-left embedding (RLE). In both cases, the embedding manifests a change in direction of embedding or zigzag embedding (ZE). Clearly, for any CNF CFG, a structure manifests CE if and only if it manifests ZE, and in both cases in 38, ZE° $=$ CE° $= 1$.[12] However, many structures associated with strings generated by G_4 manifest a ZE° greater than their CE°; an example in which ZE° $= 2$ but CE° $= 1$ is given in 41.

[12] ZE° is defined inductively as follows. Let C be an RE and LE category in a CE CNF CFG G, and let C_0 be the first occurrence of C in a derivation with respect to G, that is, $S \Rightarrow \chi\ C_0\ \omega$, where S is a start symbol of G, and for no substring ξ of $\chi\ C_0\ \omega$ properly including C_0 does $C \Rightarrow \xi$. Then C has LRE° (left-then-right-embedding degree) $= 1$ if $C_0 \Rightarrow C_i\ \psi$ and $C_i \Rightarrow \varphi\ C$. Similarly, C has RLE° (right-then-left-embedding degree) $= 1$ if $C_0 \Rightarrow \varphi\ C_i$ (where C_i, like C_0, is an occurrence of C) and $C_i \Rightarrow C\ \psi$. Next, if C_j (where C_j is also an occurrence of C) has RLE° $= n$, then C has LRE° $= n+1$ if $C_j \Rightarrow \varphi\ C$. Similarly, if C_j has LRE° $= n$, then C has RLE° $= n+1$ if $C_j \Rightarrow C\ \psi$. Finally, if C has LRE° $= k$ or RLE° $= k$, then C has ZE° (zigzag-embedding degree) $= k$.

(41) String generated by G_4 and corresponding structure in which ZE° = 2 and CE° = 1

a. they said they said they fled amused them

b. [$_S$ they said [$_S$ [$_S$ they said [$_S$ they fled]] amused them]]

'they said that it amused them that they said that they fled'

The CE° of an RE and LE category C in a derivation with respect to a CE CNF CFG G is related to its ZE° by the inequalities in 42. Accordingly, any bound on ZE° is also a bound on CE°, but a maximum bound n ($n > 1$) on ZE° results in excluding not only all structures with CE° $> n$, but also some structures with CE° $\leq n$. I discuss the significance of this difference between ZE° and CE° bounding below in §6.

(42) Relation of ZE° to CE°

a. $\frac{1}{2}(\text{ZE}° + 1) \leq \text{CE}° \leq \text{ZE}°$, if ZE° is odd

b. $\frac{1}{2}(\text{ZE}°) \leq \text{CE}° \leq \text{ZE}°$, if ZE° is even

The procedure Ω^+ proposed in §5.2 below proceeds by ZE°. Starting with the NFLC CFG J^0 that Ω constructs from a CE CNF CFG G that is strongly equivalent to G for all sentences generated by G with ZE° < 1, Ω^+ constructs an EXTENDED NFLC CFG (ENFLC CFG) J^1 that assigns structural descriptions for all sentences equivalent to those assigned by G with ZE° < 2, and so on. Moreover, for each J^i that Ω^+ constructs, the procedure Φ' constructs a strongly equivalent FT T^i.

5.2. FROM ZE° = 0 TO ZE° ≤ 1. Let G be a CE CNF CFG, and J^0 the NFLC CFG constructed from G using the procedure Ω. The procedure Ω^+ constructs the ENFLC CFG J^1 that generates all sentences of $L(G)$ with ZE° ≤ 1 as follows.

STEP 1. If C^1 and C^2 are categories of J^0, add the categories C^{12}, C^{21}, $C^{1\lambda}$, $C^{1\lambda'}$, $C^{2\rho}$, and $C^{2\rho'}$ to the 'extended' vocabulary N_e of J^1; C^{12} and C^{21} are additional start symbols if C^1 and C^2 are. If they are not, add the necessary productions to introduce them into derivations in the manner of 22 and 29.

STEP 2. Add the categories ${}^{C^1}C$, C^{C^2}, ${}^{C^1C}C$, $C^{CC^2_2}$, ${}^{C^1}\tau$, τ^{C^2}, ${}^{C^1C}\tau$, and τ^{CC^2} to the vocabulary N_n, and for each $C_i{}^C$ in J^0 add $C_i{}^{C^2}$ and $C_i{}^{CC^2_2}$, and for each ${}^CC_j$ in J^0 add ${}^{C^1}C_j$ and ${}^{C^1C}C_j$ ($1 \leq i \leq m$, $1 \leq j \leq n$). Add the productions in 43.

(43) New members of $\Pi(J^1)$ for each RE and LE category C in G

a. $C^{12} \rightarrow C^1 \quad {}^{C^1}C$

b. $C^{12} \rightarrow C^{1\lambda'} \quad {}^{C^1C}C$

c. $C^{12} \rightarrow C^{1\lambda} \quad {}^{CC}C$

d. $C^{21} \rightarrow C^{C^2} \quad C^2$

e. $C^{21} \rightarrow C^{CC^2} \quad C^{2\rho'}$

f. $C^{21} \rightarrow C^{CC} \quad C^{2\rho}$

g. $C^{1\lambda'} \rightarrow C^1 \quad {}^{C^1}C$

h. $C^{1\lambda} \rightarrow C^{1\lambda'} \quad {}^{C^1C}C$

i. $C^{1\lambda} \rightarrow C^{1\lambda} \quad {}^{CC}C$

j. $C^{2\rho'} \rightarrow C^{C^2} \quad C^2$

k. $C^{2\rho} \rightarrow C^{CC^2} \quad C^{2\rho'}$

l. $C^{2\rho} \rightarrow C^{CC} \quad C^{2\rho}$

m. ${}^{C^1}\tau \rightarrow \varepsilon$

n. $\tau^{C^2} \rightarrow \varepsilon$

o. ${}^{C^1C}\tau \rightarrow \varepsilon$

p. $\tau^{CC^2} \rightarrow \varepsilon$

STEP 3. If J^0 has the productions in 20 and 27 (with C replacing A and B), add the productions in 44; but if J^0 has the productions in 21 and 28 (with C replacing A and B), add the productions in 45.

(44) New members of $\Pi(J^1)$ if 20 and 27 (with C replacing A and B) are in $\Pi(J^0)$

a. ${}^{C^1}C \rightarrow {}^{C^1}\tau\ \ X$

b. $C^{C^2} \rightarrow X\ \ \tau^{C^2}$

c. ${}^{{}^{C^1}C}C \rightarrow {}^{{}^{C^1}C}\tau\ X$

d. $C^{C^{C^2}} \rightarrow X\ \ \tau^{C^{C^2}}$

(45) New members of $\Pi(J^1)$ if 21 and 28 (with C replacing A and B) are in $\Pi(J^0)$

${}^{C^1}C \rightarrow {}^{C^1}C_1\ \ X_1$

$C^{C^2} \rightarrow X_1\ \ C_1{}^{C^2}$

${}^{{}^{C^1}C}C \rightarrow {}^{{}^{C^1}C}C_1\ X_1$

$C^{C^{C^2}} \rightarrow X_1\ \ C_1{}^{C^{C^2}}$

. . .

${}^{C^1}C_n \rightarrow {}^{C^1}\tau\ \ X_n$

$C_m{}^{C^2} \rightarrow X_m\ \ \tau^{C^2}$

${}^{{}^{C^1}C}C_n \rightarrow {}^{{}^{C^1}C}\tau\ \ X_n$

$C_m{}^{C^{C^2}} \rightarrow X_m\ \ \tau^{C^{C^2}}$

This completes the construction of the ENFLC CFG J^1. The strongly equivalent FT T^1 can be constructed from J^1 by the procedure Φ'.

According to Ω^+, the productions of $J_4{}^1$ are those in 40 together with those in 43, with S replacing C. In particular, $J_4{}^1$ assigns the structures in 46 (with all but the S-type brackets omitted), which are equivalent to their ZE° = CE° = 1 counterparts in 39, repeated here for convenience as 47. Strictly speaking, the structures in 46 are nonembedding because the categories corresponding to the embedding categories in 39/47 have been relabeled. However, their bracketing is isomorphic to that in 39/47, so their pattern of embedding may be called PSEUDO-ZE (ψZE) and PSEUDO-CE (ψCE); moreover, their ψZE° = ψCE° = 1.

(46) ψZE° = ψCE° = 1 structures that $J_4{}^1$ associates with the string 38 that are equivalent to those in 39/47

a. $[_{S^{12}}$ $[_{S^1}$ $[_{S^S}$ they said $\tau^S]$ $[_S$ they fled]] $[_{{}^{S^1}S}$ ${}^{S^1}\tau$ amused them]]

b. $[_{S^{21}}$ $[_{S^{S^2}}$ they said $\tau^{S^2}]$ $[_{S^2}$ $[_S$ they fled] $[_{{}^SS}$ ${}^S\tau$ amused them]]]

(47) ZE° = CE° = 1 structures that G_4 associates with the string 38

a. $[_S$ $[_S$ they said $[_S$ they fled]] amused them]

'it amused them that they said that they fled'

b. $[_S$ they said $[_S$ $[_S$ they fled] amused them]]

'they said that it amused them that they fled'

In addition, G_4 generates the string 41a with the two ZE° = CE° = 1 structures in 48. $J_4{}^1$ assigns the equivalent ψZE° = ψCE° = 1 structures in 49. However, while G_4 also generates that string with the CE° = 1 structure in 41b, repeated here as 50, $J_4{}^1$ is unable to assign its structural counterpart, because its ZE° = 2.

(48) ZE° = CE° = 1 structures that G_4 associates with the string 41a

a. $[_S$ $[_S$ they said $[_S$ they said $[_S$ they fled]]] amused them]

'it amused them that they said that they said that they fled'

b. $[_S$ they said $[_S$ they said $[_S$ $[_S$ they fled] amused them]]]

'they said that it amused them that they said that they fled'

(49) ψZE° = ψCE° = 1 structures that J_4^1 associates with the string 41a that are equivalent to those in 48

a. $[_{S^{12}}\ [_{S^1}\ [_{S^{S^S}}$ they said $\tau^{S^S}]\ [_{S^S}$ they said $\tau^S]\ [_S$ they fled$]]\ [_{{}^{S^1}S}\ {}^{S^1}\tau$ amused them]]

b. $[_{S^{21}}\ [_{S^{S^{S^2}}}$ they said $\tau^{S^{S^2}}]\ [_{S^{S^2}}$ they said $\tau^{S^2}]\ [_{S^2}\ [_S$ they fled$]\ [_{{}^{S}S}\ {}^{S}\tau$ amused them]]]

(50) ZE° = 2, CE° = 1 structure that G_4 associates with the string 41a that J_4^1 cannot handle

$[_S$ they said $[_S\ [_S$ they said $[_S$ they fled]] amused them]]

'they said that it amused them that they said that they fled'

G_4 also generates the string 51 with the two ZE° = CE° = 1 structures in 51, and J_4^1 assigns the equivalent ψZE° = ψCE° = 1 structures in 53. However, G_4 also associates with that string the two ZE° = CE° = 2 structures in 54, and the two ZE° = 3, CE° = 2 structures in 55, which J_4^1 is unable to handle.

(51) they said they said they fled amused them amused them

(52) ZE° = CE° = 1 structures that G_4 associates with the string 51

a. $[_S\ [_S\ [_S$ they said $[_S$ they said $[_S$ they fled]]] amused them] amused them]

'it amused them that it amused them that they said that they said that they fled'

b. $[_S$ they said $[_S$ they said $[_S\ [_S\ [_S$ they fled] amused them] amused them]]]

'they said that they said that it amused them that it amused them that they fled'

(53) ψZE° = ψCE° = 1 structures that J_4^1 associates with the string 51 that are equivalent to those in 51

a. $[_{S^{12}}\ [_{S^1}\ [_{S^{S^S}}$ they said $\tau^{S^S}]\ [_{S^S}$ they said $\tau^S]\ [_S$ they fled$]]\ [_{{}^{S^1}S}\ {}^{S^1}\tau$ amused them$]\ [_{{}^{S^1S}S}\ {}^{S^1S}\tau$ amused them]]

b. $[_{S^{21}}\ [_{S^{S^{S^2}}}$ they said $\tau^{S^{S^2}}]\ [_{S^{S^2}}$ they said $\tau^{S^2}]\ [_{S^2}\ [_S$ they fled$]\ [_{{}^{S}S}\ {}^{S}\tau$ amused them$]\ [_{{}^{SS}S}\ {}^{SS}\tau$ amused them]]]

(54) ZE° = CE° = 2 structures that G_4 associates with the string 51 that J_4^1 is unable to handle

a. $[_S$ they said $[_S\ [_S\ [_S$ they said $[_S$ they fled]] amused them] amused them]]

'they said that it amused them that it amused them that they said that they fled'

b. $[_S\ [_S$ they said $[_S$ they said $[_S\ [_S$ they fled] amused them]]] amused them]

'it amused them that they said that they said that it amused them that they fled'

(55) ZE° = 3, CE° = 2 structures that G_4 associates with the string 51 that J_4^1 is also unable to handle

a. $[_S$ they said $[_S\ [_S$ they said $[_S\ [_S$ they fled] amused them]] amused them]]

'they said that it amused them that they said that it amused them that they fled'

b. $[_S\ [_S$ they said $[_S\ [_S$ they said $[_S$ they fled]] amused them]] amused them]

'it amused them that they said that it amused them that they said that they fled'

5.3. From ze° ≤ n to ze° ≤ n + 1. The procedure Ω^+ constructs the ENFLC CFG J^{n+1} from the ENFLC CFG J^n by following the same steps as described in §5.2 for constructing the ENFLC CFG J^1 from the NFLC CFG J^0. If n is even, replace C^1 (in

J^0) by $C^{(21)^{\frac{1}{2}n}}$ (in J^n) and C^2 by $C^{(12)^{\frac{1}{2}n}}$; if n is odd, replace C^1 by $C^{1(21)^{\frac{1}{2}(n-1)}}$ and C^2 by $C^{2(12)^{\frac{1}{2}(n-1)}}$. This completes the description of the procedure Ω^+. Viewed as a method of approximating context-free grammars using finite-state transducers, this approach is like that proposed by Nederhof (2000), who first 'transforms' a CE CFG into a form in which an FT can be efficiently designed to recognize the expressions it generates with bounded CE°. Since he is concerned with weak, rather than strong, approximation, however, he does not deal with the problem presented by RE and LE.

6. Bounding ze° vs. ce°: consequences for comprehension and acquisition. It has long been known that different types of sentences with the same CE° present varying degrees of difficulty in comprehension (Bever 1970). One reason for this is that CE° is a crude measure of grammatical difficulty, which masks some differences in ZE°. In §5.1 above, I point out that for the strings generated by CNF CFGs, CE° is never larger than ZE°, and that in certain structures CE° is as small as ½(ZE°). If difficulty increases with ZE°, then if A and B are otherwise comparable structures such that ZE°(A) > ZE°(B) but CE°(A) = CE°(B), then all things being equal A should be harder to process than B. For example, the ZE° = 2, CE° = 1 structure in 50 appears to be more difficult to comprehend than the string-identical ZE° = CE° = 1 structures in 48, though judgment is obscured by the fact that the example is ungrammatical in English.

For a grammatical example, consider the phrase in 56, which is structurally ambiguous in English. Its most natural interpretations appear to be the ones based on the structures in 57a,b, for which ZE° = CE° = 1. Its least natural interpretation appears to be the one based on the structure in 57c, for which ZE° = 2 and CE° = 1.

(56) the tallest son of the mighty queen's oldest daughter's favorite child

(57) a. [$_N$ [$_N$ the tallest son] of [$_N$ [$_N$ [$_N$ the mighty queen's] [$_N$ oldest daughter's]] [$_N$ favorite child]]]
'the mighty queen's oldest daughter's favorite child's tallest son'

b. [$_N$ [$_N$ [$_N$ [$_N$ the tallest son] of [$_N$ the mighty queen's]] [$_N$ oldest daughter's]] [$_N$ favorite child]]
'the mighty queen's tallest son's oldest daughter's favorite child'

c. [$_N$ [$_N$ [$_N$ the tallest son] of [$_N$ [$_N$ the mighty queen's]] [$_N$ oldest daughter's]] [$_N$ favorite child]]
'the mighty queen's oldest daughter's tallest son's favorite child'

Another example is the phrase in 58, which is of a type that Langendoen, McDaniel, and Langsam (1989) asked subjects to draw pictures of. Almost everyone drew them based on ZE° = CE° = 1 (consistently low or high attachment) structures, as in 59a,b; very few did so based on ZE° = 2, CE° = 1 (alternate low and high attachment) structures, as in 59c.

(58) the star below the circle beside the diamond above the square

(59) a. [$_N$ [$_N$ the star] [$_N$ [$_N$ below the circle] [$_N$ [$_N$ beside the diamond] [$_N$ above the square]]]]
'the star [is] below the circle, the circle [is] beside the diamond, the diamond [is] above the square'

b. [$_N$ [$_N$ [$_N$ [$_N$ the star] [$_N$ below the circle]] [$_N$ beside the diamond]] [$_N$ above the square]]
'the star [is] below the circle, the star [is] beside the diamond, the star [is] above the square'

c. [$_N$ [$_N$ [$_N$ the star] [$_N$ [$_N$ below the circle] [$_N$ beside the diamond]]] [$_N$ above the square]]
 'the star [is] below the circle, the circle [is] beside the diamond, the star [is] above the square'

The idea that the grammatical complexity of recursive structures is better measured by ZE° than by CE° suggests that mastery of these structures is acquired by order of ZE°. Although this question has not yet been empirically investigated, a critical test may be made using adverbial adjunction to sentence-final clauses in English, as in 60, where high and low attachment give rise to structures with ZE° $=$ 1, as in 61a,b, whereas medial attachment results in structures with ZE° $=$ 2, as in 61c.

(60) they said they said they fled yesterday

(61) a. [$_S$ they said [$_S$ they said [$_S$ [$_S$ they fled] yesterday]]]
 'they said that they said that it was yesterday that they fled'

b. [$_S$ [$_S$ they said [$_S$ they said [$_S$ they fled]]] yesterday]
 'it was yesterday that they said that they said that they fled'

c. [$_S$ they said [$_S$ [$_S$ they said [$_S$ they fled]] yesterday]]
 'they said that it was yesterday that they said that they fled'

Pearlmutter and Gibson (2001) show that under certain conditions, medial attachment is preferred to high attachment for adult speakers in structurally ambiguous sentences like 60, and contend that a processing principle they call 'predicate proximity' interacts with other principles (ZE° is of course not one of them!) to give this result. Thus it would be of interest to study the development of children's understanding of such sentences, particularly to determine whether there is a stage in which the low- and high-attachment interpretations are available, but not the medial one.

7. FTs vs. cfgs as models of knowledge of language. As noted in §1, the research on finite-state (regular) approximation of context-free languages and grammars has largely assumed, following Miller and Chomsky (1963), that the theory of CFG provides the computationally weakest potentially adequate model of human knowledge of language, and that memory or other linguistically irrelevant performance limitations prevent people from assigning structural descriptions to strings that are well formed with respect to an internalized CFG with CE° $>$ n for some n. If that is so, then if people are given more 'space' to compute these structures, the bound on CE° should increase without any change in their internalized grammar. No one, however, has yet proposed a mechanism by which people use additional memory to compute structural descriptions for strings whose CE° (or more precisely ZE°) is beyond their normal capacity to comprehend. In fact, the bound on CE°/ZE° appears to increase very little, no matter how much space and time people are given to compute structures with CE°/ZE° beyond their normal capacity. This observation suggests that overcoming the bound on CE°/ZE° requires much more mental work than simply making use of additional computation space to relabel nodes in a structural description, as Chomsky (1963:400) has suggested.

Suppose instead we adopt a version of Krauwer and des Tombe's (1979) proposal that FTs of a certain form provide the weakest potentially adequate model of human knowledge of language, and that people are equipped to extend them systematically given the right conditions. More specifically, suppose that human linguistic development can be modeled initially by FTs as defined in §3 that assign recursive coordinate (paratactic) structures of the sort that NFSC CFGs assign. Then when confronted with evidence that structures equivalent to recursive NCE (i.e. RE or LE) subordinate (hypotactic) structures are required, a mechanism for constructing FTs as defined in §4 that assign structures of the sort that NFLC CFGs assign is available. Next, when confronted with evidence that structures equivalent to those with ZE° = 1 are required, a mechanism for constructing FTs as defined in §5 that assign structures of the sort that ENFLC CFGs assign is available. The theory of FTs equivalent to ENFLC CFGs does not fix a specific upper bound on ZE°, but instead defines an infinite series of classes of transducers, starting with those like T_4^0 that assign structures without bound on RC° and LC°, but limit ZE° to 0. The transduction counterpart to procedure Ω^+ is capable of successively constructing each T^{n+1} that assigns structures with ψZE° $\leq n + 1$ from T^n that assigns structures with ψZE° $\leq n$. The application of the counterpart to Ω^+ and the burden of using the resulting transducers is predicted to be increasingly difficult as N increases if for no other reason than the rapidly increasing size and complexity of those transducers, so that a natural limit is quickly reached, but no specific threshold beyond which ψZE° cannot pass is predicted.

8. LIMITATION ON COORDINATE EMBEDDING. As shown in §3 above, NFSC CFGs and their FT counterparts constructed by Φ' are designed to assign structures with CoE° = 0 only (i.e. 'flat structures'). However, I identified structures with CoE° as great as 2, such as that in 12 (with CoE° = 1), in natural languages (Langendoen 1998). I now contend that the structures that are found exhibit bounded 'pseudo-CoE°' (ψCoE°). In this section, I show how bounded ψCoE° can be accounted for by a procedure Ξ that is similar to Ω^+.

Rather than specifying Ξ in detail, I show how it applies to the NFSC CFG J_2^0 to create an EXTENDED NFSC CFG (ENFSC CFG) J_2^1 that assigns structures with ψCoE° ≤ 1 to the strings it generates. First, the categories $S_c{}^2{}_c{}^2$, $S_c{}^2{}_cU$, and $S_c{}^2{}_c{}^\lambda$ are added to the coordinate vocabulary V_c, with $S_c{}^2{}_c{}^2$ a start symbol. Second, the productions in 62 are added. The resulting ENFSC CFG J_2^1 is weakly equivalent to J_2^0 but also assigns ψCoE° = 1 structures, such as the one in 63 to the string 12, which is equivalent to the CoE° = 1 structure in 13, repeated here as 64, to all strings in L_2 with at least three coordinate members. It is a straightforward matter to apply the procedure Ξ to iteratively create grammars that assign structures with ψCoE° $\leq N$ for any N, and structures with bounded ψCoE° of the various types discussed in Langendoen 1998, for example, those with mixed coordinators.

(62) Members of $\Pi(J_2^1)$ that are not in $\Pi(J_2^0)$

a. $\mathrm{S_c{}^2{}_c{}^2 \rightarrow S_c{}^2 \quad S_c{}^2{}_c}$

b. $\mathrm{S_c{}^2{}_c{}^2 \rightarrow S_c{}^2{}_c{}^\lambda \; S_c{}^2{}_c}$

c. $\mathrm{S_c{}^2{}_c{}^\lambda \rightarrow S_c{}^2 \quad S_c{}^2{}_c}$

d. $\mathrm{S_c{}^2{}_c{}^\lambda \rightarrow S_c{}^2{}_c{}^\lambda \; S_c{}^2{}_c}$

e. $\mathrm{S_c{}^2{}_c \rightarrow C \quad S_c{}^2}$

(63) A ψCoE° = 1 structure that J_2^1 assigns to the string 12

$[_{\mathrm{S_c{}^2{}_c{}^2}}$ $[_{\mathrm{S_c{}^2}}$ $[_\mathrm{S}$ black] $[_{\mathrm{S_c}}$ $[_\mathrm{C}$ and] $[_\mathrm{S}$ white]]] $[_{\mathrm{S_c{}^2{}_c}}$ $[_\mathrm{C}$ and] $[_{\mathrm{S_c{}^2}}$ $[_\mathrm{S}$ red]] $[_{\mathrm{S_c}}$ $[_\mathrm{C}$ and] $[_\mathrm{S}$ green]]]]

(64) A CoE° = 1 structure for the string 12

$[_{\mathrm{S_C{}^2}}$ $[_{\mathrm{S_c{}^2}}$ $[_\mathrm{S}$ black] $[_{\mathrm{S_c}}$ $[_\mathrm{C}$ and] $[_\mathrm{S}$ white]]] $[_{\mathrm{S_c}}$ $[_\mathrm{C}$ and] $[_{\mathrm{S_c{}^2}}$ $[_\mathrm{S}$ red]] $[_{\mathrm{S_c}}$ $[_\mathrm{C}$ and] $[_\mathrm{S}$ green]]]]

In Langendoen 1998, I contended that the bound on CoE° in natural languages is grammatically determined, and proposed an optimality-theoretic account for it. But if coordinate structures in natural languages are generated by the productions of ENFSC

CFGs or their FT equivalents, there is no actual CoE in natural languages, only ΨCoE, and the existence of a finite bound n on ψCoE° follows from whatever prevents people from constructing more complex FTs than those that are capable of recognizing structures with ψCoE° $\leq N$, just as I have proposed for ψZE° in §7 above.

REFERENCES

Bever, Thomas G. 1970. The cognitive basis for linguistic structures. *Cognition and the development of language*, ed. by John R. Hayes, 279–352. New York: John Wiley & Sons.

Chomsky, Noam. 1959a. On certain formal properties of grammars. *Information and Control* 2.137–67. [Reprinted in *Readings in mathematical psychology*, vol. 2, ed. by R. Duncan Luce, Robert R. Bush, and Eugene Galanter, 125–55. New York: John Wiley & Sons, 1965.]

Chomsky, Noam. 1959b. A note on phrase-structure grammars. *Information and Control* 2.393–95.

Chomsky, Noam. 1963. Formal properties of grammars. In Luce et al., 323–418.

Chomsky, Noam, and George A. Miller. 1963. Introduction to the formal analysis of natural languages. In Luce et al., 269–321.

Gazdar, Gerald; Ewan Klein; Geoffrey K. Pullum; and Ivan A. Sag. 1985. *Generalized phrase structure grammar*. Oxford: Blackwell.

Krauwer, Steven, and Louis des Tombe. 1979. Transducers and grammars as theories of language. *Theoretical Linguistics* 8.173–201.

Langendoen, D. Terence. 1961. Structural descriptions for sentences generated by non-self-embedding constituent-structure grammars. Cambridge, MA: MIT SB thesis.

Langendoen, D. Terence. 1975. Finite-state parsing of phrase-structure languages and the status of readjustment rules in grammar. *Linguistic Inquiry* 6.533–54.

Langendoen, D. Terence. 1976. On the weak generative capacity of infinite grammars. *CUNYForum* 1.13–24.

Langendoen, D. Terence. 1979. On the assignment of constituent structures to the sentences generated by a transformational grammar. *CUNYForum* 8.1–32.

Langendoen, D. Terence. 1998. Limitations on embedding in coordinate structures. *Journal of Psycholinguistic Research* 27.235–59.

Langendoen, D. Terence. 2003. Merge. *Formal approaches to functional phenomena: In honor of Eloise Jelinek*, ed. by Andrew Carnie, Heidi Harley, and Mary Willie, 307–18. Amsterdam: John Benjamins.

Langendoen, D. Terence; Dana McDaniel; and Yedidyah Langsam. 1989. Preposition-phrase attachment in noun phrases. *Journal of Psycholinguistic Research* 18.533–48.

Luce, R. Duncan; Robert R. Bush; and Eugene Galanter (eds.) 1963. *Handbook of mathematical psychology*, vol. 2. New York: John Wiley & Sons.

Miller, George A., and Noam Chomsky. 1963. Finitary models of language users. In Luce et al., 419–91.

Nederhof, Mark-Jan. 2000. Practical experiments with regular approximation of context-free languages. *Computational Linguistics* 26.17–44.

Pearlmutter, Neal J., and Edward Gibson. 2001. Recency in verb phrase attachment. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 27.574–90.

National Science Foundation
4201 Wilson Blvd., Room 1125
Arlington, VA 22230
[dlangend@nsf.gov]

[Received 10 July 2003;
accepted 27 February 2004]