Essentials of Semantic Syntax: an appetiser

Pieter A.M. Seuren
Max Planck Institute for Psycholinguistics
PO Box 310
6500 AH Nijmegen - Netherlands
pieter.seuren@mpi.nl

Lecture Abralin 30 June 2020

Abstract

Semantic Syntax (SeSyn), originally called Generative Semantics, is an offshoot of Chomskyan generative grammar (ChoGG), rejected by Chomsky and his school in the late 1960s. SeSyn is the theory of algorithmical grammars producing the well-formed sentences of a language L from the corresponding semantic input, the Semantic Analysis (SA), represented as a traditional tree structure diagram in a specific formal language of incremental predicate logic with quantifying and qualifying operators (including the truth functions), and with all lexical items filled in. A SeSyn-type grammar is thus by definition transformational, but not generative. The SA originates in cognition in a manner that is still largely mysterious, but its actual form can be distilled from the Surface Structure (SS) of the sentences of L following the principles set out in SeSyn. In this presentation we provide a more or less technical résumé of the SeSyn theory. A comparison is made with ChoGG-type grammars, which are rejected on account of their intrinsic unsuitability as a cognitive-realist grammar model. The ChoGG model follows the pattern of a 1930s neopositivist Carnap-type grammar for formal logical languages. Such grammars are random sentence generators, whereas, obviously, (nonpathological) humans are not. A ChoGG-type grammar is fundamentally irreconcilable with a mentalist-realist theory of grammar.

The body of the paper consists in a demonstration of the production of an English and a French sentence, the latter containing a classic instance of the cyclic rule of Predicate Raising (PR), essential in the general theory of clausal complementation yet steadfastly repudiated in ChoGG for reasons that have never been clarified. The processes and categories defined in SeSyn are effortlessly recognised in languages all over the world, whether indigenous or languages of a dominant culture—taking into account language-specific values for the general theoretical parameters involved. This property makes SeSyn particularly relevant for linguistic typology, which now ranks as the most promising branch of linguistics but has so far conspicuously lacked an adequate theoretical basis.

I principi della Sintassi Semantica: un assaggio

Sommario

La Sintassi Semantica (SeSyn), originariamente Semantica Generativa, è un ramo della grammatica generativa Chomskiana (ChoGG), respinta da Chomsky ed i suoi negli anni '60. SeSyn è la teoria di grammatiche algoritmiche che producono le frasi ben formate di una lingua L a partire dall'input semantico corrispondente, la Analisi Semantica (SA), rappresentata nella forma di una struttura ad albero linguistico in una lingua specifica formale per una logica del predicato incrementale con operatori di quantificazione e di qualificazione, incluse le funzioni di verità, e con gli elementi lessicali già riempiti. Una grammatica tipo SeSyn è quindi trasformazionale per definizione, ma non generativa. La SA nasce dalla cognizione in un modo

tuttora in gran parte misteriosa, ma la sua forma precisa si lascia distillare dalla Struttura Superficie (SS) delle frasi di L, secondo i principi specificati in SeSyn. In questa presentazione provvediamo un breve riassunto più o meno tecnico della teoria SeSyn. Un paragone è fatto con il modello grammatico ChoGG, il quale è scartato a causa della sua inettitudine intrinseca per rappresentare una grammatica in qualsiasi senso realistico-cognitivo. Il modello ChoGG segue il disegno di 'grammatica' presentato fra altri da Carnap negli anni '30 del secolo scorso per le lingue logiche formali. Tali grammatiche sono generatori random di frasi, mentre, ovviamente, un essere umano (non patologico) è tutt'altro che questo. Questo modello è fondamentalmente irriconciliabile con una teoria grammaticale mentalista-realista.

La maggior parte di questa presentazione consiste in una dimostrazione a titolo d'esempio della produzione di due frasi, una inglese, l'altra francese. Quest'ultima è un caso classico della regola ciclica di Salita del Predicato (Predicate Raising, PR), essenziale nella teoria generale della complementazione grammaticale ma nondimeno ripudiata sistematicamente dalla scuola Chomskiana per motivi mai chiarificati. Le procedure e le categorie della SeSyn si riconoscono immediatamente nelle lingue del mondo, siano esse lingue indigene o lingue culturali dominanti—prendendo in conto i valori specifici di ogni lingua particulare per i parametri generali della teoria. Questa proprietà conferisce alla SeSyn una rilevanza particolare per la tipologia linguistica, oramai considerata il ramo più promettente della linguistica, malgrado la mancanza palese di una base teorica adeguata.

Classification: Theoretical Essay

Keywords: algorithm, cognitive realism, Cycle, Generative Grammar, Generative Semantics, grammar model, mentalism, neopositivism, Postcycle, Predicate Raising, proposition, Semantic Analysis, Semantic Syntax, typology

algoritmo, Analisi Semantica, Ciclo, Grammatica Generativa, mentalismo, modello grammaticale, neopositivismo, Postciclo, proposizione, realismo cognitivo, Salita del Predicato, Semantica Generativa, Sintassi Semantica, tipologia

1 General background

In this lecture I intend to give you a taste of what *real* syntax amounts to and how exciting it can be as the regularities pop up and the system unfolds itself. A *Semantic Syntax* (SeSyn) grammar (Seuren 2018a) is an algorithmic system producing the sentences S of a language L from the corresponding semantic input, the *semantic analysis* (SA) of S. (SAs are formulated in a logical language (Seuren 2010), with quantifying and qualifying operators, but we will not go into such foundational matters here.) The SA is taken to originate in cognition and to be cast into the format of what is known as *a tree structure*.

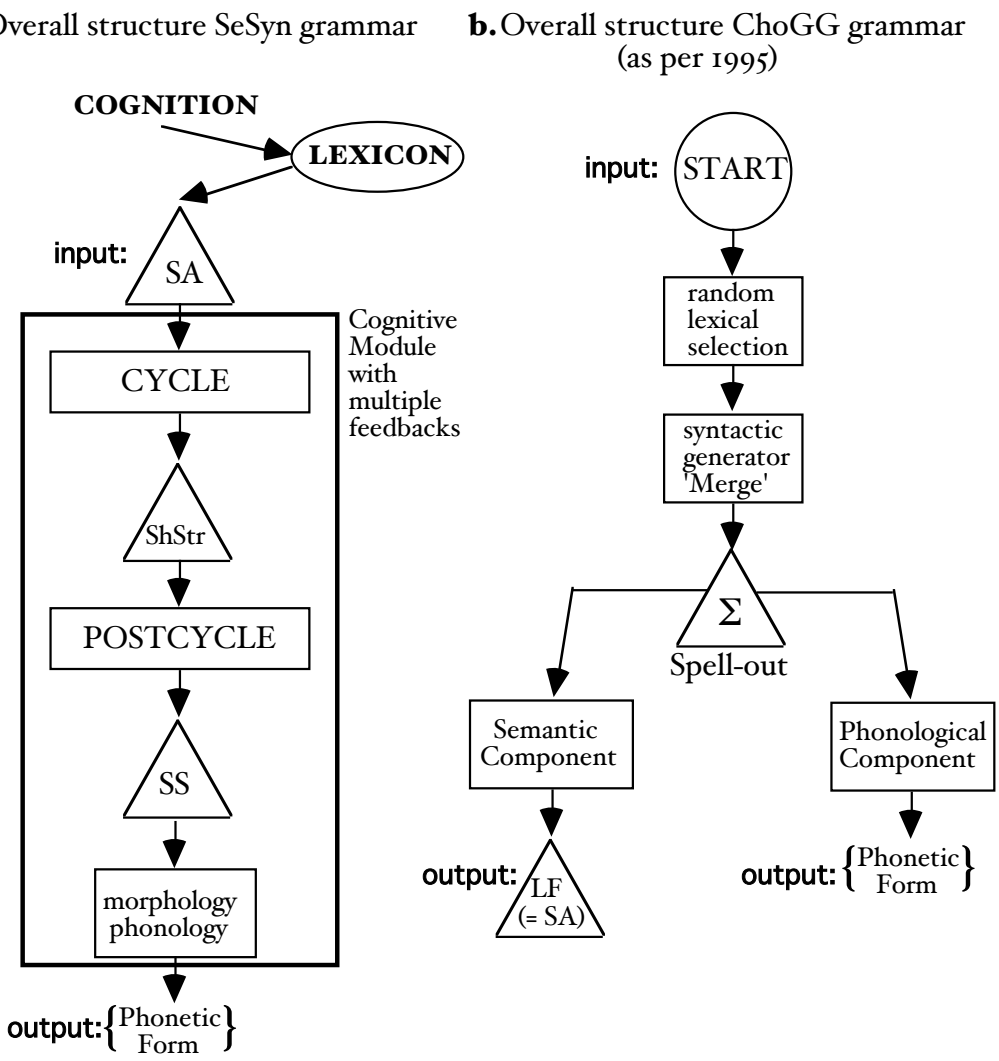

Figure 1 Overall structure of SeSyn and Chomskyan Generative Grammar (ChoGG) (triangles are structures; boxes are procedures)

A SeSyn grammar is *transformational*, as it transforms, via the CYCLE and the POSTCYCLE, SA tree structures into *surface* tree structures (SS), which serve as input to the morphology and phonology of the language L in question (Fig. 1-a). SeSyn thus makes explicit how propositional thoughts are expressed in any given L, having been generated by, or in, cognition and cast into the predicate-logical propositional form of an SA, via

the LEXICON of L, combined with a speech-act operator. SAs are not universal but already geared to the L at hand through the choice of language-specific lexical predicates, which dictate the predicate-argument structure of the sentences of L.[1]

The algorithmic procedure from the SA input to the Phonetic Form output, is taken to be a *module* (Fodor 1983), i.e. a fully automated algorithmic procedure, inaccessible to introspection or interference and open to conscious control only at the input window, though with multiple feedback possibilities allowing for correction before and after final delivery (Seuren 2009a: 234)—not unlike a (very fast) production line in a factory.[2] The *grammar* as such is thus fully automatic and meaning-independent, but its input is semantically determined, and integrated into the whole of the conscious and subconscious human mind.

SeSyn differs basically from a Chomskyan Generative Grammar (ChoGG), which generates sentences randomly, triggered by an arbitrary START symbol, and producing as yet meaningless symbol structures: ChoGG is a *random sentence generator*, whose intermediate output "Σ" is assigned a semantic representation yielding a *logical form* or LF (roughly SA), next to a phonetic interpretation producing a *phonetic form*, as shown in Fig. 1-b. A ChoGG grammar thus has a double output for every random input. Since humans are not random sentence producers, the system cannot represent a cognitive-realist grammar.

SeSyn began, during the mid-1960s, as an offshoot of ChoGG called *Generative Semantics*, a movement started by Jerrold Katz and Paul Postal in their (1964) and further developed mainly by Jim McCawley, Haj Ross and George Lakoff, with myself as a European founding member. The main argument was that a ChoGG-type grammar of the then current kind is nontrivially simplified and made empirically more adequate if it is assumed that the input consists of a well-defined SA-structure (the nature of which was developed by McCawley and myself), while the surface structure output is roughly as assumed in ChoGG. That argument still holds for all later manifestations of ChoGG.

The initiators of Generative-Semantics/SeSyn did not, at the time, realise that their theory actually originated with Wilhelm Wundt (1832–1920), who had proposed that

---

[1] It is probably necessary to postulate, in addition, a PRECYCLE, where cognitively organised logical structures are re-arranged according to lexical choice and language-specific instructions—a process known as "thinking for speaking" (Slobin 1987) or "microplanning" (Levelt 1989: 107–110). Given the still prevailing lack of empirical access and reliable theory formation, the PRECYCLE is left out of account here.

[2] The *comprehension* of sentences is considered not to be a fully algorithmic process (contrary to what is held by most philosophers and computational linguists), but largely based on 'reconstruction-by-hypothesis', codetermined by context, situation and world knowledge and, like the production process, open to multiple feedback control cycles (Seuren 2009a: 268–276). (Interestingly, morphological constructions tend to be much more open to algorithmic analysis, and thus less dependent on context, situation or world knowledge, than syntactic constructions.)

natural-language sentences are derived from an underlying cognitively real propositional structure (Seuren 1998: 219–227). Nor did they realise that the Swiss linguist Albert Sechehaye (1870–1946), junior member of staff in Ferdinand de Saussure's linguistics department in the University of Geneva, had taken up Wundt's challenge (Sechehaye 1908, 1926) and had thus in effect laid the foundations of Generative Semantics and thus also of SeSyn (see Seuren 2018b for extensive discussion).

SeSyn immediately met with excessive resistance from Noam Chomsky, who, however, failed to provide any substantial argument, pointing merely at aberrations and excesses by some overenthusiastic but undisciplined adherents. Regardless of what motivated this hostility, the net result has been that ChoGG has, in actual fact, remained faithful to its *neopositivist* (Carnapian) origin as an algorithm triggered by pushing a start button and with an output in need of a semantic interpretation. Chomsky himself has always been ambiguous about whether ChoGG still follows the mid-20th-century-type neopositivist model of a scientific theory, which excludes, as a matter of principle, any causal role of anything like 'the mind'—the notorious *horror mentis* of 20th-century science—or whether it is to be seen as a *cognitive-realist* theory, in the spirit of the *Cognitive Revolution* that took place at Harvard during the 1950s and to which Chomsky, then also at Harvard, had pledged allegiance (Chomsky 1968/1972). While the new Cognitivism promoted the mind as a prime causal factor in human behaviour, Chomsky, though considered a leading figure of the movement, never admitted that, in fact, it ran counter to the ChoGG grammar model he was in the process of developing and promoting, and whose theoretical foundations he was, apparently, unwilling to revise. What Chomsky attacked was behaviourism, not neopositivism, which makes him only a half-hearted participant in the Cognitive Revolution, despite his prominent position in it. ChoGG is thus still firmly rooted in early 20th-century neopositivism and thus runs counter to the letter and the spirit of the post-1950 Cognitive Revolution, despite claims and appearances to the contrary.

In the wake of the Cognitive Revolution, a massive shift is now under way from neopositivism towards cognitive realism, drawing cognitive neuroscience into the picture. It rests on, and is motivated by, mere principles of good science: neopositivism fails for the cognitive side of the human sciences simply because it does not deliver the goods. A journalist recently asked me the blunt question: why is ChoGG wrong? My answer was simple: ChoGG has remained caught between, on the one hand, the cognitive absurdity of its formalism, blocking a realist interpretation, and, on the other, its perverse formal tortuousness if taken in a nonrealist sense (Seuren 2004: 70). Typologists, who, sanely, study language in its ecological environment, have developed an allergy to theoretical grammar, as that field is entirely dominated by ChoGG, which only provokes shock reactions. This has proved highly deleterious to linguistics as a whole, as it blocks a sound insight into the universals of language, which are to be found in the system rather than in the surface phenomena. Normal scientific method simply requires cognitive realism in the human sciences, and SeSyn fulfills that condition as far as possible.

No matter how one looks at it, when the foundations of a theory are unsafe, then sooner or later the edifice will collapse. This is now visibly happening to ChoGG, which signally fails to link up with adjacent disciplines such as psycholinguistics, neurolinguistics, sociolinguistics, historical linguistics, linguistic typology or any viable form of semantics, all of which are crying out for a serious theory of grammar but are not receiving anything worthy of that name.

Despite the many facts, no matter how important, revealed by ChoGG practitioners in many different languages, no recognition has ever come forward of the equally important facts revealed in terms of the SeSyn framework. In particular, Verb Clustering by means of the rule of PREDICATE RAISING, discussed below, has been systematically ignored by ChoGG practitioners, who themselves have no solution to offer: there is no dialogue between the parties—a very anomalous situation.

SeSyn, being cognitive-realist, fully fits into the pattern set by Cognitive Science. It actively strives for integration with cognitive (neuro-)science, even though the gap between mental structures and processes on the one hand and their neural correlates on the other has not (yet) been bridged. Like its syntax, the semantics of SeSyn is fully cognitive-realist, the notion of *mental proposition* being basic to all cognitive, logical and (con)textual processing. For that reason, SeSyn requires sweeping and highly consequential measures with regard to standard logic, which model-theoretic semantics has been unwilling to consider. Model-theoretic semantics is thus likewise irrelevant to the study of natural language. SeSyn manifests a novel, mentalist, anti-positivist approach to linguistics, cutting through the *horror mentis* that has plagued the human sciences for over a century.

The claim is thus that the cognitive-realist approach of SeSyn is empirically superior to any existing alternative. This assertion cannot be seriously underpinned in a single presentation but it can be demonstrated by a few examples. In the present context, the focus will be on two outstanding features of SeSyn: *Auxiliation* and *Verb Clustering* through the rule of PREDICATE RAISING (PR). Just to show how SeSyn works, the algorithmic derivations of an English and a French sentence are shown from their SA input to their output *Shallow Structure* (ShStr). The interest is that these cyclic processes are found in languages all over the world, allowance being made for language-specific differences.

This makes SeSyn especially relevant for *linguistic typology*, which ranks as the most promising branch of linguistics but has so far lacked a satisfactory basis in the theory of grammar. The problem with linguistic typology is that it is happy with intuitive and impressionistic language descriptions, oblivious, it appears, of the fact that there is an underlying systematically functioning grammatical *mechanism*. As specifications of such a mechanism such descriptions are hopelessly inadequate. In this regard, SeSyn helps out: any typologist familiar with SeSyn will immediately recognise syntactic phenomena in the languages they study as instances of the rules and categories of SeSyn, just as any SeSyn theorist immediately sees what is going on in the syntax of a language unknown to them.

2 Input SA-structure and cyclic derivation of an English sentence

The tree structure of Fig. 2-a is the presumed SA of the English sentence *The cat may have eaten the mouse*. It shows the general principles underlying SA-structure for English and other European languages (with only minor adaptations for wider ranges of languages). The first thing to be noted is that the structure is *throughout propositional*: every S (below the Speech-Act Operator) consists of a predicate followed by one or more arguments, which are either nominal (NP) or themselves propositional again (S). The system is thus infinitely recursive in virtue of its property of allowing unrestricted S-embedding. The propositional nature of SA-structure establishes a direct link with logic.

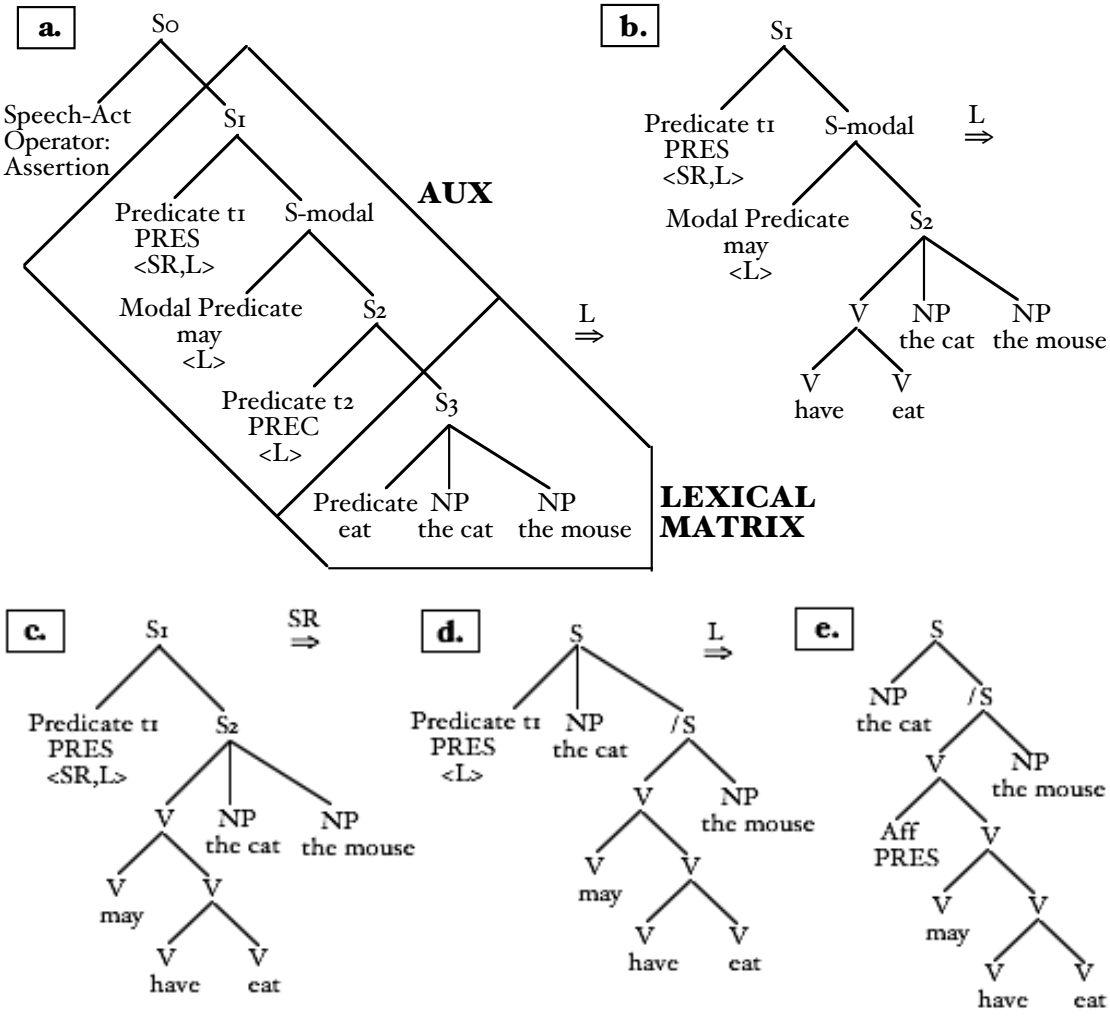

Figure 2 SA and cyclic derivation of the sentence *The cat may have eaten the mouse*

In each S-structure, the predicate is placed first, followed by the lexically defined number of argument terms (NP or S). English SA thus differs from SS, where the dominant order is: Subject-Verb-Object (SVO), that is, NP-VP. The transition to surface

SVO-order is effected by the CYCLE.[3] A predicate has minimally one and maximally three argument terms: subject, indirect object, direct object—in that order. If one, it is the subject; if two, the first is subject and the second is direct object. If three, the middle term is indirect object. Subject and direct object terms may consist of an embedded S. The embedding of Ss under a matrix predicate is known as COMPLEMENTATION.

Verb-initial, or VSO, languages are produced by leaving out one single rule feature, as will be shown in a moment. For Verb-final languages, such as Japanese or Turkish, the underlying order is, *ceteris paribus*, taken to be SOV. For the time being, two possible SA constituent orders are assumed: VSO and SOV. How to handle other possible orderings (including Ergative) in the languages of the world is a matter of current research.

The SA-structure in Fig. 2-a is divided into three parts: the Speech-Act operator, AUX and the LEXICAL MATRIX. The first is left undiscussed, given the format of this presentation. The LEXICAL MATRIX contains the main lexical predicate, which is, normally speaking, a Verb, Adjective or NP in SS, plus the argument terms. The subject and object terms may be an embedded complement-S. The *Auxiliary Area* (AUX), contains, in principle, all elements that somehow qualify the proposition expressed in the LEXICAL MATRIX, such as tense, modality, aspect. Overall, the MATRIX tends to incorporate elements of both AUX and embedded complement-S-structures—a tendency I call *matrix greed*. AUX-elements are usually, but not always, united with the MATRIX V-constituent, thus forming a complex V-cluster, as is demonstrated in the derivations presented below.

In English, as in most other European languages, AUX minimally contains two tenses, t1 and t2, in that order. Both allow for a double choice, the former between deictically defined 'Present' (PRES) and anaphorically defined 'Past' (PAST), the latter between 'Simultanuous' (SIM) and 'Preceding' (PREC) (both SIM and PREC with regard to the time segment referred to by means of t1). For English, the following correspondences hold:

PRES + SIM → Simple Present (eat)	PRES + PREC → Present Perfect (have eaten)
PAST + SIM → Simple Past (ate)	PAST + PREC → Pluperfect (had eaten)

In most languages, the precise semantics of the tenses is complex, idiosyncratic and strongly language-specific. The Futuricity operator is a modal predicate, both grammatically and semantically, not a tense, also in languages that express the future morphologically in the Verb form (for SeSyn this is just a matter of surface category specification in the LEXICON).

Modal predicates (*will*, *may*, *can* etc.) are optionally placed in SA between t1 and t2 and are thus part of AUX. This explains their 'defective paradigm': they lack infinitival and participial forms and occur only in the Simple Present and Simple Past, never in the

[3] This is based on McCawley's (1970) discovery that the cyclic machinery transforming an SA into a ShStr is nontrivially simplified if VSO-order is assumed for the SAs of English and many other languages—a hypothesis that has withstood the wear of time.

perfective tenses, which require the perfective auxiliary *have* plus the Past Participle, but they can be followed by either a SIM or a PREC infinitive, as in *may eat* and *may have eaten*.

Historically, most AUX-predicates, notably modal and aspectual predicates, but also Tense predicates (as one learns, for example, from the history of Creole languages), started life as full MATRIX predicates embedding what is now the lexical MATRIX as a complement-S, but getting 'bleached' over time, which made them become part of AUX—a universally found tendency known as *auxiliation* (Kuteva 2004). Apparently, the Dutch Futuricity verb *zullen* ('shall/will') had not yet auxiliated when I acquired Dutch some eighty odd years ago, but it seems to have auxiliated in the meantime. For me and my contemporaries, a sentence like *Hij had zullen vertrekken* ('*He had willed leave'), where *zullen* occurs in the Pluperfect, is fully grammatical. But nowadays, I find to my surprise, young speakers no longer accept that sentence, which means that *zullen* has auxiliated in the meantime. Its German equivalent *werden* auxiliated much earlier, as is shown by the fact that the equivalent German sentence **Er hätte abreisen werden* is ungrammatical for all speakers. But other, similar, verbs have not, or not yet: *Er hätte abreisen sollen* (he should have left) is impeccable, like its Dutch equivalent *Hij had moeten vertrekken*. In Old-English, the modals still had participles and infinitives, but these disappeared in Middle-English (Fischer et al. 2004), clearly because they fell victim to auxiliation.

A sharp distinction is made between the status of lexical items in the lexicon on the one hand and their surface lexical category on the other. In the lexicon, all items are *predicates*: verbs, nouns, adjectives, adverbs, conjunctions, prepositions, quantifiers, negation, are all listed as 'PREDICATE'. The lexicon specifies for each predicate—apart from its semantic specification—its surface category, its argument structure, its phonological representation(s), and the cyclic rules induced by it (its *rule features*). This accounts for the sometimes surprising differences in word class of otherwise semantically equivalent lexical elements. In Finnish, for example, the negation is realised as a modal predicate followed by an infinitive, with a defective paradigm identical to that of the English modals.

The CYCLE carries the SA to the *shallow structure* (ShStr), which is is fed into the POSTCYCLE. The latter is only touched upon here as it is highly language-specific, while the CYCLE is surprisingly uniform across the languages of the world (ShStr is typically a halfway product, already recognisable as a yet unborn surface structure). Which rules are assigned to the CYCLE and which to the POSTCYCLE is a matter of experience and judicious choice.

So now the question is: what is the CYCLE?[4] The CYCLE is the first processing box the SA is fed into. It contains a handful of universal rules either deleting (semantically recoverable) elements in a lower S, or lowering the PREDICATE of a given S into an

[4] Other than claimed by Freidin (2007: 1), who mentions (Chomsky 1965) as the originator, the CYCLE was introduced in Fillmore (1963). In Chomsky (1965) neither the term nor the notion occurs. The CYCLE has, moreover, never figured prominently in ChoGG.

embedded S, or raising an element of an embedded S into itself. Cyclic rules are induced by the predicate of any given S in SA. The CYCLE starts at the most deeply embedded S whose predicate induces a cyclic rule.[5] The LEXICON specifies for each predicate which cyclic rules it induces (obligatorily or optionally) and which surface category is assigned to it as the CYCLE passes through it. The cyclic rules induced by a predicate as specified in the LEXICON are shown in the trees as rule features between angled brackets below the predicate at issue. When the predicate induces no cyclic rule, we say that the CYCLE *passes vacuously* through the S in question. When the CYCLE has done its work in a given S, it passes on to the immediately superior S, recursively, until it reaches the highest S, at which point the CYCLE comes to an end and delivers the *shallow structure* (ShStr) of the sentence being processed. This is then fed into the POSTCYCLE.

So let us now see how the CYCLE operates for the sentence *The cat may have eaten the mouse* (Figure 2). We start with the SA-structure (Fig. 2-a). The predicate of S3, the lowest S in Fig. 2-a, induces no cyclic rule (carries no rule feature), so that the CYCLE passes vacuously through S3, but the label "Predicate" over *eat* is replaced with the surface category label "V" (Verb). Then, at the S2-Cycle, the rule L (LOWERING) applies. L occurs in a number of varieties. In this case, L left-attaches the predicate in question to the V(erb) of the lower S, forming the cluster $_{V}$[$_{V}$[have] $_{V}$[eat]], as shown in Fig. 2-b (*have*, the surface form for PREC, is specified in the LEXICON as a surface V). The original S2-node is eliminated given the general principle that an S dominating only another S is 'pruned'. At the next cycle up, the S-modal cycle, the Predicate *may*, likewise a surface V, is lowered in exactly the same way as *have*, resulting in Fig. 2-c. At the S1-Cycle, two rules apply, SR (SUBJECT RAISING) and L, in that order. SR takes the subject-NP of the embedded S and places it in the position of that embedded S, which is moved one position to the right. In virtue of a general DEMOTING PRINCIPLE, the S of the now subjectless S is relabelled "/S" (traditionally known as Verb Phrase or VP), meaning that it needs an NP to be a full S. This gives Fig. 2-d. (Without the rule feature SR for t1, stage Fig. 2-d is cut out so that the language keeps VSO as surface order.) All that remains is the application of L at the S1-Cycle, yielding Fig. 2-e, the ShStr of the SA-input (PRES is labelled for surface category 'Affix'). This last structure is now left to the mercy of the POSTCYCLE, the MORPHOLOGY and the PHONOLOGY of the grammar of English, each successive stage being more language-specific and closer to the corresponding SS than the previous one.

---

[5] Cyclicity is a general property of recursive calculus. In arithmetic, for example, to compute the value of a formula like (12 − 2) + (9 : 3), one must first compute the embedded (12 − 2) and (9 : 3), before the highest operator (+) can do its work, with 13 as the final value. The final 'value' of an SA at the end of the Cycle is its Shallow Structure, which is fed into the Postcycle.

3 Predicate Raising (PR) in French and related matters

The cyclic rule PREDICATE RAISING (PR), induced by the French causative Matrix predicate *faire*, as in Fig. 3-a, has a curious history, starting with McCawley (1968, 1970b, 1971), who proposed PR as a so-called 'prelexical rule', operative within the LEXICON. His main example, still widely quoted among linguists, was the verb *kill* as a lexicalisation of "cause-to-die" by means of the PR rule. Chomsky criticised this (Chomsky 1972: 142–143):

> This is the approach taken by McCawley in the case of words such as *kill* = "cause to die". In the proposed underlying structure, *John caused Bill to die* [...], the unit that is replaced by *kill* is not a constituent, but it becomes one by the *otherwise quite unnecessary rule of predicate raising* [emphasis mine; PAMS]. Such a device will always be available, so that the hypothesis that Q is a constituent has little empirical content.

But Chomsky, who rarely uses examples from other languages than English, was unaware of the fact (as was McCawley) that English (like Latin or Portuguese) belongs to a minority of languages that lack PR, using SR instead. PR is overwhelmingly and demonstrably present in the languages of the world, as signalled in (Seuren 1972), which takes the French *faire*-construction as a prototypical case and analyses it in great detail, besides further cases from other languages such as Dutch or German, where the PR-rule is rampant throughout the complementation system (Evers 1973, 1975). Although the arguments put forward in (Seuren 1972) were compelling, and although the paper in question was brought to the attention of all leading ChoGG linguists at the time, no public response ever came forth, as if a fatwa had been slammed on it, prohibiting any mention. One reason, one gathers, was the lack of valid counterargument or alternative analysis (typically, the empirically superior Katz & Postal 1964 is never mentioned in the ChoGG literature). In ChoGG circles, PR phenomena have, until the present day, steadfastly been ignored or treated in an offhand way. PR has thus become an irritant, rather than a stimulant for advancing the theory.

But what does PR amount to? The answer is simple:

> PR takes the V(-cluster) of the embedded S and unites it with the inducing higher predicate (relabelled "V"), either on the right or the left hand side, depending on the language (German takes left, Dutch, French, Italian take right attachment). The S-node is deleted and all remaining material is re-attached higher up, in the order given.

In Fig. 3-a, the CYCLE passes vacuously through S4, then applies PR on the S3-cycle, resulting in Fig. 3-b, which now has the V-cluster $_{V}$[$_{V}$[faire]$_{V}$[voir]] while the remaining material of S4, $_{NP}$[Didier] and $_{NP}$[la lettre], has been re-attached to the next higher S3 in the order given, turning S3 into an S with three NPs in a row: a subject, an indirect object and a direct object. This explains how the original subject of S4 becomes an indirect object (dative) in the surface structure of the sentence. (Contrary to the way the French dative is treated in Seuren 2018a, the obligatory change from internal to external dative (*à Didier*) is probably best treated as the result of a postcyclic rule.) If S4 only has a subject term, this will turn up as a direct object in SS, as in *Anne fera partir Didier* (Ann will make Didier leave), simply because the predicate-raising S will now have only two NP-terms.

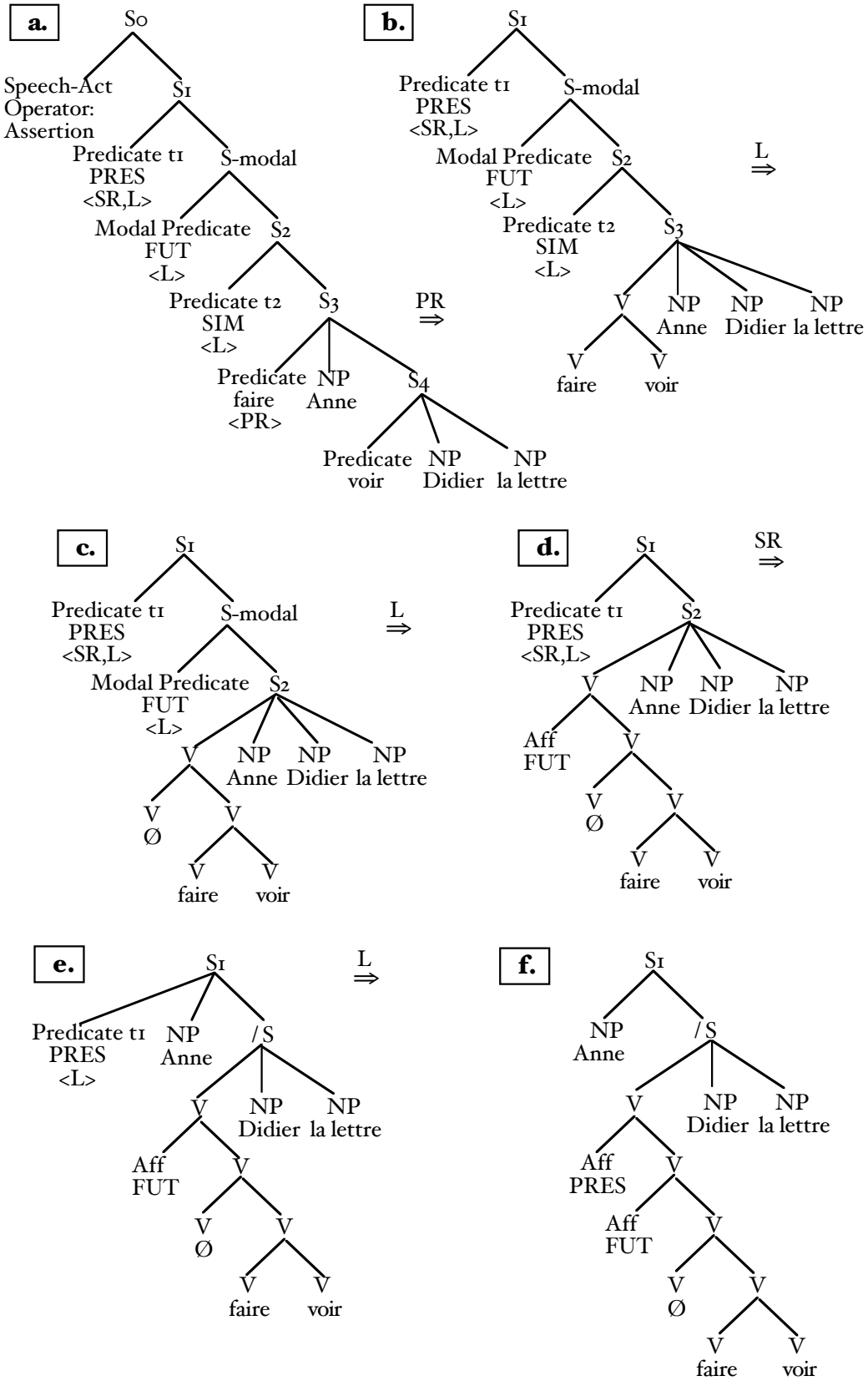

Figure 3 Cyclic derivation of the French sentence *Anne fera voir la lettre à Didier* ('Ann will make Didier see the letter')

Corresponding phenomena occur in many other languages even though they differ considerably in other aspects of their grammar, especially when the causative predicate is relabelled 'Affix'. The phenomenon known as 'valency increase' is, in most cases, simply due to PR with 'cause' relabelled as 'Affix': intransitive or transitive verb stems, augmented with a causative prefix, suffix or infix, or morphophonemically modified, acquire an extra argument in the dative case, as in the following Turkish example (Lewis 1975: 146–147):

(1) a. Mektub-u imzala-d-ım
letter-ACCUSATIVE sign-PAST-1sg
(I signed the letter)

b. Müdür-e mektub-u imzala-t-t-ım
director-DATIVE letter-ACCUSATIVE sign-CAUSATIVE-PAST-1sg
(I made the director sign the letter)

The dative *müdüre* signals that PR has been at play. If the embedded clause had been intransitive, as in *I made the director leave*, the word *müdür* (director) would be in the accusative case (*müdürü*), just as in French, Italian and the many other languages that show evidence of PR in their complementation system and have case marking. In general, the occurrence of a dative representing the semantic subject of a transitive clause embedded under a causative verb is a sure sign that the language in question has PR, rather than SR, in its complementation system. This also applies to many lexical predicates, taken by McCawley and others to incorporate a bit of syntactic structure, as in the lexical predicate *show*, prelexically analysed as 'cause-to-see': 'Ann caused-to-see Didier the letter' then becomes *Ann showed Didier the letter*. Apparently, English does have PR in its prelexical syntax, though it lacks PR in its open syntax, using SR instead, as already implied by McCawley.[6] (Ironically, ChoGG adherents have now started to take to prelexical analyses as well.)

It is not so, however, that PR always induces a dative when the embedded complement-S is transitive. In German, for example, accusative case is assigned, as in (2a) (probably as a result of 16th-century Latin influence in German schools). Overuse of PR may lead to a succession of more than three NP-terms in a row, separated from the corresponding series of PR-raised verbs, as in the Dutch subordinate clause (2b), where the accumulation of four NPs in a row strains the short-term storage capacity needed for processing the clause, even though it is produced according to the rules (luckily, Dutch makes PR optional in enough cases for the grammar to allow for more easily processable

[6] Typically, lexical items (including compounds) that are open to prelexical analysis undergo semantic specialisation, as is normal for lexical items. For example, *murder* is, like *kill*, analysable as 'cause-to-die', but it has specialised for 'cause a human to die unlawfully and with malice aforethought', while *assassinate* has further specialised for 'cause a person of public importance to die unlawfully and with malice aforethought'. Such semantic specialisations have been used as an argument against prelexical analysis, but it will be clear that this type of argument lacks any force.

alternatives, such as (2c), which has three consecutive NPs) (for a detailed analysis and description, see Seuren 2018a/1996, Chapters 5 and 6):

(2) a. ... daß Johann **das** Kind den Ball zu werfen gelehrt hat.
... that John **the**$_{ACC}$ child the$_{ACC}$ ball to throw taught has
(... that John has taught the child to throw the ball)

b. ... dat Jan het kind de hond de bal wilde laten leren halen.
... that John the child the dog the ball wanted let teach fetch
... (that John wanted to let the child teach the dog to fetch the ball)

c. ... dat Jan het kind de hond wilde laten leren de bal te halen.
... that John the child the dog wanted let teach the ball to fetch
(... that John wanted to let the child teach the dog to fetch the ball)

A further powerful argument supporting the PR-analysis is provided by the behaviour of *anaphorical clitics* in French and some other Romance languages. In these languages, unaccented pronouns and a few anaphorical adverbs occur as clitics attached to the Verb form in a strictly defined order (Seuren 2009b gives a detailed analysis of clitics in French and Italian). What is relevant here is that the placement of clitic clusters, at least in French and Italian, is sensitive to the /S (or VP) constituent they are part of. Consider the following two sentences:

(3) a. Anne la lui fera voir. (*Anne le fera la voir.)
(Ann will make him see it)

b. Anne veut la lui donner. (*Anne la lui veut donner.)
(Ann wants to give it to him)

Fig. 4-a shows what sentence (3a)—with *Didier* and *la lettre* replaced with the clitics *lui* and *la*, respectively—looks like after the postcyclic rules Ø-DELETION and AFFIX-HANDLING have applied to the ShStr shown in Fig. 3-f. Ø-DELETION takes out the Ø branch of the V-cluster of Fig. 3-f; AFFIX-HANDLING subsequently takes all affixes of the V-cluster, starting with the lowest one, and re-attaches them to the right of $_{V}$[faire], the new cluster being relabelled 'Finite Verb' (FV). Now assume a postcyclic rule *Clitic Movement* (CliticMov), left-attaching clitic argument terms within any given /S-constituent to the FV-constituent or, failing an FV, the V-constituent, of the same /S (the order in which this is done is of no concern here). Since there is only one /S in Fig. 4-a, the pronominal clitics *lui* (to him) and *la* (it) are left-attached to the FV-cluster, as shown in Fig. 4-b. This explains the position of *la* + *lui* in (3a) and the ungrammaticality of **Anne le fera la voir*.

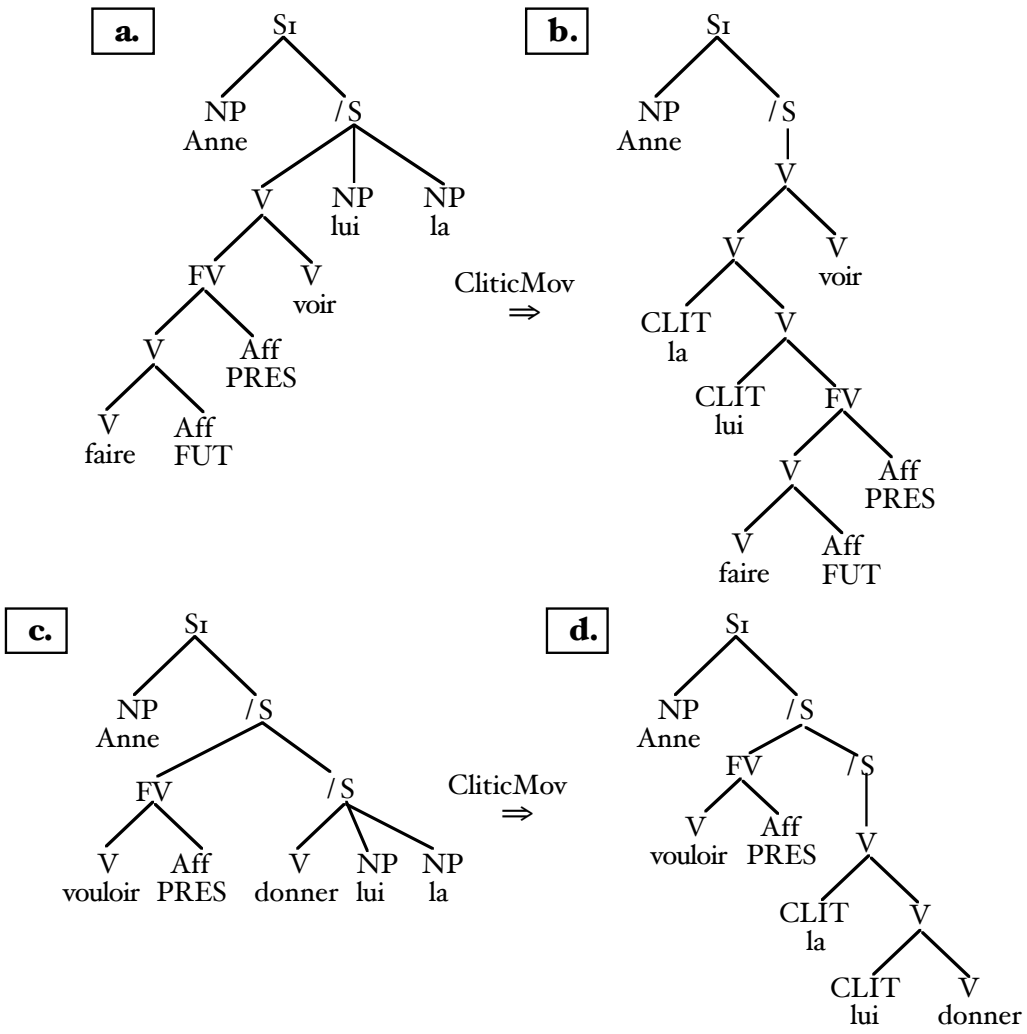

Figure 4 French Clitic Movement under PR and under SD

Now consider (3b), presented in Fig. 4-c at the same postcyclic stage as Fig. 4-a. The difference with Fig. 4-a is explained by simply listing the French predicate *vouloir* (want) in the LEXICON as inducing the cyclic rule of SUBJECT DELETION (SD) (once known by the awkward name of EQUI-NP-DELETION). SD deletes the subject-NP of the complement-S under the condition of referential binding by the subject of the higher S (identical reference of the two subject-NPs). Since the complement-S has now been stripped of its subject-NP, it is demoted to /S status, as in Fig. 4-c, where the lower /S lacks a subject. The structure now contains two /S-constituents, one /S as part of the main S and one /S as a constituent of the latter. The postcyclic rule CliticMov, now left-attaches the clitics *la* + *lui* to the V-constituent of the same /S, resulting in Fig. 4-d.

Italian behaves in much the same way as French does, except that the verb *volere* ('want')—semantically and historically equivalent to French *vouloir*—is listed in the LEXICON as being free to choose between SD and PR—as was likewise the case in early modern French. Consequently, Italian allows for both (4a) and (4b), both meaning 'Ann wants to give it to him', while the French sentence marked as ungrammatical in (3b) was grammatical in early modern French, exactly as its Italian counterpart:

(4) a. Anna gliela (= gli + la) vuole dare.
b. Anna vuole dargliela.[7]

Early on, adherents of ChoGG have tried to account for the difference between (4a) and (4b) by assuming a rule of 'Clitic Raising', obligatory for French *faire* and Italian *fare*, forbidden for modern French *vouloir* but optional for Italian *volere*. This rule, however, was invented only to accommodate the observed facts and did not help explain other facts or fit into any general pattern or system. Devoid as it was of any explanatory power, it has not or hardly been heard of since the 1970s, leaving the facts unexplained.

It appears that PR is one of the main complementation strategies in the languages of the world, in competition with SR, while SD is combinable with both, as shown in Figure 5. Many linguists have confused PR with SR. Had they engaged in more systematic observation, they would have seen, for example, that, without PR, there is no well motivated way to explain why in the English sentence *Ann will make Didier leave* $_{NP}$[Didier] comes between $_{V}$[make] and $_{V}$[leave}, while in the French sentence *Ann fera partir Didier* $_{NP}$[Didier] must occur after the V-cluster $_{V}$[$_{V}$[faire]$_{V}$[partir]].

The English predicate *want* induces SD in cases where the higher, commanding, subject referentially binds the lower subject of the complement S, leading to *John wants to eat the cake*, but it induces SR when there is no referential binding, giving *John wants Lucy to eat the cake*. PR does not occur in English open syntax (other than in the semilexicalised *let go*, which can even be passivised, as in *She has been let go by her employer*).

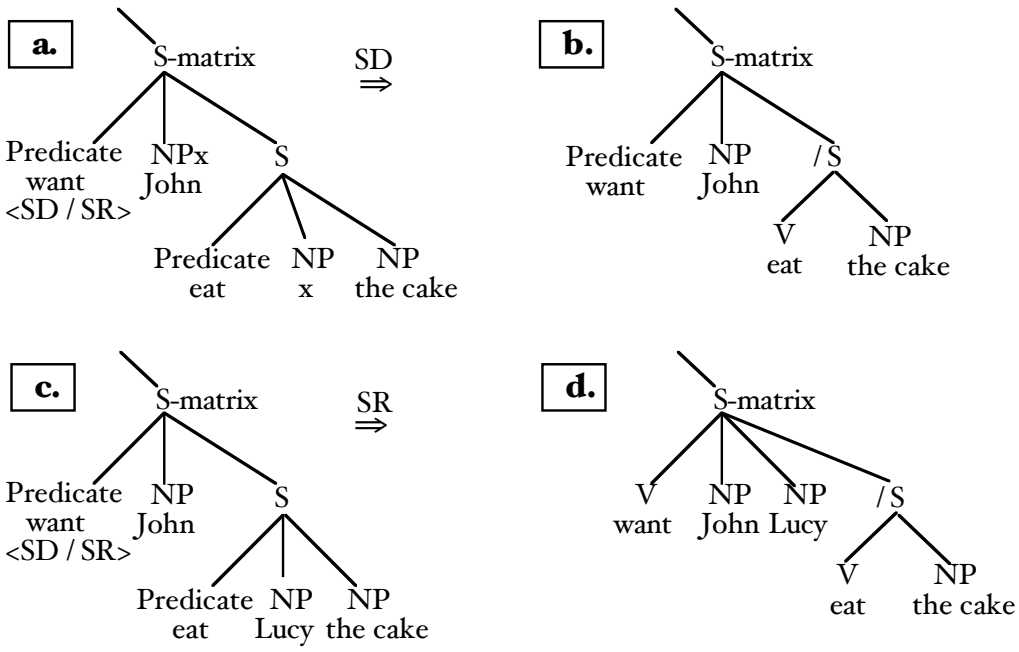

Figure 5 The interplay between SD and SR for English *want*

[7] In Italian, clitics in construction with infinitives and participles are right-attached as affixes, not as clitics. The form *dargliela* is thus composed of *dare* (give) + *gli* (to him) + *la* (it).

French has PR for complement object-Ss under *faire* and a small number of other predicates (often with idiosyncratic restrictions), whereas complement subject-Ss undergo SR, as in (5a). But SR never occurs in French complement object-Ss. Italian is like French, in this respect, except that complement subject-Ss are left the choice between PR and SR, as in (5b,c). French *sembler* and Italian *sembrare*, like their English equivalent *seem*, are listed in the LEXICON as taking a complement S as subject term. SR yields English *Conrad seems to have found it*, French *Conrad semble l'avoir trouvé* (= (5a)) and Italian *Corrado sembra averlo trovato* (= (5b).

By contrast, Italian *Corrado lo sembra aver trovato* (= (5c)) is the result of PR induced by *sembrare* as a permitted alternative to SR (authoritative native speakers of Italian have assured me that any difference between the two is merely stylistic):

(5) a. Conrad semble l'avoir trouvé. (*Conrad le semble avoir trouvé.)
Conrad seems it_{ACC} have found
(Conrad seems to have found it)

b. Corrado sembra averlo trovato.
Conrad seems have- it_{ACC} found
(Conrad seems to have found it)

c. Corrado lo sembra aver trovato.
Conrad it_{ACC} seems have found
(Conrad seems to have found it)

Portuguese is like English in that it has no PR but only SR, again in combination or alternaton with SD. Dutch and German only have PR, assigned to almost all complement-taking predicates, sometimes combined with SD but never as an alternative to it. SR does not occur in the complementation system of Dutch and German—a fact still unacknowledged in most of the literature, owing to deficient observation and lack of sound theory. Typological studies suggest that PR is more frequent than SR in the complementation systems of the languages of the world, frequently penetrating into the morphology of the verb forms and thus giving rise to what is widely known as 'valency increase'.

This concludes our brief tour of SeSyn. It has provided a mere glimpse of what the full system already amounts to, and of what it will amount to when it is applied to more categories and more languages. Small as it is, however, this glimpse should whet the appetite of those linguists who are on the lookout for a theory of grammar that is in step with adjacent disciplines such as cognitive science, psycholinguistics, language typology or sociolinguistics, and at the same time provides them with a more solid theoretical background . The arguments and analyses provided show the explanatory power and the solid theoretical foundations of the SeSyn system as a whole. Taken together, SeSyn constitutes a massive challenge to any possible alternative theory of grammar. As long as that challenge is not met in a level playing field, the claim can be upheld that, at least in principle, the SeSyn approach to syntax reflects, approximately at least, the mental, and perhaps even the neural, mechanism underlying the production of sentences—even though in practice, as in any factory production line, shortcuts may be taken, ready-made

spare parts may be inserted, and parallel processing may be practiced so as to speed up or simplify the process. The human mind strikes one as extremely efficient in this regard.

References:

Chomsky, Noam (1965). *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

———(1972). *Language and Mind. Enlarged Edition*. New York: Harcourt Brace Jovanovich. (First edition 1968)

——— (1995). *The Minimalist Program*. Cambridge, MA: MIT Press.

Evers, Arnold (1973), The syntactic motivation of Predicate Raising. *Spektator* 3: 69–94.

——— (1975). *The Transformational Cycle in Dutch and German*. PhD-Diss. Utrecht University.

Fillmore, Charles J. (1963). The position of embedding transformations in a grammar. *Word* 19: 208–231

Fischer, Olga, Ans van Kemenade, Willem Koopman and Wim van der Wurff (2004). *The Syntax of Early English*. Cambridge: Cambridge University Press.

Fodor, Jerry A. (1983). *The Modularity of Mind*. Cambridge, MA: MIT Press.

Freidin, Robert (2007). *Generative Grammar: Theory and its History*. London/New York: Routledge.

Katz, Jerrold J. and Paul M. Postal (1964). *An Integrated Theory of Linguistic Descriptions*. Cambridge, MA: MIT Press.

Kuteva, Tania (2004). *Auxiliation. An Enquiry into the Nature of Grammaticalization*. Oxford: Oxford University Press.

Levelt, Willem J. M. (1989). *Speaking. From Intention to Articulation*. Cambridge, MA: MIT Press.

Lewis, Geoffrey L. (1975). *Turkish Grammar*. Oxford: Clarendon Press.

McCawley, James D. (1968). Lexial insertion in a transformationl grammar without deep structure. In: *Papers from the Fourth Regional Meeting, Chicago Linguistic Society*, Chicago: Linguistics Department of the Universty of Chicago: 71–80. (Also in McCawley 1973: 155–166)

——— (1970a). English as a VSO-language. *Language* 46.2: 286–299. (Also in McCawley 1973: 211–228)

——— (1970b). Semantic representation. In: Paul M. Garvin (ed.), *Cognition: A Multiple View*, New York: Spartan Books: pp. 227–247. (Also in McCawley 1973: 240–256)

——— (1971). Prelexical syntax. *Georgetown University Monograph Series on Language and Linguistics* 24: 19–33. (Also in McCawley 1973: 343–356)

——— (1973). *Grammar and Meaning. Papers on Syntactic and Semantic Topics*. Tokyo: Taishukan.

Sechehaye, Ch. Albert (1908). *Programme et méthodes de la linguistique théorique: psychologie du langage.* Paris: Champion / Leipzig: Harrassowitz / Genève: Eggiman. (Facsimile reprint 2011, Nabu Press, Charleston, South Carolina)

——— (1926). *Essai sur la structure logique de la phrase*. Paris: Champion. (Reprint 1950)

Seuren, Pieter A.M. (1972). Predicate Raising and Dative in French and sundry languages. Circulated by LAUT (Linguistic Agency University Trier). Published in Seuren (2001: 139–184).

——— (1998). *Western Linguistics. An Historical Introduction*. Oxford: Blackwell.

——— (2001). *A View of Language*, Oxford: Oxford University Press.

——— (2004). *Chomsky's Minimalism*. New York: Oxford University Press.

——— (2009a). *Language in Cognition*. (= Language from Within, Vol 1). Oxford: Oxford University Press.

——— (2009b). The clitics mechanism in French and Italian. *Probus* 21.1: 83–142.

——— (2010). *The Logic of Language*. (= Language from Within, Vol 2). Oxford: Oxford University Press.

——— (2018a). *Semantic Syntax*. 2nd edition. Leiden: Brill. (1st edition: 1996, Oxford: Blackwell)

——— (2018b). *Saussure and Sechehaye: Myth and Genius. A Study in the History of Linguistics and the Foundations of Language*. Leiden: Brill.

Slobin, Dan (1987). Thinking for speaking. In: Jon Aske, Natasha Beery, Laura A. Michaelis and Hana Filip (eds), *Berkeley Linguistics Society. Proceedings of the 13th Annual Meeting. Febr. 14–16, 1987.* Berkeley CA: Berkeley Linguistics Society: 435–445.