Journal of Universal Language 3
September 2002, 1-30

The Universality of Morpho-Syntax: Synthetic Compounding in French, English, Dutch and Korean*

Benjamin Hebblethwaite
Indiana University

Abstract

This paper argues that synthetic compounding SC accesses phrasal syntax. The split-Inflectional Phrase (Split-IP) supplied by Universal Grammar provides a parsimonious representation of SC in French, English, Dutch and Korean. In SC, nouns and verbs are attracted leftward up the tree to check and eliminate strong nominal or verbal features located in higher functional projections. Split-Infl and leftward-only Move α provide a minimal and learnable framework for SC word formation. Kayne's Universal Base Hypothesis (1994) imposes a rigid syntactic theory (section 2). Data from French, English, Dutch and Korean exemplify SC in (3). A literature review in (4) evaluates recent work on SC found in Lieber (1992), Barbaud (1994) and Di Sciullo (1996). In (5), a minimalist syntactic solution provides a unified analysis of the data set. Finally, an X-bar

* Special thanks to Professor Barbara Vance for her constructive criticism. I am grateful to Changhee Rhee for her help with the Korean section, to Professor Stuart Davis for his helpful suggestions, to Ben Kloda who corrected a final draft, and to two anonymous reviewers for their perspicuous comments.

theory of categorical conversion and affixation in (6) argues in favor of a post-syntactic, morphological module that precedes PF in word formation. The conclusion suggests how generative morpho-syntactic theory might be applied to the development of a universal artificial language.

Keywords: morphology, syntax, synthetic compounding, universal base hypothesis, artificial language, minimalism, distributed morphology

1. Introduction

This paper is a contribution to the ongoing cartography of Universal Grammar. Generative syntactic theory provides an explanation for deverbal synthetic compounding SC. SC accesses syntactic processes in order to satisfy expressive needs minimally. Synthetic compounds are produced to act as nuanced NPs in sentential syntax.[1] The lexicon calls on syntax for semantic expansion, thus avoiding the unnecessary labor of creating and storing new phonological forms. Word formation thrives on syntax because this module of the language faculty permits the production of semantically salient nouns by means of familiar vocabulary.

This paper is a critique and an extension of Lieber (1992), Barbaud (1994), and Di Sciullo (1996) from a Kaynian minimalist framework (Kayne 1994; see also Chomsky 1995, Zwart 1997,

[1] Examples of synthetic compounds in ordinary usage:

(1) French: *Le tire-bouchon se trouve dans l'armoire.* (The cork-screw is located in the drawer.)

(2) English: *Truck drivers earn decent pay.*

(3) Dutch: *Bolletjeslikkers reizen tussen Suriname en Nederland.* (Cocaine smugglers travel between Surinam and the Netherlands.)

Roberts 1997a, Fischer 2000). Synthetic compounds are analyzed by means of a single phrase marker. I argue that this phrase marker, the "split-Inflectional system" [split-INFL] supplied by Universal Grammar, provides a parsimonious representation of SC in French, English, Dutch and Korean.[2]

Split-INFL splices the classic IP structure [$_{IP}$ [$_{I'}$ [$_{VP}$ [$_{V'}$]]]] into distinct functional projections. A verb's functional projections are construed as follows: [$_{\text{Agreement Subject Phrase}}$ [$_{AgrS'}$ [$_{\text{Tense Phrase}}$ [$_{T'}$ [$_{\text{Agreement Object Phrase}}$ [$_{AgrO'}$ [$_{\text{Verb Phrase}}$ [$_{V'}$]]]]]]]].[3]

SC word formation accesses phrasal syntax. In underived base structure, the verb assigns a θ-role to its internal argument, the complement. Attracted by strong Noun features NF in the specifier of a higher functional projection FP, called AgrOP, the N^{o} complement moves leftward to check and eliminate the NF. AgrOP, the landing-site, assigns structural Accusative Case to direct objects universally in sentential syntax (Pollock 1989) and acts as a structural receptacle devoid of Case-assignment in SC word formation. V is likewise susceptible to the attraction of strong V-features VF in a V-related FP called AgrSP. AgrSP checks the V's Φ-features in sentential syntax and acts as a structural receptacle for the production of synthetic compounds in word formation. The V or N movement operation only applies if strong VF or NF force movement. In word formation, syntactic operations feed Morphological Structure MS and MS, in

---

[2] See Chomsky (1991 in 1995) and Roberts (1997).

[3] In recent work vp-shells dominating AgrOP are employed universally (Koizumi 1993; also see Lasnik 1999). I do not employ the vp-shell analysis here, however. Koizumi's structure is: [$_{AgrSP}$ [$_{AgrS'}$ [$_{TP}$ [$_{T'}$ [$_{vp}$ [$_{v'}$ [$_{AspP}$ [$_{Asp'}$ [$_{VP}$ [$_{V'}$ [$_{DP}$]]]]]]]]]]]. Koizumi proposes that the V moves to v' in overt Modern English syntax (as opposed to only DP moving up to AgrOP). This idea doesn't necessarily pose problems for the analysis presented here if SC (as opposed to sentential syntax) has the property of delayed verb movement. This amounts to an additional stipulation, however. Bowers (2002: 183-224) argues for three main phrase marker types cross-linguistically. Transitive type A: [$_{PredicateP}$ [$_{DP}$] [$_{TransitiveP}$ [$_{VP}$ [$_{DP}$]]]]; Unergative type B: [$_{PredicateP}$ [$_{DP}$] [$_{TransitiveP}$ [$_{VP}$ ([$_{PP}$])]]]; Unaccusative type C: [$_{PredicateP}$ [$_{VP}$ [$_{DP}$]]].

turn, feeds Phonological Form PF.

French and Korean-type SC is derived as shown below:[4]

(1) French and Korean[5]

Base $[_{AgrSP}\ [_{AgrS'}VF\ [_{AgrOP}\ [_{AgrO'}\ [_{VP}\ [_{V'}\ \alpha\ V\ [_{DP/NP}$ DP/NP]]]]]]]

DS[6] $[_{AgrSP}\ [_{AgrS'}\ \alpha\ V_i\ [_{AgrOP}\ [_{AgrO'}\ [_{VP}\ [_{V'}\ t_i\ [_{DP/NP}$ NP/DP]]]]]]]

MS/PF $[_{NP}$ *verb* $[_{N'}$ *noun* + *ø* (i.e., no affix necessary)]]

English, Dutch and Korean-type SC is derived as shown below:[7]

(2) English, Dutch and Korean:

Base $[_{AgrSP}\ [_{AgrS'}\ [_{AgrOP}$ NF $[_{AgrO'}\ [_{VP}\ [_{V'}$ V $[_{N}\ \alpha$NP]]]]]]]

DS $[_{AgrSP}\ [_{AgrS'}\ [_{AgrOP}\ \alpha\ NP_j\ [_{AgrO'}\ [_{VP}\ [_{V'}$ V $[_{NP}\ t_j$]]]]]]]

MS/PF [NP *noun* [N' *verb* + N Affix]]

I present my theoretical assumptions in section 2. Section 3 provides SC data from French, English, Dutch and Korean. Section 4 is a literature review that addresses the advances and problems in recent analyses of Old and Modern English (Lieber 1992) and Modern French and English (Barbaud 1994, Di Sciullo 1996). Section 5 examines the mechanics of my syntactic solution. In my morphological solution, section 6, I analyze categorical conversion, i.e., how $[_{AgrSP}\ [_{AgrOP}\ [_{VP}\ [_{DP}\ [_{NP}\ ]]]]]$ becomes $[_{NP}\ [_{N'}\ ]]$, and propose a theory of affixation. The conclusion suggests how minimalist syntactic theory might be applied to the development of an artificial language

[4] The functional projection TP is omitted for ease of exposition.

[5] This type of SC can occur in English and Dutch, but it is not considered to be productive.

[6] DS stands for Deep Structure, i.e., a level of representation between Base Structure and Morphological Structure.

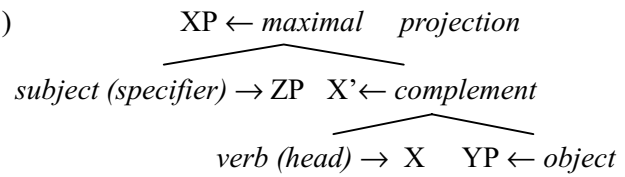
[7] Note that the English, Dutch and Korean phrase marker does not include Determiner Phrase DP dominating the direct object NP. The reason for this is elaborated in section 5.

acceptable to the international community.

2. Basic Assumptions

This theory of SC adopts the Universal Base Hypothesis, UBH, (Kayne 1994, Zwart 1997). The UBH states that UG supplies an SVO syntactic base as in (3) below. The UBH holds that all natural languages share this one syntactic foundation. The argument for this model of syntax is based on language acquisition, i.e., the language acquisition device must minimally generate all syntactic phenomena exhibited in the species. Poverty of L_1 stimulus in acquisition argues that universal principles must account for the mastery of L_1 in children (see Chomsky 1981, Cook & Newson 1996). The UBH phrase marker, presented in (1) below, is the most restricted formulation yet of these acquisition intuitions.

(3)



In the UBH phrase marker shown in (3) above, the specifier precedes the head and the head precedes the complement. Non-SVO surface word order results from applications of Move α where α may be a head X or a phrase XP, ZP or YP. In syntactic theory, the Base (Theta-marking) → Phonological Form PF (Feature-checking) → Logical Form LF (semantic interface) form a continuum of derivational representations in which the restricted (i.e., shortest movement), yet generalized, application of Move α dislodges α from its base position and lands α in LF. Either movement occurs overtly in PF or covertly, after Spell-Out, in LF. Move α operations are obligatory. Movement operations occur in order for lexical items to check their own features against abstract features (categorical, morpho-

logical or Case-related) located in the (Spec or head) nodes of higher functional projections (Chomsky 1995).[8]

In the word formation theory proposed here, the levels of derivation proceed as follows: Base → Deep Structure DS → Morphological Structure MS → Phonological Form PF. SC word formation accesses syntax and dislodges α from the Base and lands α in DS. DS feeds α to MS where a nominal X-bar structure concatenates an affix, if needed. The complex NP, the synthetic compound, is next realized in PF.[9]

Section 2 provides examples of synthetic compounds in Old and Modern French and English, and in Modern Dutch and Korean.

3. Data

The examples (4-11) presented below will be the object of analysis in sections 4, 5 and 6.

(4) French VO
a. un tire-bouchon ~ "a pull-cork," i.e., a cork-screw
b. un couvre-feu ~ "a cover-fire," i.e., a curfew
c. un gratte-ciel ~ "a scrape-sky," i.e., a sky-scraper
d. un porte-parole ~ "a carry-word," i.e., a spokesperson
e. une garde-malade ~ "a keep-sick," i.e., care-giver, someone who cares for the sick
(Barbaud 1994: 5)

(5) Old French VO
a. un gäaignepain ~ "earn-bread," i.e., type of sword used in

[8] The base structure position and the derived structure target (PF) may both be involved in the pressures that give rise to movement (Lasnik 1999).

[9] LF is not discussed with respect to synthetic compounds because the topic has little bearing on the argumentation I adopt.

tournaments
b. un gaaigne-oble ~ “an earn-penny,” i.e., a wage slave
c. un portepais ~ “a carry-peace,” i.e., case that contains the paten; i.e. someone who announces peace.
d. un porte fouet ~ “a carry-whip,” i.e., someone who carries a whip in hand
(Tobler-Lommatzsch 1969, Baldinger 1995)

The examples in (4) and (5) reflect the structure presented in (1) above. These synthetic compounds reflect the most productive categorical ordering found in French, i.e. V-N. French SC is treated in detail in examples (13-15) below.

(6) Modern English OV
a. a truck driver
b. a sky scraper
c. a call-minder (British) ~ voicemail
d. a blood-sucker
e. a mouse catcher
(Lieber 1992)

(7) Modern English VO
pickpocket

(8) Old English OV
a. dom settend ~ “doom setter,” i.e., a lawyer
b. faeder-swicca ~ “father traitor,” i.e., traitor to one’s father
c. ecg-bana ~ “sword’s edge commander,” i.e., a slayer
d. beorn-wiga ~ ‘braveman fighter,” i.e., a warrior
e. ealo-wosa ~ “ale-soaker,” i.e., a drunkard
(Sweet 1896)

(9) Dutch OV

a. een bolletjeslikker ~ “a little balloon swallower,” i.e., a cocaine trafficker
b. een wijndrinker ~ a wine drinker
c. een boekverkoper ~ a book seller
d. een wedstrijdspeler ~ “a competition player,” i.e., a competitor

e. een lesgever ~ “a lesson giver,” i.e., a teacher, instructor
(Booij & van Santen 1998: 171-182)

The examples in (6), (8) and (9) reflect the structure presented in (2) above. The synthetic compounds in (6), (8) and (9) reflect the most productive categorical ordering in English, Old English and Dutch, i.e., N-V. Example (7) shows V-N ordering like the French examples in (4) and (5). This is a rare and unproductive ordering. English, Old English and Dutch SC is treated in detail in examples (16-21) below.

(10) Korean VO
a. cep-khal ~ “fold-knife,” i.e., pocketknife
b. hel-kaps ~ “cheap-price,” i.e., inexpensive
c. nal-cimsung ~ “fly-animal,” i.e., winged animal
d. tah-soli ~ “touch-sound,” i.e., consonant
(Sohn 1994, 1999)

(11) Korean OV
a. cec-mek-i ~ “milk eat-i,” i.e., baby
b. chilphan-ttakk-i ~ “blackboard-erase-i,” i.e., blackboard eraser
c. hay-tot-i ~ “sun-rise-i,” i.e., sunrise
d. pap-pel-i ~ “rice-earn-i,” i.e., job

The examples in (10) and (11) reflect the structure presented in

(1) and (2) above. While O-V ordering as in (10) is more productive, V-O synthetic compounds like (11) also exist. Korean SC is treated in detail in examples (22-25) below.

4. Literature Review

Recent work on SC advances in descriptive adequacy, but problems remain. I adopt the strengths of Lieber (1992), Barbaud (1994), and Di Sciullo (1996), but raise counterarguments on decisive points.

Lieber assumes that SVO[10] "Licensing Conditions" form the default setting for English phrase structure. OV-ordered synthetic compounds reflect a marked setting. Lieber argues that direct object DO θ-role assignment discharged to the left in Old English embedded clauses and synthetic compounds via a "subdomain parameter," i.e., a directionality parameter.

Lieber takes SC in Modern English to be a residue of this OE setting, however, she assumes that Modern English only assigns the DO θ-role from left to right. In ME synthetic compounding the DO complement moves to the pre-nominal landing-site to satisfy Accusative Case-licensing (Lieber 1992).[11] I argue that object shift in SC is more related to morpho-lexical economy principles and that Accusative Case-licensing is not the issue.

In section 5, I argue that Lieber's "subdomain parameter" does not provide a straightforward explanation of Old English SC or its embedded clause syntax on the basis of the variable word ordering

[10] SVO = Specifier-Head-Complement order.

[11] In effect, this means that a movement operation derives OV SC in ME, whereas no movement occurred in OE due to a parameterized base. A parameterized SOV/SVO OE base requires stipulations with undesired consequences. Assuming that OE *gold-gifa*, i.e., "gold giver," was originally generated without movement implies the existence of a θ-role assignment "switch system" capable of discharging to the right in main clauses and to the left in embedded clauses.

(i.e., SOV *and* SVO) attested in embedded clauses.

Barbaud (1994) observes that the base verb for a French synthetic compound is usually 3rd person singular, present indicative (10-11). Compounds such as *un cessez-le-feu* and *un m'as-tu-vu* show that the verb is most generally finite. Masculine gender agreement suggests a syntactic operation (12).[12] An operator, the "hyperonomic subject," generalizes the gender morphology of lexical items that enter its domain.[13] The masculine output is a memory-saving device (12).[14]

Barbaud's theory of θ-roles and Case presents problems under currently standard assumptions. In French, he claims, I' θ-marks SpecIP "[+ θ]." Accusative Case-assignment discharges post-verbally.[15] Example (12) is a reproduction of his view:

---

[12] The many noun endings and suffixes that trigger predictable genders are an instance of economy in word formation. For example, of the 117 words ending in [-ro], as in *bureau*, 100% are masculine; of the 697 words ending in [-izm] as in *communisme*, 100% are masculine; and of the 79 words ending in [-az], as in *base*, 91.1% are feminine, etc. (for a list and analysis see Valdman 1976: 144-147). Another example of an economy "path" can be seen in serialization where adjective classes are inserted in predictable specifier nodes and not arbitrary ones (see Cinque 1994).

[13] Situating his argument in an X-bar system where verbs move to higher functional projections, Barbaud argues that all subjects deprived of person marking realize the (masculine) 3rd P.Sg. by default. He proposes the notion of the hyperonomic subject (a kind of "archisubject") as an account of the uniformity of the masculine, 3rd P.Sg. The hyperonomic subject may be understood as an abstract operator that generalizes the gender morphology of all forms that enter the domain of a particular process (compounding, suffixation, etc.). The output of French SC, thus, is a default setting based on an ([masc.], 3P.Sg.) archisubject.

[14] Note that "exceptions" in the uniformity of gender assignment like ***une*** *garde-malade* belong to a class of [animate] synthetic compounds that take up the gender of their referent, whereas [inanimate] French synthetic compounds converge on masculine grammatical gender as the unmarked option (Vance, personal communication).

[15] Barbaud claims that Romance deverbal compound nouns invariably exhibit a transitive (V-O) Case structure. In the minimalist framework, the DO receives its θ-role from the head V that dominates it. Case assignment, on the contrary, is structural. AgrOP structurally assigns Accusative Case in [$_{AgrOP}$ + Case [$_{VP}$ [$_{DP}$

(12) $[_{SN}\ [\ +\ \theta]_{SPEC}\ [_{I'}\ \textit{tire}_i\ [_{V'}\ e_i\ [\ \textit{bouchon}]\ _{N\ +\ accusative}\ ]]]$

In his account of OV English SC, Barbaud claims that it is impossible for English I' to Theta-mark SpecIP because, unlike French, V to I movement does not occur overtly. Barbaud assumes that Case assignment hinges on θ-role assignment, and since no V to I movement can θ-mark IP, the object itself cannot receive Case. This eliminates the requirement of a post-verbal object. Consequently, the object is inserted in SpecIP, producing the English-type OV synthetic compound.

According to the VP-internal subject hypothesis, however, subjects are θ-marked before moving to SpecIP (see Kitagawa 1986: Radford 1997: 329). Neither movement-dependent theta-marking nor the blocked Accusative Case assignment that ensues fit into the model of grammar I assume.

Di Sciullo (1996) interprets SC in a modular theory. She proposes the "Modularity of Computational Space MCS" principle. MCS characterizes the language faculty as a computational space in which derivations interact until the optimal target is obtained. The MCS does not feed components but is a single computational space where derivations work in parallel and interface.

The nominalization of the verb *work* in *John is a worker*, for example, involves two X-bar categories, Q^{-er} dominating VP^{work}. The suffix *–er* is the head of functional Q. Di Sciullo takes all derivational affixes to be functional projections dominating lexical projections, i.e., *–er* (Q, N), *-able* (MOD, A), etc.[16] She assumes that V *work* moves up and adjoins to the functional head Q^{-er}, giving *worker*. To be discussed in section 6, my solution involves superimposing an NP X-bar structure onto the output of split-INFL in Mor-

---

α]]] where α DP gets Case by movement.

[16] Prefixes that do not affect the head's category right-adjoin to the head, rather than induce movement.

phological Structure MS.

Di Sciullo's analysis of Italian *porta-documenti* involves inserting VP^{porta} in SpecNP, the head of which is $N^{documenti}$. She proposes movement of the head $N^{documenti}$ to the complement position V', i.e., $[_{NP}\,[_{VP}\,[_{V'}\,porta\,[_{DP}\,[_{D}\,[_{N}\,documenti_i\,]]]]]]\,[_{N'}\,t_i]]$. The problem with her proposal is that it assumes word formation accesses X-bar syntax in a system distinct from sentential syntax. If SC can be accounted for with the principles of sentential syntax, it is unnecessary to stipulate structures like $[_{NP}\,[_{VP}\,[_{DP}\,[_{N}\,\alpha_i\,]]]\,[_{N'}\,t_i\,]]$ (i.e., a head need not move to a complement position in its own specifier).

To explain the difference between V-N in French and N-V-ordering in English, Di Sciullo claims that weak V-features in the English DP cannot attract the DO. As a result, the DO moves higher than D to eliminate Case features. Unlike her French (VP in SpecNP) structure, two NP maximal projections dominate the English VP. The topmost $SpecNP^{bottle}$ is the DO's landing-site, it dominates the affixal NP^{-er} which is the landing site for the V^{open}, i.e., $[_{NP}\,bottle_j\,[_{NP}\,[_{N'}\,open_i\text{-}er\,[_{VP}\,[_{V'}\,t_i\,[_{DP}\,t_j\,]]]]]]$. Like Barbaud and Lieber, Di Sciullo (1996) analyzes synthetic compounding as the application of movement operations. Like them, she adopts X-bar form without the rigid principles that mediate the minimalist instantiation. Salient ideas guide all three authors, however, they fall short of providing a framework that is unified theory-internally or cross-linguistically.

In section 5 I will argue for a theory that encompasses the data uniformly.

5. Analysis: Syntax

In this section, I will argue that a simpler solution is to expand the functional framework by means of the split-INFL phrase marker and to reduce the SC model to the V-to-AgrS and N-to-AgrOP

movement parameters.[17] Following recent generative work, Accusative Case is structural and assigned uniformly in AgrOP at PF or LF (Chomsky 1995: 149-50). θ-role assignment is uniformly discharged by the lexical head X to its YP complement and by X' to its specifier ZP.

Synthetic compounds are modeled on patterns in sentential syntax. Differences can be accounted for straightforwardly with Split-INFL. The bracketed French phrase marker in (13) below illustrates the sentential basis of the synthetic compound in (14).

(13) French sentential structure: "Elle tire le bouchon."

Base [$_{AgrSP}$ [$_{AgrS'}$ [$_{AgrOP}$ [$_{AgrO'}$ [$_{VP}$ *Elle* [$_{V'}$ *tire* [$_{DP}$ [$_{D'}$ *le* [$_{NP}$ [$_{N'}$ *bouchon*]]]]]]]]]]

PF [$_{AgrSP}$ *Elle*$_j$ [$_{AgrS'}$ *tire*$_i$ [$_{AgrOP}$ [$_{AgrO'}$ [$_{VP}$ t$_j$ [$_{V'}$ t$_i$ [$_{DP}$ [$_{D'}$ *le* [$_{NP}$ [$_{N'}$ *bouchon*]]]]]]]]]]

LF [$_{AgrSP}$ *Elle*$_j$ [$_{AgrS'}$ *tire*$_i$ [$_{AgrOP}$ *le bouchon*$_k$ [$_{AgrO'}$ [$_{VP}$ t$_j$ [$_{V'}$ t$_i$ [$_{DP}$ t$_{k...}$ [$_{D'}$ [$_{NP}$ [$_{N'}$]]]]]]]]

"She pulls the cork."

In (13), the subject is θ-marked AGENT in SpecVP and moves to AgrSP to get Nominative Case and check its Φ-features (EPP). The finite V *tire* moves to check its Φ-features, adjoining to the functional head AgrS. The direct object, *le bouchon*, remains in-situ at Spell-Out and moves covertly at LF to check Accusative Case.[18]

---

[17] V to I as in [+movement] in French and [-movement] in English Dutch (embedded clauses).

[18] Vance points out that overt object shift in French is an undesirable analysis because it increases the number of parameters required. For example, if one analyzes a sentence such as, *Elle a deja tiré le bouchon,* which means, *She has already pulled the cork*, as in (1) below, a parallel analysis for Germanic is impossible, i.e. (2): **hij heeft al geslikt de bolletje*, meaning, *he has already swallowed the balloon*.

(1) [$_{AgrSP}$ *Elle*$_j$ [$_{AgrS'}$ *a*$_i$ [$_{TP}$ *deja* [$_{T'}$ *tiré*$_z$ [$_{AgrOP}$ *le bouchon*$_k$ [$_{AgrO'}$ [$_{VP}$ [$_{V'}$ t$_i$ [$_{VP}$ t$_j$ [$_{V'}$ t$_z$ [$_{DP}$ t$_{k...}$ [$_{D'}$ [$_{NP}$ [$_{N'}$]]]]]]]]]]]]]

Examples (14) and (15) show that in French little distinguishes SC word formation from the sentential derivation in example (13).

(14) French synthetic compounding structure: "un tire-bouchon"

Base	$[_{AgrSP}\ [_{AgrS'}\ [_{AgrOP}\ [_{AgrO'}\ [_{VP}\ [_{V'}\ \textit{tire}\ [_{NP}\ [_{N'}\ \textit{bouchon}\ ]]]]]]]]$
DS	$[_{AgrSP}\ [_{AgrS'}\ \textit{tire}_i\ [_{AgrOP}\ [_{AgrO'}\ [_{VP}\ [_{V'}\ t_i\ [_{NP}\ [_{N'}\ \textit{bouchon}\ ]]]]]]]$
MS/PF	$[_{NP}\ \textit{tire}\ [_{N'}\ \textit{bouchon}\ ]]$
	"cork-screw"

(15) French synthetic compounding structure: "un cessez-le-feu"

Base	$[_{AgrSP}\ [_{AgrS'}\ [_{AgrOP}\ [_{AgrO'}\ [_{VP}\ [_{V'}\ \textit{cessez}\ [_{DP}\ [_{D'}\ \textit{le}\ [_{NP}\ [_{N'}\ \textit{feu}\ ]]]]]]]]$
DS	$[_{AgrSP}\ [_{AgrS'}\ \textit{cessez}_i\ [_{AgrOP}\ [_{AgrO'}\ [_{VP}\ [_{V'}\ t_i\ [_{DP}\ [_{D'}\ \textit{l}\ [_{NP}\ [_{N'}\ \textit{feu}\ ]]]]]]]]]$
MS/PF	$[_{NP}\ \textit{cessez}\ [_{N'}\ \textit{le feu}\ ]]$
	"cease-fire"

In (14) and (15) no subject is inserted. Subject insertion is obviated since the output is an NP, i.e., the speaker has no use for an extra subject since she is creating a subject: *tire-bouchon*, *truck-driver*, etc. In (14) and (15) the verb has moved to AgrS' on the same path as (13) above. The N *bouchon* remains in-situ.[19]

A difference between (14) and (15) is the absence of a definite determiner in (14) and its presence in (15). Di Sciullo (1996) claims that the DP in French SC does not have specific reference because it cannot project its referential features to the root. I do not reach this

---

(2) *$[_{AgrSP}\ \textit{Hij}_j\ [_{AgrS'}\ \textit{heeft}_i\ [_{TP}\ \textit{al}\ [_{T'}\ \textit{geslikt}_z\ [_{AgrOP}\ \textit{de bolletje}_k\ [_{AgrO'}\ [_{VP}\ [_{V'}\ t_i\ [_{VP}\ t_j\ [_{V'}\ t_z\ [_{DP}\ t_{k\ldots}\ [_{D'}\ [_{NP}\ [_{N'}\ ]]]]]]]]]]]]]]$
Thus, by leaving the French direct object in-situ in overt syntax, it is not necessary to stipulate (cost-ineffectively) AgrO as the landing-site for the Germanic past participle (personal correspondence).

[19] I assume SC formation occurs before the N checks Accusative Case in AgrOP. This poses no problems since the output of SC is a compound noun and not a sentence, in other words, in SC, *bouchon*, is not an argument of *tire.*

conclusion. In (14) I assume that a bare NP such as α [$_{NP}$ [$_{N'}$ *bouchon*]] may merge directly to V' without a DP.[20] Semantically, a *tire-bouchon* is used on "corks," not on "*the* cork," so D is not required. Example (15), *cessez-le-feu*, shows that merging with D is possible in SC. The D *le* has first merged with NP *feu* and subsequently merged to V' as a DP phrase, i.e., α [$_{DP}$ [$_{D'}$ *le* [$_{NP}$ [$_{N'}$ *feu*]]]]. Semantically, unlike the tool *tire-bouchon*, which uncorks an indefinite number of *bouchons*, a *cessez-le-feu* refers to a particular state of war, *le feu*, which must cease, rather than some indefinite number.[21] The synthetic compound *couvre-feu*, i.e., "curfew," shows that the absence or presence of the determiner in French SC is not strictly related to semantic effects.

The analysis of Dutch below will lead us to Modern and Old English. As Zwart (1997) argues, a superficial sentential difference between Dutch and English phrase structure is the absence of overt verb and object movement in English.[22]

---

[20] Some classes of bare NPs may operate syntactically (i.e., receive a θ-role, move, etc.) without a DP dominating them. D-less NPs exist with certain classes of noun phrases in French, i.e., *Marie embrasse Jean* and not **La Marie embrasse le Jean.*

[21] Barbaud notes that determiners in synthetic compounds are unproductive in French and he provides the following closed list: *un trompe-la-mort*, *un traîne-la-patte*, *un tire-au-flanc*, *un trompe-l'oeil*, *un baise-la-piastre*.

[22] Note that while the surface ordering in Dutch is SVO like English, overt movement of the V and the Object is required in (1) in order to account for object shift in matrix clauses with an auxiliary verb in (2):

(1) Dutch sentential structure: "Hij slikt bolletjes."
Base [$_{AgrSP}$ [$_{AgrS'}$ [$_{AgrOP}$ [$_{AgrO'}$ [$_{VP}$ *Hij* [$_{V'}$ *slikt* [$_{NP}$ [$_{N'}$ *bolletjes*]]]]]]]]
PF [$_{AgrSP}$ *Hij*$_i$ [$_{AgrS'}$ *slikt*$_j$ [$_{AgrOP}$ *bolletjes* [$_{AgrO'}$ [$_{VP}$ t$_i$ [$_{V'}$ t$_j$ [$_{NP}$ t$_k$ [$_{N'}$]]]]]]]]
LF [$_{AgrSP}$ *Hij*$_i$ [$_{AgrS'}$ *slikt*$_j$ [$_{AgrOP}$ *bolletjes*$_k$ [$_{AgrO'}$ [$_{VP}$ t$_i$ [$_{V'}$ t$_j$ [$_{NP}$ t$_k$ [$_{N'}$]]]]]]]]
"He swallows little balloons."

(2) Dutch matrix clause with an auxiliary verb
PF/LF [$_{AgrSP}$ *Hij*$_i$ [$_{AgrS'}$ *heeft*$_j$ [$_{AgrOP}$ ***de bolletjes***$_k$ [$_{VP}$ [$_{V'}$ t$_j$ [$_{VP}$ t$_i$ [$_{V'}$*geslikt* [$_{DP}$. . . t$_k$]]]]]]
He had the little balloons swallowed.
"He swallowed the little (cocaine) balloons."

Dutch SC derives from the syntax of embedded clauses. Example (16) shows the embedded sentential correlate of the Dutch synthetic compound:

(16) Dutch embedded sentential structure: "Ze denken dat hij de bolletjes slikt"

Base *Ze denken* [$_{CP}$ [$_{C'}$ *dat* [$_{AgrSP}$ [$_{AgrS'}$ [$_{AgrOP}$ [$_{AgrO'}$ [$_{VP}$ *hij* [$_{V'}$ *slikt* [$_{DP}$ [$_{D}$ *de* [$_{NP}$ [$_{N'}$ *bolletjes*]]]]]]]]]

PF *Ze denken* [$_{CP}$ [$_{C'}$ *dat* [$_{AgrSP}$ *hij*$_{j}$ [$_{AgrS'}$ [$_{AgrOP}$ *de bolletjes*$_{i}$ [$_{AgrO'}$ [$_{VP}$ t$_{j}$ [$_{V'}$ *slikt* [$_{DP}$ t$_{i}$]]]]]]]]]

They think that he the little balls swallows.

"They think that he swallows the little balls."

In Dutch embedded clauses, the direct object moves up past the main verb (and its auxiliary, if it surfaces), revealing a sentential syntactic process that mirrors the SC operation. As shown in (17), the embedded clause in (16) suggests itself as the pattern of Dutch SC word formation.

(17) Dutch synthetic compounding structure: "bolletjeslikker"

Base [$_{AgrSP}$ [$_{AgrS'}$ [$_{AgrOP}$ [$_{AgrO'}$ [$_{VP}$ [$_{V'}$ *slikt* [$_{NP}$ [$_{N'}$ *bolletjes*]]]]]]]]

DS [$_{AgrSP}$ [$_{AgrS'}$ [$_{AgrOP}$ *bolletje*$_{k}$ [$_{AgrO'}$ [$_{VP}$ [$_{V'}$ *slik(ker)* [$_{NP}$ t$_{k}$ [$_{N'}$]]]]]]]]

MS/PF [$_{NP}$ *bolletje* [$_{N'}$ *slikker*]]

"little balloon swallower" ~ "smuggler who swallows cocaine-filled balloons"

The embedded analysis works neatly for Dutch, but consider now Modern English. I assume the following sentential correlate for English synthetic compounds:

See Zwart (1997) for further discussion.

(18) English matrix (or embedded) sentential structure: “She drives trucks.”

Base	$[_{AgrSP}\ [_{AgrS'}\ [_{AgrOP}\ [_{AgrO'}\ [_{VP}$ *She* $[_{V'}$ *drives* $[_{NP}\ [_{N'}$ *trucks*]]]]]]]]
PF	$[_{AgrSP}$ *She*$_i$ $[_{AgrS'}\ [_{AgrOP}\ [_{AgrO'}\ [_{VP}\ t_i\ [_{V'}$ *drives* $[_{NP}\ [_{N'}$ *trucks*]]]]]]]]
LF	$[_{AgrSP}$ *She*$_i$ $[_{AgrS'}$ *drives*$_j$ $[_{AgrOP}$ *trucks*$_k$ $[_{AgrO'}\ [_{VP}\ t_i\ [_{V'}\ t_j\ [_{NP}\ t_k\ [_{N'}\ ]]]]]]]]

In (18) above, the verb and the object remain in situ at Spell-Out. The V moves to AgrS and the object moves to AgrOP covertly in LF to check Case and Φ-features. The most productive English SC structure is shown below in (19):

(19) English synthetic compounding structure: “truck driver”

Base	$[_{AgrSP}\ [_{AgrS'}\ [_{AgrOP}\ [_{AgrO'}\ [_{VP}\ [_{V'}$ *drives* $[_{NP}\ [_{N'}$ *trucks*]]]]]]]]
DS	$[_{AgrSP}\ [_{AgrS'}\ [_{AgrOP}$ *truck*$_i$ $[_{AgrO'}\ [_{VP}\ [_{V'}$ *drive(r)* $[_{NP}\ t_i$ $[_{N'}\ ]]]]]]]]
MS/PF	$[_{NP}$ *truck* $[_{N'}$ *driver*]]

Why is it that English synthetic compounds are OV when both matrix and embedded clauses are (S)VO in Modern English syntax?[23] In the next paragraphs, I will argue that Modern English SC is a vestige of the SOV word order that was “predominant” (but not exclusive) in derived Old English embedded clauses (Roberts 1997a: 400; Fischer 2000: 46-53, 138-72).

This discussion requires briefly revisiting Lieber’s claims. Due

[23] English and Dutch do not produce many VO-ordered synthetic compounds like French. The evidence suggests that Old English and Dutch synthetic compounds were largely modeled on embedded clause object shift. Object shift around the embedded V became a “path” of word formation. Note, however, that it is possible that the embedded clause pattern (OV) coexisted with a more French-type VO pattern. Evidence for this is sparse but suggestive: *pickpocket*.

to the fact that SOV was optional in OE embedded clauses (as opposed to Dutch, where it is obligatory), as mentioned in section 4, Lieber's OE "subdomain parameter" cannot be predetermined at the clausal level.[24] The word order of embedded clauses is in fact unpredictable, therefore her theory must assume that verbs are subcategorized with information regarding the directionality of θ-role assignmen.[25] Such subcategorization of verbs is undesirable if the behavior of objects can fall out from more general processes. The prose examples (20) and (21) below, taken from Fischer (2000: 51-52), show the *optional* shift of the DO in embedded OE clauses:

(20) paet hi *mihton* swa bealdlice Godes
that they might so boldly God's
geleafan$_{AgrOP}$ *bodian*
faith preach
"that they *might preach* God's faith so boldly"

In embedded (20), the direct object, *Godes geleafan*, appears before the main verb, *bodian*, i.e., (20) is OV. In embedded (21), the direct object appears after both the auxiliary, *magon*, and the main verb, *geoffrian*, i.e., (21) is VO:

(21) paet hi urum godum *geoffrian* *magon*
that they our gods offer may
dancwurde onsaegednysse
grateful sacrifice
"that they may offer a grateful sacrifice to our gods"

Given only (20), Lieber's claim that base structure embedded clauses are parameterized SOV is perhaps tenable. Example (21),

[24] Clausal determination means: embedded clause = automatic SOV.
[25] I do not believe she addresses this problem.

however, poses problems for her analysis. The main verb *geoffrian* would presumably have a subcategorization frame stipulating the rightward assignment of the DO θ-role, i.e., (*ge*)*offrian*: [NP__].[26]

In the approach adopted here, on the contrary, the direct objects in (20) and (21) both receive their θ-roles as internal arguments of the main V. In (20) the main verb's internal argument *Godes geleafan* moves to AgrOP by PF. In (21) the main verb's internal argument *dancwurde onsaegednysse* remains in situ in PF.[27] The structural assignment of θ-roles in base structure coupled with the massive, but restricted, application of leftward Move α yields a more learnable model of syntax since operations are structural and lexical items are not idiosyncratically subcategorized.

Therefore, while object shift is extinct in ME embedded syntax, it remains robust in SC word formation.[28]

Setting aside the debate over the applicability of Kayne (1994) to Korean and Japanese, I will assume Korean is SVO in base structure in order to test the analysis developed above. OV Synthetic compounds like *cec-mek-i*, "milk-eat-er," i.e., "baby," are analyzed as in Dutch (17) or English (19) (Korean examples from Sohn 1999: 246):

(22) Korean synthetic compounding structure: *cec-mek-I* ~ "milk eat-er," i.e., "baby"

Base $[_{AgrSP}\ [_{AgrS'}\ [_{AgrOP}\ [_{AgrO'}\ [_{VP}\ [_{V'}\ \textit{mek}\ [_{NP}\ [_{N'}\ \textit{cec}\ ]]]]]]]]$

DS $[_{AgrSP}\ [_{AgrS'}\ [_{AgrOP}\ \textit{cec}_k\ [_{AgrO'}\ [_{VP}\ [_{V'}\ \textit{mek(-i)}\ [_{NP}\ t_k\ [_{N'}\ ]]]]]]]]$

[26] Her other option is to stipulate the *rightward movement of the DO.

[27] Note, however, that the main verb itself has moved to a position higher than the auxiliary verb in the syntax. Roberts (1997) observes the particle-verb order possible in OE and Dutch embedded clauses and also analyzes it as verb raising (400-1).

[28] Roberts describes the reasons for the extinction of SOV in OE as follows, "[. . .] loss of movement dependencies [. . .] caused by changes in abstract features of functional heads" (1997: 399).

MS/PF $[_{NP}$ *cec* $[_{N'}$ *mek-i*]]

Unlike French, English or Dutch, Korean PF is SOV in both matrix and embedded clauses. In Korean both auxiliary and main verbs follow the subject and object. I assume that compounds like *cec-mek-i* in (22) are modeled on sentential movement patterns as shown in (23):

(23) Korean sentential structure: *cecul mekta* ~ "She drinks milk."
Base $[_{AgrSP}\ [_{AgrS'}\ [_{AgrOP}$ *(-ul)* $[_{AgrO'}\ [_{VP}\ [_{V'}$ *mekta* $[_{NP}\ [_{N'}$ *cec*]]]]]]]
PF $[_{AgrSP}\ [_{AgrS'}\ [_{AgrOP}$ *cecul*$_k$ $[_{AgrO'}\ [_{VP}\ [_{V'}$ *mekta* $[_{NP}$ t$_k$ $[_{N'}$]]]]]]]

In (23) the base structure V *mekta* assigns the complement *cec* a THEME θ-role, the complement then moves to AgrOP to check Accusative Case by PF.

Interestingly, while Korean clauses are uniformly SOV, V-O synthetic compounds exist, although they are less productive (Sohn 1999: 246). A synthetic compound like *cep-khal*, "fold-knife," i.e., "pocketknife," suggests two possible solutions. One involves moving the direct object *khal* to AgrOP and the verb *cep* to AgrS, the other, which I will adopt, procrastinates the DO in base structure, as in the French structure in examples (14-15). This maintains verb movement as the crucial determinant of SC word order.

(24) Korean synthetic compounding structure: *cep-khal* ~ "fold knife," i.e., "pocketknife"
Base $[_{AgrSP}\ [_{AgrS'}\ [_{AgrOP}\ [_{AgrO'}\ [_{VP}\ [_{V'}$ *cep* $[_{NP}\ [_{N'}$ *khal*]]]]]]]
DS $[_{AgrSP}\ [_{AgrS'}$ *cep*$_j$ $[_{AgrOP}\ [_{AgrO'}\ [_{VP}\ [_{V'}$ t $_j$ $[_{NP}\ [_{N'}$ *khal*]]]]
MS/PF $[_{NP}$ *cep* $[_{N'}$ *khal*]]

The sentential equivalent of (24) in (25) below illustrates the sentential requirement that direct objects move up to AgrOP in Korean. Given the consistently higher landing-site of the direct object

in Korean sentential syntax, the synthetic compound in (24) begs an explanation. Speculatively, N-features in AgrOP may simply be weak in this type of Korean SC. The resultant absence of object shift could boost semantic saliency by setting such synthetic compounds apart from sentential syntax. VO synthetic compounding in SOV sentential syntax (Korean) is the mirror image of OV SC in SVO sentential syntax, i.e., English.[29]

(25) Korean sentential structure: *hakseng-i khal-ul cepta* ~ "The tudent folds the knife."

PF [$_{AgrSP}$ [$_{AgrS'}$ [$_{AgrOP}$ [$_{AgrO'}$ [$_{VP}$ *hakseng* [$_{V'}$ *cepta* [$_{NP}$ [$_{N'}$ *khal*]]]]]]]]

LF [$_{AgrSP}$ *hakseng-i* [$_{AgrS'}$ [$_{AgrOP}$ *khal-ul*$_k$ [$_{AgrO'}$ [$_{VP}$ t$_j$ [$_{V'}$ *cepta* [$_{NP}$ t$_k$ [$_{N'}$]]]]]]]]

It is important to be able to explain the presence or absence of affixes in SC. Korean SC does not employ the same affixation in SC as it does in sentential syntax. As shown in (22), OV-order in SC does not trigger the Accusative *–ul/-rul* suffixation on *cec* in *cec-mek-i*, whereas the sentential OV-order does, i.e., (23) *cecul mekta*. I will suggest that rather than marking the object, Morphological Structure applies nominal suffixation to the verb when it enters the NP conversion process. Like English N-V SC, a (Spec-head ordered) NP-structure concatenates an affix to any [–N] head. The V-N-ordered synthetic compound needs no nominal affixation, i.e., *cep-khal* but not **cep-khal-i* since the head *khal* is a noun and matches with the head N features of the NP-structure. This process is taken up in greater detail in section 6.

I have argued that SC can be subsumed under split-INFL phrase structure and the V-to-AgrS **and N-to-AgrOP** movement parameter.

[29] The strength feature, by hypothesis, may be, like the one operative in Modern English, the vestige of an extinct sentential pattern.

Section 6 adopts a theory of categorical conversion and affixation in order to account for the finishing touches that Morphological Structure applies to the output of split-INFL.

6. Analysis: Morphology

Conversion theory proposes that the output of a word formation operation in syntax converts into an easily stored categorical unit. It is unlikely that movement operations continue to generate synthetic compounds once they have become productive in the lexicon. Conversion from syntax to lexicon involves superimposing NP X-bar categorical structure on the output of syntax in order to minimize storage in and maximize retrieval from the lexicon (see Barbaud 1994: 7).

A guiding assumption behind categorical superimposing is the concept of isomorphism. As defined by Bussman (1998), isomorphism is the notion that structural equivalence exists with respect to relations between elements of two or more sets. Crystal (1997) describes isomorphism as a property of two or more structures whose constituent parts are in one-to-one correspondence with each other at a given level of abstraction. Isomorphism suggests how the syntactic phrase marker $[_{AgrSP}\ [_{AgrS'}\ [_{AgrOP}\ [_{AgrO'}\ [_{VP}\ [_{V'}\ [_{DP}\ [_{NP}\ ]]]]]]]]$ reduces to the simplified categorical lexical entry $[_{NP}\ [_{N'}\ ]]$.

The syntactic operation applied at the inception of SC is not useful once the synthetic compound attains currency. The conversion of a synthetic compound from syntax, $[_{AgrOP}$ *truck*$_i$ $[_{AgrO'}\ [_{VP}\ [_{V'}$ *drive(-r)* $[_{NP}\ [_{N'}\ t_i]]]]]]$, to the lexicon, $[_{NP}$ *truck* $[_{N'}$ *driver*$]]$, is an isomorphic process in which syntactic output is made available to an atomic X-bar NP structure. The lexicon/morphology superimposes an NP X-bar structure over the $O^{modifier}V^{head}$ or $V^{head}O^{modifier}$ output.

The NP-structure lexicalizes the syntactic output as an 'atomic' NP. The operation is employed to satisfy processing constraints in

an economical way. The operation can be illustrated by the superimposition of derived structure square-bracketing upon syntactic structure round-bracketing, i.e., "[. . .] + {. . .}". English-type N-V → NP isomorphism is shown below with *truck-driver*:

Syntactic output: $\{_{\text{AgrOP}}$ *truck* $\{_{\text{Agr}}, \{_{\text{VP}} \{_{\text{V}}, drive_i \{_{\text{DP}} \{_{\text{D}}, t_i \}\}\}\}\}\}$
Synthetic compound atomization structure: $[_{\text{NP}} [_{\text{N}},]]$
The process: $[_{\text{NP}} \{_{\text{AgrOP}}$ *truck* $\{_{\text{VP}} \{_{\text{V}}, [_{\text{N}},$ *drive –r* $\}\}\}\}]]$
'Visible' derived structure output of process: $[_{\text{NP}}$ *truck* $[_{\text{N}},$ *driver*$]]$

The French-type V-N → NP isomorphism is shown below with *tire-bouchon*:

Syntactic output: $\{_{\text{AgrS}},$ *tire* $_i \{_{\text{AgrOP}} \{_{\text{Agr}}, \{_{\text{VP}} \{_{\text{V}}, t_i \{_{\text{NP}} \{_{\text{N}},$ *bouchon* $\}\}\}\}\}\}$
Synthetic compound atomization structure: $[_{\text{NP}} [_{\text{N}},]]$
The process: $[_{\text{NP}} \{_{\text{AgrS}},$ *tire*$_i \{_{\text{AgrOP}} \{_{\text{VP}} \{_{\text{V}}, t_i [\{_{\text{NP}} \{_{\text{N}},$ *bouchon* $\}\}\}\}\}]]$
'Visible' derived structure output of process: $[_{\text{NP}}$ *tire* $[_{\text{N}},$ *bouchon*$]]$

The end result of conversion is an "atom" to which syntax has no more access.[30]

The conversion analysis above occurs in a post-Move α (post-syntactic) morphological domain. Such a morphological domain has been dubbed Morphological Structure MS in Distributed Morphology DM (Halle & Marantz 1993: 112). In DM, Morphological Structure interfaces syntax and phonology. The output of the syntax module relates with "bundles" of morphophonological features in MS. Terminal nodes lack such morphophonological features until entering MS. My analysis follows DM in assuming that effects in MS are predictable after syntax has applied, i.e., the presence or ab-

[30] Bisetto & Scalise (1999) provide examples of synthetic compound atomicity tests such as insertion.

sence of affixation results from the ordering N-V versus V-N.

Linear order predicts the presence or absence of morphology in SC. Note that none of the French synthetic compounds exhibit affixation on either verb or object, for example, *gratte-ciel*, but not, **gratteur-ciel* or **gratte-cieleur*, as opposed to English, *sky-scraper*, but not, **sky-scrape* or **skyer-scrape*. This difference suggests that the affixation of *–er* only occurs once the output of split-INFL has concatenated (shared information) with the NP conversion structure (see Epstein 1999). In the French *gratte-ciel*, the N *ciel*, adjoins to N, so no nominal suffix is needed. In English, however, the head V adjoins to N and projects its features into N. The head N, in response, concatenates a nominal affix in the presence of V. Observe also that when English SC follows the French-type V-N order, no noun morphology is needed on the N, i.e., *pickpocket* not **pickpocketer*. Likewise, as shown above, Korean SC displays French-type traits when ordered V-N, i.e., *cep-khal* not **cep-khal-i*, and English traits when ordered N-V, i.e., *cec-mek-i* not **cek-mek*. Therefore Deep Structure (i.e., pre-MS) ordering allows one to predict the outcome of MS.

Conversion theory suggests that word formation output from the syntactic module of grammar is processed by other modules, i.e., morphology, phonology and the lexicon. It also suggests that "NP superimposition" inserts nominalization suffixes such as *–er* in *truck-driver* as needed. This argues that morphology is post-syntactic in the vein of Distributed Morphology. Most striking is the possibility that a specifier-head ordered NP X-bar conversion-structure has the potential to account for SC morphological facts from a wide array of languages. In the conclusion I examine the significance of these findings with regard to the planning of an international artificial language.

7. Conclusion: Implications for a Universal Artificial Language

Theoretical syntax in the generative tradition may prove to offer unique insights in the development of an international auxiliary language.

Sapir, Bloomfield & Boas (1925) mistakenly argued that a universal artificial language UAL "should be built as far as is possible from materials which are familiar to speakers of West European languages" (cited in Large 1985: 183). Such views are impossible today, as Chung (1996) notes, an AL like Esperanto is a poor choice for a UAL because it derives from one language family out of at least seven recognized in the world. Gledhill (2000) provides the following statistics on the sources of Esperanto vocabulary from a corpus of 350,000 words: Latinate 70 %, Esperanto 12 %, Germanic 10 %, Indo-European 5 %, Greek 2 %, Balto-Slavic 1 %. A UAL derived from a single language family alienates the speakers of other families.

Sapir, Bloomfield, & Boas (1925) rightly called for the UAL to have the "simplest grammatical structure" and greatest "flexibility of structure." With respect to Esperanto syntax, Gledhill (2000: 87-90) reports that SVO predominates and SOV may optionally occur with preverbal pronouns as in French. OSV corresponds to interrogatives. The remaining clause types are rare and associated with poetry. His corpus yields the following number of clause-types: SVO (705), OSV (258), SOV (51), OVS (27), VOS (3), and VSO (1). These numbers suggest that Esperanto syntax does not accommodate speakers of obligatory SOV natural languages such as Korean and Japanese, to name but two.

In a generative interpretation, the split-INFL phrase marker accommodates any conceivable linear order in syntax. The most adaptable UAL would offer syntactic processes familiar to the majority of natural language speakers. Two coexisting patterns might

exist. SVO word order could be available with *optional* Case morphology on the argument structure. SOV/OSV/VSO/VOS/OVS word order could be available with *obligatory* Case morphology on the argument structure. Thus, in SVO utterances Case would be marked by position and in non-SVO utterances, by affix. A flexible UAL should accommodate the syntactic resources of the maximal number of natural grammars.

With respect to synthetic compounds, the natural language findings reported in this paper suggest that a (Spec-head ordered) X-bar NP structure converts the output of split-INFL into the compacted form of a compound noun. If the lowest constituent is a noun, no affix is needed, if it is a verb, suffixation must establish nominal categorical status. In Esperanto, like Korean, both patterns exist. Examples for (1), that is, $[_{NP}$ *verb* $[_{N'}$ *noun* + *ø*], are *manghoro* ~ 'eat hour,' i.e., meal-time, time for eating, *skribmasino* ~ 'write machine,' i.e., typewriter. Examples for (2), that is, [NP *noun* [N' *verb* + N Affix], are *labordonanto* ~ 'work-giver,' i.e., employer, *laborprenanto* ~ 'work-taker,' i.e., employee, *paperfaristo* ~ papermaker (Krause 1987: 36, Gledhill 200: 67, Kellerman 190: 132). It is plausible that both patterns be available in UAL SC just as they are in many natural languages.

In conclusion, this contribution to comparative linguistics encompasses data from French, English, Dutch and Korean within a single theory, the computational system for human language C_{HL}. The Universal Base Hypothesis (Kayne 1994) imposes a restricted theory on syntactic analysis: by means of a single base, the language learner maps and generates language via (language-specific) applications of left-only Move α. The lexical items (α) of languages K, Z and Y Move from a common base up to functional landing-sites which feed semantic interpretation. The variation in word order exhibited in the language set examined here reduces to the parameter Move or Do Not Move α. Finally, a morphological module MS straightforwardly accounts for affixation and the conversion from

syntax to Phonological Form and the lexicon.

References

Baldinger, K. 1995. *Dictionnaire Etymologique de L'ancien Français*. Tübingen: Max Niemeyer Verlag.

Barbaud, P. 1994. Conversion Syntaxique. *Lingvisticae Investigationes* 18.1, 1-26.

Bisetto, A. & Scalise, S. 1999. Compounding: Morphology or Syntax? *The Boundaries of Morphology and Syntax* 31-48. Amsterdam: Benjamins.

Booij, G. & A. van Santen. 1998. *Morfologie: De Woordstructuur van het Nederlands*. Amsterdam: Amsterdam University Press.

Botha, R. 1984. *Morphological Mechanisms: Lexicalist Analyses of Synthetic Compounding.* Oxford: Pergamon.

Bowers, J. 2002. Transitivity. *Linguistic Inquiry* 33.2, 183-225.

Bussman, H. (ed.). 1998. *Routledge Dictionary of Language and Linguistics*. New York: Routledge.

Chomsky, N. 1981. *Lectures on Government and Binding*. Berlin: Walter de Gruyter.

Chomsky, N. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

Chung, Y. 1996. An International Language for the World to Come. *Journal of Universal Language* 3, 56-70.

Cinque, G. 1994. On the Evidence for Partial N-movement in the Romance DP. In G. Cinque et al (eds.), *Paths toward Universal Grammar* 85-110. Washington, DC: Georgetown University Press.

Cook, V. & M. Newson. 1996. *Chomsky's Universal Grammar*. Malden, MA: Blackwell.

Crystal, D. 1997. *A Dictionary of Linguistics & Phonetics.* Oxford: Blackwell.

Di Sciullo, A. 1996. Modularity and X°/XP Asymmetries. *Linguistic Analysis* 26, 3-28.

Epstein, D. 1999. Un-Principled Syntax: The Derivation of Syntactic Relations. In S. Epstein & N. Hornstein (eds.), *Working Minimalism* 317-345. Cambridge, MA: MIT Press.

Fischer, O., W. van Kemenade, & W. van Der, 2000. *The Syntax of Early English.* Cambridge: Cambridge University Press.

Franks, S. 2002. Advanced Syntax L643 Class Notes. Bloomington, IN: Indiana University.

Gledhill, C. 2000. *The Grammar of Esperanto: A Corpus-based Description.* Muenchen: Lincom Europa.

Hacken, P. 1994. *Defining Morphology: A Principled Approach to Determining the Boundaries of Compounding, Derivation and Inflection*. Zürich: Geort Olms Verlag.

Halle, M. & A. Marantz. 1993. Distributed Morphology and the Pieces of Inflection. In K. Hale, S. Keyser, & J. Samuel (eds.), *The View from Building* 20, 111-176. Cambridge, MA: MIT Press.

Kayne, R. 1994. *The Antisymmetry of Syntax*. Cambridge, MA: MIT Press.

Kellerman, I. 1910. *A Complete Grammar of Esperanto*. New York: D.C. Heath and Company.

Kitagawa, Y. 1994. *Subjects in Japanese and English.* New York: Garland.

Koizumi, M. 1993. Object Agreement Phrases and the Split VP Hypothesis. *MIT Working Papers in Linguistics* 18, 99-148.

Krause, E. 1983. Wörterbuch: Deutsch-Esperanto. Leipzig: VEB Verlag Enzyklopädie.

Large, A. 1985. *The Artificial Language Movement.* Oxford: Blackwell.

Lasnik, H. 1994. Verbal Morphology: Syntactic Structures Meets the Minimalist Program. *Minimalist Analysis* 97-115. Malden, MA: Blackwell.

Lasnik, H. 1999. On Feature Strength: Three Minimalist Approaches to Overt Movement. *Linguistic Inquiry* 30.2, 197-217.

Lardiere, D. 1998. Parameter-resetting in Morphology: Evidence from Compounding. *Morphology and Its Interfaces in Second Language Knowledge* 283-305. Amsterdam: Benjamins.

Lieber, R. 1992. *Deconstructing Morphology*. Chicago, IL: University of Chicago Press.

Pollock, J. 1989. Verb Movement, Universal Grammar, and the Structure of IP. *Linguistic Inquiry* 20.3, 365-424.

Roberts, I. 1997a. Directionality and Word Order Change in the History of English. *Parameters of Morphosyntactic Change* 397-426. Cambridge: Cambridge University Press.

Roberts, I. 1997b. *Comparative Syntax.* London: Arnold.

Radford, A. 1997. *Syntactic Theory and the Structure of English*. Cambridge: Cambridge University Press.

Sohn, H. 1994. *Korean*. New York: Routledge.

Sohn, H. 1999. *The Korean Language.* Cambridge: Cambridge University Press.

Sweet, H. 1967. *Sweet's Anglo-Saxon Reader in Prose and Verse*. Oxford: Oxford University Press.

Sweet, H. 1928. *The Student's Dictionary of Anglo-Saxon*. Oxford: Oxford University Press.

Tobler, A & E. Lommatzsch. 1969. *Altfranzösisches Wörterbuch*. Wiesbaden: Franz Steiner Verlag.

Valdman, A. 1976. *Introduction to French Phonology and Morphology*. Rowley, MA: Newbury House Publishers.

Vance, B. 2002. French Morphology F678 Class Notes. Bloomington, IN: Indiana University.

Zink, G. 1987. *L'ancien Français (XI^{e}-XIII^{e} Siècle)*. Paris: Presses Universitaires de France.

Zwart, J. 1997. The Germanic SOV Languages and the Universal Base Hypothesis. In C. Haegeman & A. Liliane (eds.), *The New Comparative Syntax* 246-67. New York: Longman.

Zwart, J. 1997. *The Morphosyntax of Verb Movement*. Dordrecht: Kluwer Academic Publishers.