

The Subject Chain in a Compositional Cognitive Grammar

Hsin-I Hsieh

University of Hawai'i

Subject-verb agreement in Mandarin is an Interactive phenomenon, involving semantics and syntax as two Interactive grammatical modules. The semantic Subject Chain, coding the primary participant in an event, maps onto the syntactic First Argument, functioning as the Subject of the sentence expressing this event. The subject chain refers to the primary participant itself, and not to how it relates to the event actions normally coded as verbs or co-verbs, making corresponding subject-verb relation non-crucial. This explains why Mandarin subject-verb relation shows no agreement and requires no morphological markings for agreement. We discuss the subject chain in the framework of a Compositional Cognitive Grammar (CCG), which in turn is a particular realization of the general concept of an Emergent Grammar (EG).

Key words: subject chain, Compositional Cognitive Grammar, Mandarin

1. Introduction

A grammar is a multi-modular system, depending on the mechanism of Interface for its static organization, and on the mechanism of Interaction for its dynamic operation.¹ In organization, grammatical representations of varying complexity and patterns group into modules and submodules of the grammar. A grammatical representation Emerges from the interface of two Compatible or Congruous representations of lesser complexity. The emergent representations in a module (or submodule) are related by a set of rules that (together with these emergent representations) constitute the structure of the module. As these emergent representations slowly and quietly become stabilized and established in the grammar, their once-emergent nature is obscured and they appear as conventional constructions.

In operation, two grammatical rules (or sets of rules) may Interact within a single module or across two modules. Whether in an intramodular or intermodular Interaction, the two Interactive rules may be (critically) sequentially ordered or simultaneously ordered. If sequentially ordered, one rule may (i) *feed* the other by applying before it to

¹ I wish to thank two anonymous reviewers for their very valuable and helpful suggestions for revision.

create an additional input for it (e.g. $a \rightarrow b, b \rightarrow c$), or may (ii) *bleed* the other by applying before it to take away an input otherwise available to it (e.g. $a \rightarrow b, \bar{a} \rightarrow c$) (Kiparsky 1978). If simultaneously ordered, the two rules may (iii) *complement* each other by applying to two separate inputs to yield separate outputs (e.g. $a \rightarrow c, d \rightarrow f$), or may (iv) *conflict* (or *contradict*) each other by applying to the same input to yield two conflicting (or contradicting) outputs (e.g. $a \rightarrow b, a \rightarrow c$). Finally, two rules may (v) *conspire* if they apply jointly to create a single output from either identical or distinct inputs. Conspiracy occurs in three typical situations. If the two conspiring rules apply sequentially, then one of them feeds the other, just as in interaction mode (ii). If they apply simultaneously, one non-vacuous or effective and the other vacuous or non-effective, then they complement each other, just as in mode (iii). If the input to one rule intersects the input to the other rule, then the crucial intersection leads to an identical output for the two rules (e.g. $ac \rightarrow e, bc \rightarrow e$).

Agreement in general, and subject-verb agreement in particular, is an interactional phenomenon between semantics and syntax as two interactive modules, with one feeding to the other in terms both of representations and rules. In Mandarin and other Chinese dialects, the semantic part of the subject-verb agreement is the Subject Chain, and the syntactic part is the First-Argument Sequence. The subject chain codes the primary participant in an event, and transforms into the First Argument of the first-argument sequence in a sentence, where it acquires the status of the Subject. Since the subject is ultimately rooted in the primary participant itself, and not in how it relates to the event actions typically coded as verbs or co-verbs, the subject-verb agreement that would code this relation is non-crucial. This explains why subject-verb agreement in Mandarin employs no overt morpho-syntactic markings, although the subject clearly has a semantic root.

The central concern of this paper has the subject chain as its foreground issue, and has an Emergent Grammar (EG) as its background theory. The background theoretic framework must be sketched in sufficient detail to allow us to fully appreciate the foreground problem and its solution. Fundamental in EG are the key concepts of Interface, Interaction, and Emergence, which must be all made precise first. Furthermore, these key concepts all depend for their elucidation on a Compositional Cognitive Grammar (CCG), serving as a specific realization of an EG. The above dependency relations lead us to the following organization of this paper. Section 2 is devoted to emergent grammar. Section 2.1 discusses interface and interaction, Section 2.2 Emergent state, Section 2.3 Interaction modes. Section 3 concentrates on CCG, sketching it and using it to illustrate an emergent grammar. Section 3.1 discusses the interactive modules, Section 3.2 Semantic structure in CCG, Section 3.2.1 ‘form’ semantic structure, Section 3.2.2 ‘meaning’ semantic structure, and Section 3.3

Emergent constructions. Section 4 is the culminating section on subject chain, followed by a short conclusion as Section 5.

2. Emergent grammar

2.1 Interface and interaction

We make a technical distinction between interface and interaction which elsewhere are usually treated as interchangeable synonyms. This distinction depends on the concept of compatibility between two grammatical units, whether they be two representations or two rules. Two grammatical units are Compatible if (i) they have an *identical* category or pattern, or if (ii) they have two *mutually complementary* categories or patterns. For simplicity, we call the first type of compatibility an Identity Compatibility, and the second type a Reciprocity Compatibility.

Let us explain these two types of compatibility with two artificial examples. Consider the system of usual addition of non-negative integers. Since no compatibility condition is imposed, these integers can be added together, regardless of their status of being even or odd. Thus, the equation (i) $2 + 4 = 6$ is ‘well-formed’ or ‘grammatical’, so is the equation (ii) $2 + 5 = 7$. But now suppose that we adopt an identity compatibility condition, which requires the two integers in the addition to be either both even or both odd. Then equation (i) remains well-formed but equation (ii) becomes ‘non-wellformed’ or ‘ungrammatical’. On the other hand, suppose that we instead adopt a reciprocity compatibility condition, which requires the two integers in the addition to be one even but one odd number. Then equation (i) becomes non-wellformed but equation (ii) remains well-formed.

If two grammatical units are not compatible, then they are Incompatible. Only two compatible grammatical units can be in a state of interface or interaction. Interface and interaction can now be made technically distinct on the basis of compatibility. Two grammatical representations (or sets of representations) are in an Interface if they are compatible and compose into a more complex representation (or set of representations). By contrast, two grammatical rules (or sets of rules) are in an Interaction if they are compatible and compose into a more complex set of rules. If two grammatical representations are in an interface, the rules governing them can often be expected to be in an interaction. Interface and interaction are thus often two sides of the same coin, one static and one dynamic. Since a module has both a set of representations and a set of rules, two modules can be compatible and can on that account interface or interact.

Compatibility, whether based on identity or reciprocity, has a deep, mathematical foundation. Let us explicate by invoking the mathematical concepts of a function and a homomorphism. A function f from set A into set B is a many-to-one (possibly

one-to-one) relation from set A to (into) set B, which assigns to each element a in set A an element b in set B. For example, if set $A = \{1, 3, 5, 7, \dots\}$ and set $B = \{2, 4, 6, 8, \dots\}$, then a function f from set A into set B may assign to each odd number in set A an even number in set B, and if so, it may take the following form: $1 \rightarrow 2, 3 \rightarrow 4, 5 \rightarrow 6, 7 \rightarrow 8, 9 \rightarrow 10, \dots$. If expressed as a set of ordered pair, in this case $f = \{ \langle 1, 2 \rangle, \langle 3, 4 \rangle, \langle 5, 6 \rangle, \langle 7, 8 \rangle, \langle 9, 10 \rangle, \dots \}$. Actually, neither set A nor set B has an intrinsic structure such as an order among its elements. We have tacitly assigned to both set A and set B an order in terms of the relation ' $m < n$ ' (' m is less than n ') among the elements in them. Since this order relation is the same in both set A and set B, when f maps set A into set B, the same order relation ' $m < n$ ' in set A is of course preserved in set B.

If a function f from set A into set B preserves the structure (such as the order relation) of set A in set B, then f is moreover a homomorphism. By this definition, f in our case is not merely a function, but also a homomorphism. Now suppose that instead f , we have g , which maps set A into itself, so that $g = \{ \langle 1, 1 \rangle, \langle 3, 3 \rangle, \langle 5, 5 \rangle, \langle 7, 7 \rangle, \langle 9, 9 \rangle, \dots \}$. Then g like f is obviously both a function and a homomorphism. Since a compatibility between set A and set B can be stated in either direction, the homomorphism f from set A into set B has an inverse homomorphism f^{-1} from set B into set A. Similarly, the homomorphism g from set A into set A has an inverse homomorphism g^{-1} , again from set A into set A. The homomorphism f mapping set A into set B, as well as its inverse f^{-1} , illustrates the formal property of the *reciprocity* compatibility between two grammatical representations or rules. The homomorphism g mapping set A into set A itself, as well as its inverse g^{-1} , illustrates the formal property of the *identity* compatibility between two grammatical representations or rules. Under either condition, compatibility requires a homomorphism, which may be f or g , having f^{-1} or g^{-1} as its inverse. Thus, two grammatical representations or rules, x and y , are compatible if x and y have identical or reciprocal features on a particular analysis. For example, *singing lessons* contains two identically compatible elements (both nouns), and *singing boys* is composed of two reciprocally compatible elements (with the adjective *singing* and the noun *boys* needing each other), but *suddenly boys* consists of two incompatible elements (since *suddenly* is an adverb and *boys* is a noun).

2.2 Emergent state

Two interfacing or interfaced elements or units in a composition can enter into any one of the three progressively more well-connected states of interface: (i) Connect or Connection, (ii) Mix or Mixture, and (iii) Blend or Blending. An interface state is a necessary but not sufficient condition for an Emergent State, or Emergence, to reesult. A sufficient condition for an emergent state must be determined for each individual case.

Some scientists such as Holland (1998) believe that the complexity theory (Waldrop 1992), which studies how order may develop from chaos under specific conditions, can help us identify the sufficient as well as necessary conditions, under which a particular state will emerge. Hopefully, as the complexity theory makes progress, our understanding of emergence may also improve. In this paper we will be concerned primarily with the interface state as a necessary condition for emergence. Three linguistic examples illustrate the three states of interface in terms of well-connectedness: (i) *black board* is a noun phrase which emerges upon the connection state formed by the adjective *black* and the noun *board*; (ii) ‘black board’ is a semantic composite which emerges upon the mixture state formed by the semantic elements ‘black’ and ‘board’; and (iii) ‘blackboard’ is a semantic entity which emerges upon the blending state formed by the semantic elements ‘black’ and ‘board’. As these examples suggest, the boundary among these three types of interface states is not clear and sharp. In general, connection preserves the categories of the composing elements, mixture obscures them, and blending neutralizes them. For convenience, we will refer to both mixture and blending by Fusion, or Fuse, to be distinct from connection, or connect.

To recapitulate, two elements in a composition can enter into an interface state if there is a homomorphism from one of the two elements to the other, making them compatible or congruous. This interface state can be a fusion or connection. Unless an interface state holds, no emergent state results. Emergence thus depends on an interface state, which in turn depends on the two compatible elements in a composition. Moreover, the two elements in an interface can occur within the same module or submodule or across two different modules or submodules. Hence, interface yields grammar-wide emergence, making grammar an inherently emergent system.

2.3 Interaction modes

A legacy which Wang (1969) and his associates (Cheng and Wang 1972, Chen and Wang 1975, Wang and Lien 1993, Lien 1987, 1992, Yue-Hashimoto 1993, Zhang 1995, Hsieh 1972) bequeathed to the field of general and Chinese linguistics is the ever-inspiring insight that grammatical rules compete, causing lexical diffusion in phonological and syntactic changes. Expanding on Wang’s concept of competition, Hsieh (1992, 1996b) and other scholars (Her 1994, Chang 1991, Cheng 1994, M-L. Hsieh 1992, Huang 1990) have offered approaches to an Interaction Theory. Hsieh’s version of the Interaction Theory recognizes three broad modes of Grammatical Interaction by fine-tuning Wang’s notion of competition into three sub-notions. Two grammatical rules may (i) Complement each other, or are Complementing or In Complementation with each other, if each one of them applies to a separate input to

produce a distinct output, may (ii) Conflict with each other, or are Conflicting or In Conflict with each other, if they apply to the same input to yield two co-existing conflicting outputs, and may (iii) Conspire with each other, or are Conspiring or In Conspiracy with each other, if they jointly produce a single output from either identical or distinct inputs.

Hsieh's three modes of grammatical interaction were initially formulated as types of interaction among rules within the same module or submodule. However, it is easy to extend these intramodular interactional patterns to intermodular interactional patterns, by allowing rules to apply across modules. Since the two interactive modules are compatible, a third and intermediate module can be constructed as an emergent module composed of or compromised from these two original modules. In this emergent intermediate module, rules *analogical* to those in one or both modules can be allowed to apply to representations in the intermediate module. For an illustrative example, consider the semantic configuration, syntactic configuration, and the emergent 'semantic-syntactic' configuration fused from the former two configurations, for the English sentence (1):

- (1) John pushed Bill down. (surface form)
 - a. 'John pushed Bill, then Bill fell'. (semantic form)
 - b. < John, << pushed, Bill >, down >>. (syntactic form)
 - c. (Cause ('secondary'): John pushed Bill; Effect ('primary'): Bill fell).
(semantic-syntactic form)

The intermediate-modular, semantic-syntactic form (1c) for sentence (1) emerges upon fusing the semantic form (1a) and syntactic form (1b) to create a Cause-Effect sequence in (1c), with Cause being ranked as a 'secondary' element and Effect as a 'primary' element. This emergent form (1c) allows an element such as 'twice' to be added to it to modify its primary element, Effect. As a result, (2) *John pushed Bill down twice* has the semantic-syntactic form (2c) (Cause ('secondary'): John pushed Bill; Effect ('primary'): Bill fell twice). Neither (1a) nor (1b) alone would allow (2) to be interpreted as (2c), but their interface (1c) does. Let us see why. In (1b) *pushed* as the verb is the head and *down* as the adverb is the modifier, so *twice*, if added, would modify not the modifier *down* but the head *pushed* on the syntactic level. On the other hand, (1a) is a temporal sequence in the sense of Tai (1985), and *twice*, if added, could modify either the first event or the second event in the temporal sequence. However, when *twice* is added to (1c), it modifies (1c) through modifying its second event, which is its Effect. Assume that a syntactic rule requires a modifying frequency word *twice* to modify the head-verb *push*, and yet an interacting semantic rule requires 'twice' to

instead modify the second temporal event, expressed as the syntactic modifier-adverb *down*. These two rules would disagree, making *twice* in (2) ambiguous. But the semantic-syntactic form in (1c) resembles the semantic form (1a) more than it does the syntactic form (1b), since it reconstrues Tai's temporal sequence as a Cause-Effect sequence. Hence the semantic rule, or rather its semantic-syntactic reinterpretation, would dominate in (1c) and (2c). And since the semantic rule applies in a domain in which the syntactic rule does not apply, the interaction is in the mode of complementation. An intermodular configuration like (1c) or (2c) may seem too abstract, but it still serves a good purpose. When scholars discuss an interaction between syntax and semantics, they usually pretend that the rules interact conveniently in the domain of a surface sentence. But of course the surface sentence has no genuine theoretic status, and it is mainly for avoiding excessively abstract postulations such as a semantic-syntactic module containing (1c) or (2c) that this convenient practice is condoned.

3. Compositional cognitive grammar

3.1 Interactive modules

We adopt a fairly strict modular view on grammar. In this view, the non-lexical and non-phonological part of the grammar consists of Pragmatics, Semantics, Syntax, and Morphology as modules, each subdivisible into submodules. Modules (and submodules) meeting a compatibility condition enter into an interface state, which may under favorable conditions yield an Emergent Module, having emergent configurations or constructions. On these emergent configurations or constructions in the emergent module, the rules of the two original interface modules interact in the three interactional modes: complementation, conflict, and conspiracy. To state somewhat more precisely, let A and B be two interfacing modules, and AB be their emergent module. Then there is a homomorphism, $f: A \rightarrow B$, from A into B, and also a homomorphism, $f^1: B \rightarrow A$, from B into A. However, there is no homomorphism relating A to AB or B to AB, because emergent configurations are unpredictable with respect both to their creation and to their structure. Quite often, rules operating on the emergent module AB are of an analogical nature based on the rules of module A or B or both, as our previous example involving sentence (1) and (2) in English illustrates.

If for some reason we believe that a number of modules are ranked in terms of structural complexity or elaborateness or richness into the decreasing sequence $M_1 > M_2 > M_3, \dots, M_n$, we may conveniently save our expository energy by giving only the typical homomorphism $f: M_k \rightarrow M_{k+1}$. Since f here may stand for a simplification of M_k into M_{k+1} , there may not be an f^1 , unlike the f relating the two compatible elements in

an interface representation. This sequence of homomorphisms is a usual device in current theories including GB (Chomsky 1986, Tang 1991, Huang 1982, Li 1990, Cheng, Huang, Li, and Tang 1997, Chiu 1993, Tsai 1994, Cole, Herman, and Sung 1990) and LFG (Kaplan and Bresnan 1982, Huang 1993). This sequence is also what Hsieh (1994, 1995, 1996a, 1997) has used to organize the several modules in his Compositional Cognitive Grammar (CCG). As Hsieh saw it, a more meaning-revealing representation of a sentence is more complex or elaborate than a less meaning-revealing but more form-economizing representation of the same sentence. Hsieh therefore postulated five non-phonological, non-morphological components of the grammar and relate them by a chain of homomorphisms that sequentially maps a more complex or elaborate configuration in one module into its less complex or elaborate reconfiguration in the next module. Sequenced, these decreasingly complex modules are: Image Structure (IS) \rightarrow Semantic Structure (SS) \rightarrow Thematic Structure (TS) \rightarrow Functional Structure (FS) \rightarrow Constituent Structure (CS). Each sentence has a representation (r) in each such complexity-ranked module, and the various representations (rr) are related by the chain of homomorphisms: $ISr \rightarrow SSr \rightarrow TSr \rightarrow FSr \rightarrow CSr$. This chain of homomorphisms as it applies to the modules or the representations therein is essentially due to what is explicitly or implicitly proposed in LFG.

As Hsieh viewed it, the ISr of a sentence s , or $ISr(s)$ is a fine-grained image underlying s . The $SSr(s)$ is the coarse-grained image of s , and it is where the more iconic dimension (Tai 1989) of s meets its more abstract dimension. The $TSr(s)$ is where the iconic force recedes to allow for the abstract force to prevail, thereby providing a transition from the realm of concrete images to the realm of abstract representations, beginning with $FSr(s)$, continuing with $CSr(s)$, and ending with the morphological, phonological, and phonetic shapes of s .

Let us now focus on SSr , which are a unique feature of CCG. The SSr are some sort of ‘deep-structure’ sentences in which the iconic force, primarily associated with semantics, and the abstract power, primarily associated with syntax, meet each other. However, since SSr are by design primarily ‘syntactic-minded’ configurations, they cannot fully address the question of semantic emergent constructions in an Emergent Grammar, which is what we now want to recast CCG as one. As we saw earlier, there are three interface states: connection, mixture, and blending. An emergent syntactic representation is primarily based on connection, which enables the ‘head’ category in a syntactic composition to ‘percolate up’ continuously to become the category of the entire composition. But an emergent semantic representation is primarily based on mixture and blending, or simply fusion, which creates an unpredictable meaning. Thus, we now find it useful to divide the original $SSr(s)$ for a sentence s into ‘form’ $SSr(s)$ (f - $SSr(s)$) and ‘meaning’ $SSr(s)$ (m - $SSr(s)$). Hence, what is called $SSr(s)$ previously is

now f-SSr(s), and what is added to indicate the fusion-dependent meaning of *s* is now m-SSr(s).

Because a syntactically oriented formula depends on connection, it is easy to build an f-SSr(s) using categories that compare with each other as ‘primary’ / ‘head’ and ‘secondary’ / ‘modifier’, much as it is practiced in the X-bar Theory in GB, or in the Head-driven mechanism in HPSG (Pollard and Sag 1994, Huang 1989), or in the unificational method in Categorical Grammar (Oehrle, Bach, and Wheeler 1988, Moortgat 1988). Moreover, recursiveness can be achieved by using the scheme of endocentric categorical expansion based on the head element of a compositional configuration having the connection property. Hence, a syntactic formula can be indefinitely complex and still be easily assessed as grammatical or not. Grammaticalness condition can thus be elegantly stated: if a compositional configuration is marked by a categorical symbol indicating that it is some sort of a head, or, failing to be a head, is nevertheless composed of a head and a modifier, then the configuration is grammatical or well-formed. By contrast, the characterization of well-formedness for semantically oriented formulas does not enjoy such elegance. It is perhaps next to impossible to identify a non-arbitrary meaning category that would signal a well-formed, complete thought or complete event or complete image. Because of this limitation, in our revised CCG, m-SSr(s) must depend on f-SSr(s) for asserting its status as a well-formed, complete image on the ground that the matching f-SSr(s) is a well-formed formula. This fundamental difference in terms of well-formedness status between a syntactic and a semantic formula could lead one to better understand why Chomsky persistently views syntax as a more tractable and rewarding module of the grammar than semantics.

3.2 Semantic structure

3.2.1 f-SSr(s)

The ‘building blocks’ of an f-SSr(s) for a sentence *s* consist of 28 Action Frames (ACFs). Each ACF is a general-pattern simple Action (s-AC), composed of an Act (A) as a (‘deep-structure’) verb, which composes with a Receiver (R) as an object to form a Complex Act (A’) as a predicate, which in turn composes with an Initiator (I) as the subject to form an Action (AC) as a (simple) sentence. Thus, the general form of an s-AC is $s\text{-AC} = \langle I, A' = \langle A, R \rangle \rangle$. The A is (represented by) an Abstract Verb (AV) having one of three possible categories: Full Verb (FV), Half Verb (HV), and Grammatical Verb (GV). Depending on the ACF type or a larger context, each AV type occurring in an ACF has a varying Concrete Shape in the CSr, such as verb, adjective, or adverb in the traditional analysis. A particular lexical AV enters an ACF to make it a

Particularized ACF, or PACF, which is an s-AC. Hence, based on the ‘verbal-type’ (‘v-type’) determined by the AV category, an s-AC falls into three categories: Full Verb Action (FAC), Half Verb Action (HAC), and Grammatical Verb Action (GAC). The I and R, if not empty, are each (represented by) an indexed variable v_k (for convenience written x_k for I and y_k for R), or a simple constant c or complex constant $f(c)$ (c is for convenience written h for I and k for R).

An indexed variable v_k refers to an event-participant indexed by k . A simple constant c in an s-AC indicates the site of an embedding AC, for technical reasons externally placed and marked by ‘=c’. A complex constant $f(c)$ in an s-AC is a remark that modifies an external AC, marked by ‘=c’, with respect to its selected feature such as ‘time’, ‘place’, ‘manner’, ‘aspect’, or ‘agreement’. The external AC marked by ‘=c’ composes with the s-AC containing c or $f(c)$ to ‘saturate’ it, or more precisely to saturate the c or $f(c)$ in it. If an s-AC is not saturated, or ‘unsaturated’, it will stand alone and contain one or even two $c/f(c)$ ’s. Thus, depending on the number of unsaturated c ’s or $f(c)$ ’s, an s-AC falls into three ‘nominal-type’ (‘n-type’) categories: Solitary Action (SAC), which contains no unsaturated constants, and Receptive Action (RAC), which contains one unsaturated constant, and Warm Action (WAC), which contains two unsaturated constants. The three v-type categories and the three n-type categories combine to give each s-AC one of the nine possible Composite Categories: FAC-SAC, FAC-RAC, FAC-WAC, etc. (see Table 1).

Two s-ACs compose into a complex Action (c-AC), and in general two less complex ACs compose into a more complex AC, recursively. When two ACs compose into a more complex AC, the composite category of the resulting AC is inherited from the composite categories of these two ACs, according to a set of stipulation rules (Tables 2 and 3). Each element within an s-AC is specified as ‘primary’ (‘p’) or ‘secondary’ (‘s’) relative to its co-composing element. By stipulation, we have $s\text{-AC} = \langle I('s'), A('p') = \langle A('p'), R('s') \rangle \rangle$. Within a c-AC, the ‘p’ or ‘s’ status of any AC is obtained through computation based on the composite categories of its two composing ACs, which may each be simple or complex (Tables 4 and 5). Thus, for example, FAC-SAC compares with HAC-RAC as ‘p’ compares to ‘s’, since by Table IV FAC versus HAC is $\langle 2,0 \rangle$, and by Table V SAC versus RAC is $\langle 0,1 \rangle$, and $\langle 2,0 \rangle$ and $\langle 0,1 \rangle$ add up to $\langle 2,1 \rangle$, in favor of FAC-SAC as ‘p’. The ‘p’-to-‘s’ relation of two co-composing elements in a composition on all levels translates into their linear order in SSr and in CSr (Tai 1973, Li 1990, C-T. Huang 1994).

Table 1. Action frames (ACFs) classified.

	< I, < A, R >>				< I, < A, R >>				< I, < A, R >>			
SAC	(1)	x_i	FV	y_j	(10)	x_i	HV	y_j	(19)	x_i	GV	y_j
	(2)	x_i	FV	0	(11)	x_i	HV	0	(20)	x_i	GV	0
	(3)	0	FV	y_j	(12)	0	HV	y_j	(21)	0	GV	y_j
	(4)	0	FV	0	(13)	0	HV	0	(22)	0	GV	0
RAC	(5)	h	FV	y_j	(14)	h	HV	y_j	(23)	h	GV	y_j
	(6)	h	FV	0	(15)	h	HV	0	(24)	h	GV	0
	(7)	x_i	FV	k	(16)	x_i	HV	k	(25)	x_i	GV	k
	(8)	0	FV	k	(17)	0	HV	k	(26)	0	GV	k
	(27)	x_i	k	0								
WAC	(9)	h	FV	k	(18)	h	HV	k	(28)	x_i	GV	k
	FAC				HAC				GAC			

FV = Full Verb (verb, adjective).

HV = Half Verb (preposition, conjunction, adverb, auxiliary, aspect, tense, etc.).

GV = Grammatical Verb (demonstrative, determiner, grammatical particles, etc.).

F/H/GAC = Full/Half/Grammatical Verb AC.

SAC = Solitary AC, containing no unsaturated constants.

RAC = Receptive AC, containing one unsaturated constant.

WAC = Warm AC, containing two or more unsaturated constants.

Notes:

Within an ACF, we may find (i) an indexed variable v_m , written as x_i for I and y_j for R, with i and j referring to distinct participants, or (ii) a simple constant c, written as h for I and k for R, or (iii) a complex constant $f(c)$, written as $f(h)$ for I and $f(k)$ for R, or (iv) an empty symbol 0, or (v) an abstract verb FV, HV, or GV. To simplify the table, an ‘h’ represents either h or $f(h)$, and a ‘k’ represents either k or $f(k)$. Filled with a particular AV (Abstract Verb), an ACF becomes an s-AC (simple AC). The c in an ACF or s-AC indicates an embedded AC externally placed, and the $f(c)$ refers to a feature (such as ‘time’, ‘place’, ‘manner’, ‘agreement’) of a co-composing, externally placed AC, identified as c and marked by ‘=c’. The external AC marked by ‘=c’ co-composes with an s-AC containing a constant c or $f(c)$ to ‘saturate’ it, or more precisely its constant. Each ACF or s-AC has a verbal type based on its FV, HV, or GV and also a nominal type based on its 0, 1, or 2 ‘unsaturated’ constants. The three verbal types, FAC, HAC, and GAC, combine with the three nominal types, SAC, RAC, and

WAC, to form nine possible composite types for an ACF or s-AC. ACF (27) combines with an FAC-SAC to form a topic-comment construction.

Table 2. Computation of v-types.

	FAC	HAC	GAC
FAC	(i) FAC	(ii) FAC	(iii) (a) FAC, if GAC is an SAC (b) GAC, if GAC is an RAC or WAC
HAC		(iv) HAC	(v) (a) HAC, if GAC is an SAC (b) GAC, if GAC is an RAC or WAC
GAC			(vi) GAC

Table 3. Computation of n-types.

	SAC	RAC	WAC
SAC	(i) SAC	(ii) SAC	(iii) RAC
RAC		(iv) RAC/WAC	(v) RAC/WAC
WAC			(vi) RAC/WAC

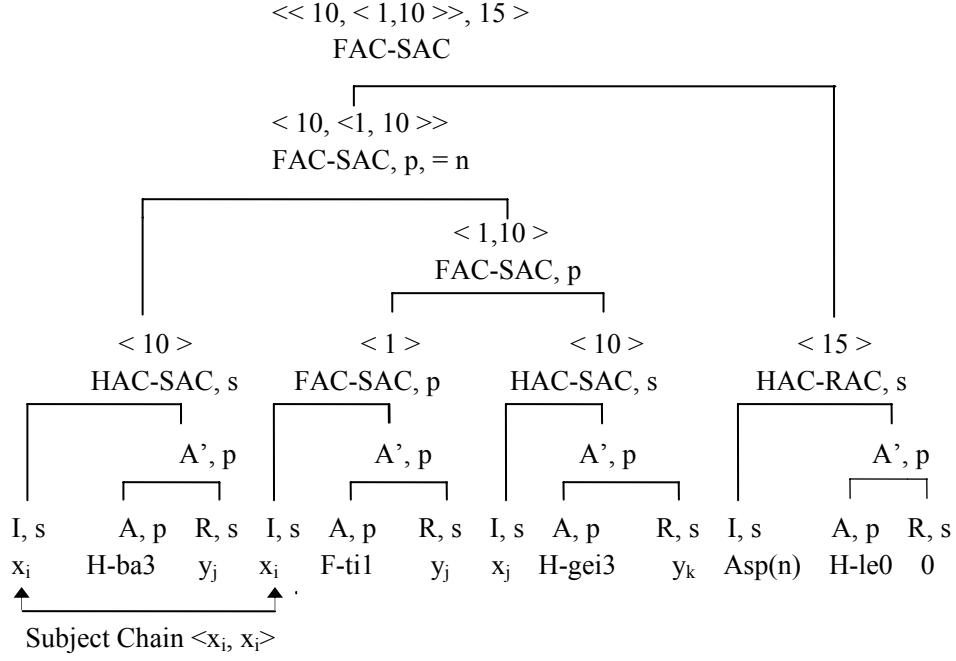
Table 4. Computing the primacy degrees in v-types.

	FAC	HAC	GAC
FAC	(i) <2,2>	(ii) <2,0>	(iii) <2,0>
HAC	(iv) <0,2>	(v) <2,2>	(vi) <2,0>
GAC	(vii) <0,2>	(viii) <0,2>	(ix) <2,2>

Table 5. Computing the primacy degrees in n-types.

	SAC	RAC	WAC
SAC	(i) <1,1>	(ii) <0,1>	(iii) <0,1>
RAC	(iv) <1,0>	(v) <1,1>	(vi) <0,1>
WAC	(vii) <1,0>	(viii) <1,0>	(ix) <1,1>

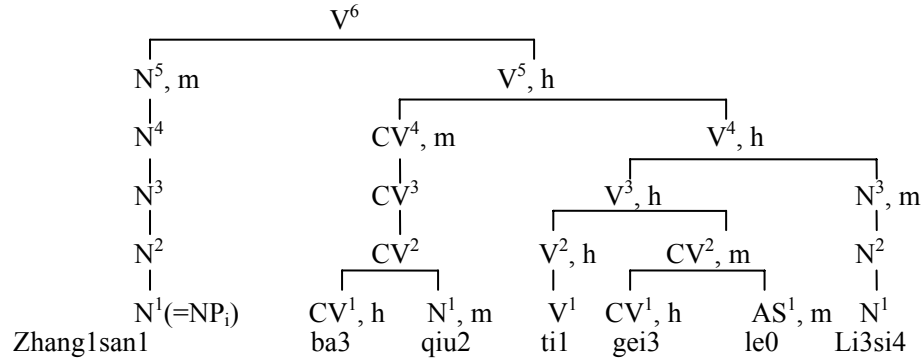
Figure 1. f-SSr(3).



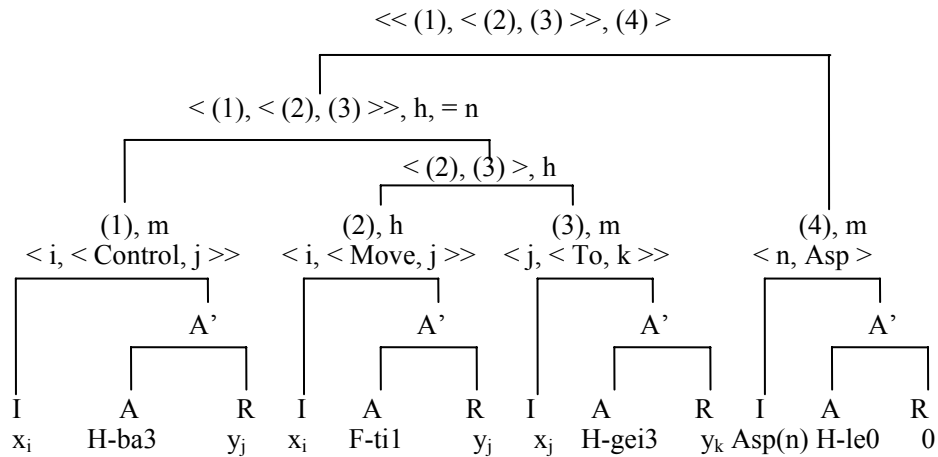
(3) Zhang1san1(=i) ba3 qiu2(=j) ti1 gei3 le0 Li3si4(=k).
kicked the ball to L'.

Note:

x_i = The First Thematic Proto-role in SSr(3); NP_i = The First Argument NP in CSr(3); Subject is the x_i or NP_i, or more precisely, the function mapping x_i onto NP_i, here Zhang1san1.

Figure 2. CSr(3).**Notes:**

After the indexed variables x_i, y_j, y_k are instantiated by their co-indexed NP_i, NP_j, NP_k , which are suitably placed in the f-SSr, and after simplification of the instantiated f-SSr, CSr results. For each CSr category, $X = N, CV, V, AS(pect)$, etc., $X^k = X$ phrase of degree k . $X^1 = \text{lexical } X \text{ (or minimal } X \text{ phrase)}$. The composite category X^{m+1} is derived from $\{X^m, Y^m\}$, where X^m is head (h) and Y^m is modifier (m), much as in the X-bar theory in GB. The linear order of the two elements marked as 'h' and 'm' in a composition is achieved by interpreting the 'h'-to-'m' relation as 'h' preceding or following 'm', according to specific rules.

Figure 3. m-SSr(3).

If an AC, either s-AC or c-AC, has the composite category FAC-SAC, it in principle anticipates a 'well-formed' CSr and consequently a 'grammatical' surface

sentence. Tense, aspect, agreement, and other grammatical features may often be required to make this virtual grammatical sentence actually grammatical in particular languages. Thus, FAC-SAC is essentially synonymous with syntactic well-formedness or grammaticality. Although technically an FAC-SAC as an f-SSr(s) cannot map directly onto its CSr(s), in the current practice this must be done, because the TSr(s) and FSr(s) still wait for a full articulation. With this adjustment, we map each f-SSr(s) onto its CSr(s), according to a five-step procedure: (i) replacing each AV with its concrete shape, (ii) ‘instantiating’ the indexed variables v_k with a suitably placed co-indexed NP_k , (iii) simplifying the instantiated f-SSr tree in a structure-preserving manner into the CSr tree, and (iv) renaming the f-SSr categories as simple CSr categories (in terms of verb, adjective, adverb, etc.) and using them to generate complex CSr categories by the familiar head projection technique.

Figure 1 provides a concrete example of an f-SSr(s) in the form of f-SSr(3) as it represents sentence (3) *Zhang1san1 ba3 qiu2 ti1 gei3 le0 Li3si4* ‘Z kicked the ball to L’. As Figure 1 shows, PACF<1> and <10> compose into AC<1, 10>, then PACF<10> and AC<1,10> compose into AC<10, <1, 10>>, and finally AC<10, <1, 10>> and PACF<15> compose into AC<<10, <1, 10>>, 15>, which is the f-SSr(s) for (3), i.e., f-SSr(3). Within each PACF, particularizing an ACF into an s-AC, the I and R are either represented by the indexed variables (x_i, y_j in <10>, x_i, y_j in <1>, x_j, y_k in <10>), or a complex constant (Asp(ect)(n) in <15>), or are empty (0 in <15>). Each s-AC has a composite category, and when two ACs compose into another AC, their composite categories also compose into another composite category for the resulting AC. For example, when <1> and <10> compose into <1, 10>, their FAC-SAC and HAC-SAC also compose into FAC-SAC for <1, 10>. Also, the co-composing ACs in a c-AC are assessed for their ‘p’ or ‘s’ status by computation, so that, for example, within <1, 10> FAC-SAC is ‘p’ and HAC-SAC is ‘s’. By a sequence of transformations, f-SSr(3) in Figure 1 maps onto CSr(3) in Figure 2.

3.2.2 m-SSr(s)

Since the m-SSr(s) depends on f-SSr(s) by design, once we have set up the f-SSr(s), we are ready to obtain m-SSr(s) with minimal further maneuvering, as in Figure 3. We simply assign to each lexical AV in a PACF, such as H-ba3, F-ti1, H-gei3, and H-le0 in the m-SSr(s) one of a finite number of simple Construction (s-Con) types. Thus, for example, H-ba3 has the s-Con Control, F-ti1 Move, H-gei3 To, and H-le0 Aspect. Having inherited the s-Con type from its AV, a PACF now is reinterpreted as a typed s-Con, such as Control in PACF <10>. We then elaborate this typed s-Con by reinterpreting the indexed variables x_i and y_j (representing I and R) in the PACF as its Thematic Proto-roles. Thus, we obtain an Elaborated s-Con, such as <i,

<Control, j>> in Figure 3, based on PACF <10> in Figure 1. In reinterpreting the x_i and y_j as thematic proto-roles, three general rules apply: (i) x_i compared to y_j is Proto-agent, (ii) y_j compared to x_i is Proto-patient, (iii) x_i or y_j alone is Proto-theme. For example, the s-Con Control originated as PACF<10> in Figure 1 but elaborated as (1) <i, <Control, j>> in Figure 3, has $i(=x_i)$ and $j(=y_j)$ as its proto-agent and proto-patient (Dowty 1991).

As we also see in Figure 3, the next step is to continuously compose these s-Cons into a complex Construction (c-Con), each time increasing its degree of complexity. This recursive procedure is based on the compositional structure provided by an already-formulated f-SSr(s), in this case f-SSr(3) in Figure 1. Since the s-Con types or patterns are finitely small, we can in principle rank every two s-Cons under comparison by seeing which one expresses a more dynamic force transmission (Talmy 1988) than the other. The more dynamic s-Con is then chosen as the ‘head’ (‘h’) element and the less dynamic one as the ‘modifier’ (‘m’) element. Thus, a conceivable descending ‘dynamicness’ ranking of the four s-Cons in m-SSr(3) would be: Move > To > Control > Aspect. Applying the familiar head projection or head-driving, we preserve the relative dynamicness rank of a ‘head’ s-Con in a composition as the relative dynamicness rank of that composition. For example, Move is the head in <Move, To>, since Move ranks higher than To. <Move, To> is the head in <Control, <Move, To>>, since its inherited rank is the rank of Move, but Move ranks higher than Control. In this way, the head status of Move is continuously inherited or preserved as the head status of the compositional structure on every level of the m-SSr(3) tree.

All this means that we generate a ‘headed’ c-Con by means of a head-driven procedure working on the dynamicness-ranked s-Cons. Moreover, the head-driven procedure enables us to characterize a well-formed image. As the case of m-SSr(3) illustrates, each c-Con has a head that originates as the head in the smallest-size or lowest-level c-Con, where its status as the head is determined by prior ranking. If we can characterize this head status in general terms, we would in principle be able to characterize a complete image, or a well-formed image on that basis. This we can achieve by setting up four Self-sufficient Super s-Cons (SSCs), each one corresponding to a major sentential type: BE (for Subject-Predicate sentence), BECOME (for Subject-Become-Complement sentence), DO (for Intransitives), and AFFECT (for Transitives). A c-Con and especially an m-SSr(s) is then well-formed if its head s-Con belongs to one of these four SSCs. By this definition, in Figure 3, s-Cons (1), (3) and (4) are not well-formed, since none of them belongs to an SSC. But s-Con (2), <i, <Move, j>>, is, since it belongs to AFFECT, and any c-Con having s-Con (2) as its head, especially the entire c-Con constituting an m-SSr(s) is well-formed.

Jackendoff (1983) and, following him, Pinker (1989) have proposed a semantic

well-formedness strategy essentially along this line of thinking, but using a technique that rewrites what amounts to a primordial SSC into an elaborate construction having a prescribed internal structure. Just as in Jackendoff's and Pinker's approaches, our m-SSr(s) depends on an f-SSr(s) for its homomorphic syntactic counterpart, and in this sense m-SSr(s) constitutes a syntax-dependent semantic configuration. Furthermore, primarily due to this dependency on syntax, m-SSr(s) is a coarse-grained image, unlike the protean fine-grained image based on a syntax-independent semantic configuration, such as Langacker (1987, 1995) and Talmy (1985, 1988) have focused their insightful studies on. From a modular view on grammar, the clear contrast among a fine-grained image, a coarse-grained image, and an abstract syntactic representation would not be surprising but revealing instead. Modules having varying degrees of iconicity or abstraction serve precisely to make a huge and complicated system such as grammar capable of effective functioning and adaptation to ever new needs.

Although we have not postulated any fusion within any c-Con in the m-SSr(3) in Figure 3, a c-Con in principle allows fusion to take place in it, as we shall see later. Since a Con is both compositional and fusion-prone, it simultaneously addresses the problems of composition and emergence in a grammar. It helps to explain why grammar is not only orderly operational like a mathematical structure, but at the same time unpredictably emergent, having the kind of evolutionary thrust and historical force that Wang (1991, ch. 6, 7, 8) insightfully pointed out to be present in language as a living organism. Once we can choose a compositional *and* emergent grammar, such as CCG hopefully serves to illustrate, we no longer have to accept Saussure's forced convenient division of a language into structure and history as two vaguely related and yet irreconcilable parts, to be studied separately. Language has a structure that, as it develops, creates its own history through emergent processes, just as any organism.

3.3 Emergent constructions

Emergence is a grammar-wide state or phenomenon and not limited only to privileged modules. In a strictly compositional grammar, such as CCG, emergence can be expected to start at the very beginning of the composition of a sentence, namely, within an s-AC. Thus, for example, in Mandarin the AV F-*da3* composes with an y_j to form $(\alpha) \langle F\text{-}da3, y_j \rangle$ in the PACF $\langle 1 \rangle \langle x_i, \langle F\text{-}da3, y_j \rangle \rangle$. The semantic content of α varies according to the meaning of the NP_j that later will instantiate y_j . For example, if NP_j = *qiu2* 'ball', then (α) = 'play' y_j ; if NP_j = *dian4hua4* 'telephone', (α) = 'operate' y_j ; if NP_j = *Li3*, (α) = 'beat' y_j ; if NP_j = *xin4hao4* 'signal', then (α) = 'make' y_j , and so on (see S. Huang 1994 for the thesis that 'metaphorical' extension of verbal meaning creates verbal polysemy). A common practice has been to postulate a division of *da3*

into various shapes in anticipation of the varying senses of (α) conditioned by NP_j , thus skipping the emergent state at this beginning level of composition. However, this common lexical-division alternative to emergence fails as soon as we move onto higher levels of composition, to which lexicon has no plausible access. And this failure occurs as early as when two s-Cons fuse into an unpredictable c-Con. As the Construction Grammarians (Fillmore 1988, Fillmore, Kay, and O'Conner 1988, Goldberg 1995, Cheng 1997) have rightly claimed, such emergent c-Constructions, or simply 'Constructions', defy the treatment of a purely compositional or generative grammar. However, the Fillmorean group and others are primarily interested in arguing for the need to recognize Constructions, without also being interested in providing a grounding compositional grammar, such as CCG. Therefore, the sketchiness and mystery evoked by proposed specific Constructions may mystify those who believe in the orderly and reliable operation of a non-Constructional, formal grammar. This drawback can be removed. The Fillmorean Constructions or their like can be naturally and systematically embedded in a CCG, making grammar essentially or primarily compositional and yet with emergent capability, everywhere in the structure and anytime in the history of a language. To see how we can achieve this goal with CCG, consider a pair of typical Goldberg sentence in (4) and (5):

(4) Sally sneezed.

(5) Sally sneezed the napkin off the table.

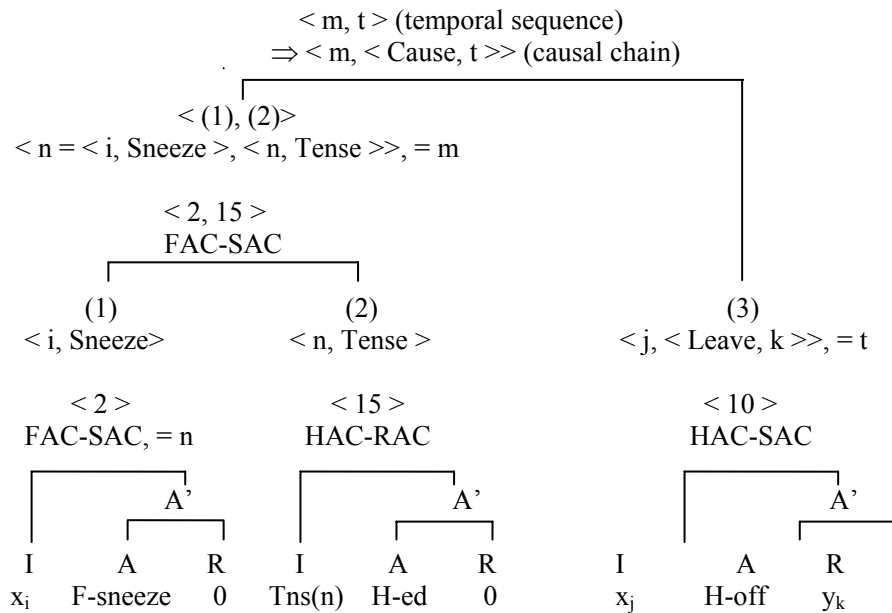
Goldberg (1995) argued that *sneeze* is normally intransitive as in (4) and there is thus no compelling reason to postulate a transitive sense solely for the purpose of accounting for the transitive-causative pattern in (5). Hence, she reasoned, the grammar of English must have, in some mysterious and capricious way so it seems, created a special causative Construction, of which (5) is merely a particular example. When we analyze (5) into an m-SSr(5) as in Figure 4, the emergent construction postulated for (5) by Goldberg is easily stated as the result of a fusion of $\langle m, t \rangle$ as a temporal sequence (Tai 1985) into $\langle m, \langle \text{Cause}, t \rangle \rangle$ as a causal chain. As shown, $m = \langle n = \langle i, \text{Sneeze} \rangle, \langle n, \text{Tense} \rangle \rangle$, and $t = \langle j, \langle \text{Leave}, k \rangle \rangle$, and (5) therefore means that Sally's sneezing caused the napkin to leave the table (by being blown at).

Fortunately or unfortunately, much of linguistics is still under the shadow of the Chomskyan revolution. Chomsky inherited an essentially Cartesian static view of the world and language. But as Deacon (1997) has persuasively argued, language is the product of the biological evolution of the human brain, hence universal grammar (such as Chomsky's x-bar syntax) need not be innate. Rather, it could have resulted from the convergence of separately emergent specific grammars that nevertheless are facilitated

by ‘an innate *bias* for learning in a way that minimizes the cognitive interference that other species encounter when attempting to discover the logic behind symbolic reference’ (Deacon 1997:141). The innateness issue aside, grammar and discourse are likely emergent systems, just as the biological views espoused by Deacon, Wang (1991) and others have suggested.

If grammar is rigidly compositional (or generative) without any flexibility provided by an emergence capability, grammar as an organism cannot evolve from a less complex symbolic system into an increasingly more complex one. Nor can it undergo any structural change in history once it has become fully evolved. If discourse were structurally predetermined just like a fancied rigid compositional grammar, discourse structure cannot change historically, either. But as Hopper (1987) and Hopper and Thompson (1984) have argued, and as Huang (1996), Chang (1996), and Heine (1993) and others have documented, developing new patterns in discourse and grammar are continuously created from old ones by the seemingly magic mechanism of emergence. In fact, many of the fine results achieved in Chinese diachronic syntax (Mei 1994, Li 1980, Zhang 1994, Yue-Hashimoto 1993) can be alternatively interpreted as results showing the ubiquitous grammaticalization process caught in its historical mode in an emergent grammar.

Figure 4. Fusion in a Goldberg sentence.



(5) Sally sneezed the napkin off the table (Goldberg 1995).

4. Subject chain

We have spent much energy and space to prepare ourselves now for a precise characterization of the subject chain. This subject chain will enable us to define the subject function, or subject, in Mandarin, and will help us understand why the Mandarin subject is grammatically genuine, even though it virtually lacks any morphological markings for agreement with its verb, in terms of person, number, and case.

The subject chain is actually a tactically convenient alternative name for the thematic proto-role chain. In the f-SSr(s), several distinct (thematic) proto-roles may occur. A proto-role k is a reinterpretation of a v_k representing I or R in an s-AC. The k may occur once or in several copies, in which case these copies of k form a proto-role chain $\langle k, k, \dots, k \rangle$, abbreviated as (k) . This chain is for our purpose isomorphic to the indexed-variable chain $\langle v_k, v_k, \dots, v_k \rangle$, abbreviated as (v_k) . To avoid unnecessary complication, we will sometimes simply speak of the indexed variable v_k or the chain (v_k) , with v_k varying between x_k and y_k , when we may actually mean the proto-role k or the proto-role chain (k) . If a small number of proto-role chains, (k) , $(k+1)$, $(k+2)$, etc. are present in an f-SSr(s), the chain (k) generated by x_k , or by y_k if x_k is absent, in the first s-AC is chosen as the relevant proto-role chain (k) . This relevant proto-role chain (k) or (x_k) is the long-awaited subject chain. Corresponding to each subject chain (x_k) in the SSr(s), there is an NP_k in the CSr(s), and the two are of course related. There is a function, called subject function, that maps each (x_k) onto a corresponding NP_k , making it the subject of s . This subject function depends on the (x_k) and not the AVs in the f-SSr(s) hosting this (x_k) . Hence, the relation of the (x_k) (mapping onto CSr(s) subject) to the AVs (mapping onto CSr(s) verbs or coverbs) is only of secondary importance as far as the subject function is concerned. This is why explicit morphological markings for subject-verb agreement is not necessary, and in fact not employed, in Mandarin.

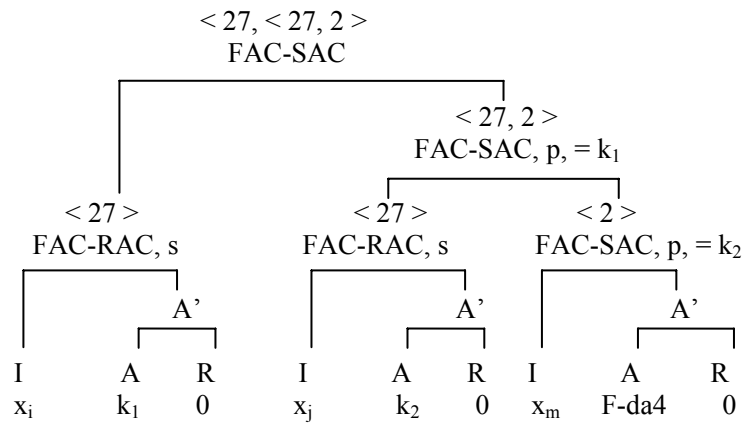
For an illustration of the subject chain and the subject, let us look back at Figure 1. In the f-SSr(3) in Figure 1, we assign a linear order to the s-ACs and c-ACs by interpreting a pair $\{ 'p', 's' \}$ of elements in a composition as 'p' preceding 's' or as 'p' preceding 'p', according to rules essentially analogous to those proposed by C-T. Huang (1994) for the CSr(s) word order. The f-SSr(3) tree in Figure 1 tacitly shows the resulting f-SSr(3) word order, which is essentially preserved as the CSr(3) order in Figure 2. This f-SSr(3) word order determines PACF $\langle 10 \rangle$ (H-ba3) as the first s-AC. Now PACF $\langle 10 \rangle$ has an x_i , which generates the proto-role chain (x_i) , composed of one x_i in PACF $\langle 10 \rangle$ and another x_i in PACF $\langle 1 \rangle$. This subject chain (x_i) of f-SSr(3) in Figure 1 maps onto the NP_i (written N^1 and lexically represented by Zhang1san1) of CSr(3) in Figure 2. This mapping is the subject function, and the NP_i , Zhang1san1, is the subject of sentence (3). Notice that neither f-SSr(3) nor CSr(3) provides any means

for subject-verb agreement morphology. This is because subject-verb agreement is a secondary and not a primary factor in the subject function creating the subject. If someday Mandarin should develop a subject-verb agreement marking in terms perhaps of particles, we could add a PACF <15> of the form <Agreement (m), <H-Arg, 0>> to the original f-SSr(3), and this hypothetical PACF <15> would spread the agreement feature ‘Arg’ to the subject and the verb and even co-verbs, to be there realized as attached particles.

This view of Mandarin subject-verb agreement does not contradict a GB view, which would postulate a subject-verb agreement parameter, with some languages requiring inflectional marking, some particle marking, and some no marking at all. Our view tacitly concurs with a GB view, but offers a reasonable explanation as to why Mandarin opts for this ‘no marking’ choice in the universal subject-verb agreement parameter.

A discussion of the subject in Mandarin would not be complete without a comment on the Topic construction which Tsao (1990), Li and Thompson (1981), and others have helped to identify and articulate as a unique grammatical pattern in Chinese and other languages. We need at least to distinguish between the Stacked-topics construction and the Topicalized NP construction. For the Stacked-topics construction, CCG provides a special device in ACF<27> that, as Figure 5 shows, can through embedding stack as many topics as desired. In this pattern as illustrated in Figure 5, every x_k anticipating a CSr NP_k is Topic, except the $x_m(=NP_m)$ originating in the ACF<2>, which is itself the subject, provided that subject chain is revised with a qualifying clause that effectively excludes any PACF<27> from consideration.

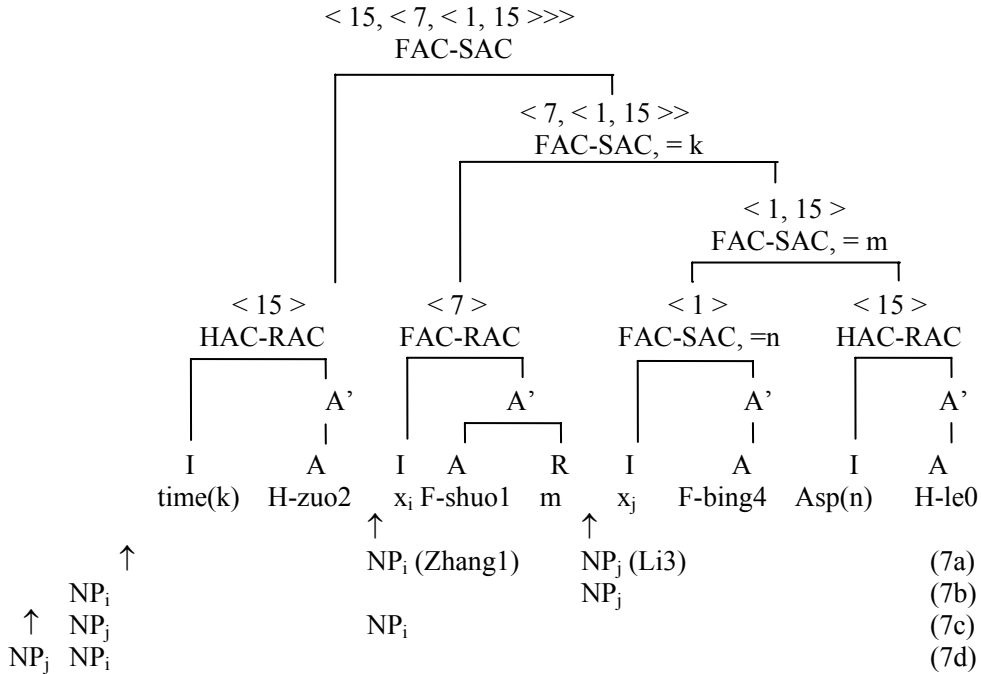
Figure 5. Stacked topics.



- (6) Zhang1san1(=i) mei4mei0(=j) yan3jing1(=m) da4.
 ‘Z’s sister has big eyes’.

Unlike stacked topics, topics derived from topicalization NPs are obtained by extraordinary or abnormal patterns of instantiation on the variables in the SSr(s) that yield emphatic NPs as topics in the CSr(s). As Figure 6 shows, sentence (7)(a) and (b) are subject-sentences. Through left-attachments to suitable ACs, the subject NP_i (=Zhang1) alternately precedes PACF <7> in (7)(a) or PACF <15, <7, <1, 15>>> in (7)(b). The time expression *zuo2tian1* originating as PACF <15> favors this alternation. However, in both cases, NP_j (= Li3) precedes <1, 15>, as is normal. By contrast, (7)(c) and (d) are topicalized-sentences, in which the topicalized NP_j (= Li3) is left-attached to the tree already containing NP_i as the subject, or in common transformational parlance, moves to precede the subject NP_i (= Zhang1). Although we have offered no novel insights on topics beyond those offered by Tsao (1990) and others, we have nevertheless shown that the subject chain can explicate established ideas on topics beyond accounting for the subject.

Figure 6. Subject and topic.



- (7) a. Zuo2tian1 Zhang1san1 shuo1 Li3-si4 bing4 le0.
 b. Zhang1san1 zuo2tian1 shuo1 Li3-si4 bing4 le0.
 c. Li3si4, zuo2tian1 Zhang1san1 shuo1 ta1 bing4 le0.
 d. Li3si4, Zhang1san1 zuo2tian1 shuo1 ta1 bing4 le0.
 ‘Yesterday Zhang1san1 said that Li3si4 is sick’. (a,b)
 ‘As for Li3si4, Zhang1san1 said yesterday that he is sick’. (c,d)

Note:

(7a-d) result from four varying patterns of instantiation; the subject is the first NP (Zhang1), and a Topic is any NP (Li3) preceding the subject.

5. Conclusion

The information about the subjecthood of a sentence \underline{s} in Mandarin is stored in f-SSr(s) as the relevant thematic proto-role chain (x_k), which is conveniently called the subject chain. When the subject function maps f-SSr(s) onto CSr(s), the subject chain (x_k) transforms into the first argument NP_k there, which then becomes the subject of \underline{s} . If a sentence expresses an event having variously ranked participants taking parts in various sub-events, then the subject chain (x_k) codes the primary participant in the entire event. It merely refers to the primary participant interpreted as the thematic proto-role k, without indicating how k is related to the SSr(s) AVs, which code the sub-events expressed by the CSr(s) verbs or co-verbs. Therefore, in the CSr(s), the relation of the subject to the verbs or co-verbs is only of secondary importance. Consequently, Mandarin requires no explicit subject-verb agreement markings.

The subject chain feeding into the subject function is made precise in terms of CCG, which is now freshly endowed with the key features of an emergent grammar, including interface, interaction, emergence, and modules. No matter how much we may desire a formally elegant grammar, a somewhat messy emergent grammar is something we can no longer avoid if, as Deacon and Wang have suggested, language is primarily an evolutionary organism, teeming with disorderly and unpredictable patterns alongside a fairly stable core structure.²

² The few linguistic dialogues which I occasionally enjoyed with Shuanfan Huang has been instrumental in making me see emergence as one basic principle in grammar and discourse. If competition as William S-Y. Wang sees it, or interaction as I see it, is another basic principle, then in a fluid, complex structure underlying language, order can emerge from chaos, through competition and interaction.

References

- Chang, Claire Hsun-huei. 1991. Verb copying: Toward a balance between formalism and functionalism. *Journal of Chinese Language Teachers Association* 26.1:1-32.
- Chang, Miao-hsia. 1996. The grammaticalization of *beh* in Taiwanese Hokkien. *Proceedings of NACCL8* 1:29-51.
- Chen, Matthew, and William S-Y. Wang. 1975. Sound change: Actuation and implementation. *Language* 51:225-81.
- Cheng, Chin-chuan, and William S-Y. Wang. 1972. Tone change in Chao-zhou Chinese. *Papers in Linguistics in Honor of Henry and Renee Kahane*, ed. by B. Kachuru et al., 99-113. Urbana: University of Illinois Press.
- Cheng, Lisa L.-S., C.-T. James Huang, Y.-H. Audrey Li, and C.-C. Jane Tang. 1997. Hoo, hoo, hoo: Syntax of the causative, dative, and passive constructions in Taiwanese. Manuscript.
- Cheng, Robert. 1994. Lexical diffusion in syntactic change: The case of Taiwanese aspect and phase markers. In *Honor of William S-Y. Wang: Interdisciplinary Studies on Language and Language change*, ed. by Matthew Y. Chen and Ovid J. L. Tzeng, 87-103. Taipei: Pyramid Press.
- _____. 1997. Lexical and morphological phenomena: Pronouns in Taiwanese. *Chinese Languages and Linguistics* 3:1-44.
- Chiu, Bonnie. 1993. *The Inflectional Structure of Mandarin Chinese*. Los Angeles: UCLA dissertation.
- Chomsky, Noam. 1986. *Knowledge of Language: Its Nature, Origin, and Use*. New York: Praeger.
- Cole, P., G. Hermon, and L-M. Sung. 1990. Principles and parameters of long-distance reflexives. *Linguistic Inquiry* 21:1-22.
- Deacon, Terrence W. 1997. *The Symbolic Species: The Co-evolution of Language and the Brain*. New York: Norton.
- Dowty, David. 1991. Thematic proto-roles and argument selection. *Language* 67.3:549-619.
- Fillmore, Charles J. 1988. The mechanisms of 'constructional grammar'. *BLS* 14:35-55. Berkeley: Berkeley Linguistics Society.
- _____, Paul Kay, and Catherine O'Connor. 1988. Regularity and idiomaticity in grammatical constructions: The case of *Let Alone*. *Language* 64:501-538.
- Goldberg, Adele E. 1995. *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago and London: The University of Chicago Press.
- Heine, Bernd. 1993. *Auxiliaries: Cognitive Forces and Grammaticalization*. New York and Oxford: Oxford University Press.

- Her, One-soon. 1994. Interaction of syntactic changes. *Chinese Languages and Linguistics II: Historical Linguistics*, 263-292. Symposium Series of the Institute of History and Philology no. 2. Taipei: Academia Sinica.
- Holland, John H. 1998. *Emergence: From Chaos to Order*. Reading, Massachusetts: Addison-Wesley.
- Hopper, Paul. 1987. Emergent grammar. *BLS* 13:139-157.
- _____, and Sandra Thompson. 1984. The discourse basis for lexical categories in universal grammar. *Language* 60:703-52.
- Hsin, Hsin-I. 1972. Lexical diffusion: Evidence from child language acquisition. *Glossa* 6:89-104.
- _____. 1992. In search of a grammatical foundation for dialect subgrouping. *Chinese Languages and Linguistics I: Chinese Dialects*, 333-377. Symposium Series of the Institute of History and Philology, no. 2. Taipei: Academia Sinica.
- _____. 1994. A computational syntax and its application to parsing. *Proceedings of ROCLING* 7:243-252. Hsinchu: Tsinghua University.
- _____. 1995. Thematic control and cross-dialectal comparison. Paper presented at ISOLIT-II. Taipei: National Taiwan University.
- _____. 1996a. Compositional cognitive grammar: An outline. Manuscript.
- _____. 1996b. A taxonomy of grammatical interactions. *Proceedings of ICCL4/NACCL* 7, ed. by Tsai-fa Cheng, Yafei Li, and Hongming Zhang, 1:123-140. Los Angeles: University of Southern California.
- _____. 1997. Lexicon and morphology in a compositional grammar: With particular reference to Chinese verbal compounds. *Chinese Languages and Linguistics II: Morphology and Lexicon*, ed. by Feng-fu Tsao and H. Samuel Wang, 333-368. Symposium Series of the Institute of History and Philology, no. 2. Taipei: Academia Sinica.
- Hsieh, Miao-ling. 1992. Analogy as a type of interaction. *Journal of Chinese Language Teachers Association* 28.3:75-92.
- Huang, Chu-Ren. 1989. *Mandarin Chinese NP de: A Comparative Study of Current Grammatical Theories*. Special publication no. 93 of the Institute of History and Philology. Taipei: Academia Sinica.
- _____. 1993. Mandarin Chinese and the lexical mapping theory: A study of the interaction of morphology and argument changing. *Bulletin of the Institute of History and Philology, Academia Sinica*, 62.2:337-88.
- Huang, C-T. James. 1982. Move *wh* in a language without *wh*-movement. *The Linguistic Review* 1:369-416.

- _____. 1990. Modularity and explanation. *Interdisciplinary Approaches to Linguistics: Essays in Honor of Yuki Kuroda*, ed. by C. Georgopoulos and R. Ishihara, 305-322. Dordrecht: Kluwer Academic Publishers.
- _____. 1994. More on Chinese word order and parametrix theory. *Syntactic Theory and First Language Acquisition: Cross Linguistic Perspectives: Head, Projection, and Learnability*, ed. by Barbara Lust, Margarita Suner, and John Whitman, 1:15-36. Hillsdale, NJ: Lawrence Erlbaum.
- Huang, Shuanfan. 1994. Chinese and a metonymic language. *In Honor of William S-Y. Wang: Interdisciplinary Studies on Language and Language Change*, ed. by Matthew Y. Chen and Ovid J. L. Tzeng, 223-252. Taipei: Pyramid Press.
- _____. 1996. On the emergence of a definite article in spoken Chinese. Paper presented at NACCL-8. Urbana: the University of Illinois at Urbana-Champaign.
- Jackendoff, Ray. 1983. *Semantics and Cognition*. Cambridge: MIT Press.
- Kaplan, Ronald M., and Joan Bresnan. 1982. Lexical-functional grammar: A formal system for grammatical representation. *The Mental Representation of Grammatical Relations*. Cambridge: MIT Press.
- Kiparsky, Paul. 1978. Rule ordering. *Reading in Historical Phonology*, ed. by P. Baldi and R. Werth. University Park: Pennsylvania State University Press.
- Langacker, Ronald W. 1987. *Foundations of Cognitive Grammar, vol. 1: Theoretical Prerequisites*. Stanford: Stanford University Press.
- _____. 1995. Raising and transparency. *Language* 71:1-62.
- Li, Audrey Yen-hui. 1990. *Order and Constituency in Mandarin Chinese*. Dordrecht: Kluwer.
- Li, Charles N., and Sandra A. Thompson. 1981. *Mandarin Chinese: A Functional Reference Grammar*. Berkeley: University of California Press.
- Li, Y. C. 1980. The historical development of the coverb and coverbial phrase in Chinese. *Journal of Chinese Linguistics* 8:73-293.
- Lien, Chinfa. 1987. *Coexistent Tone Systems in Chinese Dialects*. Berkeley: University of California dissertation.
- _____. 1992. Lexical diffusion. *Encyclopedia of Language and Linguistics*, ed. by E. Asher, 4:2141-2144. Oxford: Pergamon Press.
- Mei, Tsu-lin. 1994. The grammar of T'ang-sung *koine* and the grammar of modern Chinese dialects. *Chinese Languages and Linguistics* 2:61-97.
- Oehrle, Richard, Emmon Bach, and Deirdre Wheeler. 1988. *Categorical Grammar and Natural Language Structures*. Boston: D. Reidel Publishing Co.
- Pinker, Steven. 1989. *Learnability and Cognition: The Acquisition of Argument Structure*. Cambridge: MIT Press.

- Pollard, Carl, and Ivan A. Sag. 1994. *Head-driven Phrase Structure Grammar*. Chicago and London: University of Chicago Press.
- Tai, James H-Y. 1973. Chinese as a SOV language. *CLS* 9:659-71. Chicago: Chicago Linguistic Society.
- _____. 1985. Temporal sequence and Chinese word order. *Iconicity in Syntax*, ed. by John Haiman, 49-72. Amsterdam: Benjamins.
- _____. 1989. Toward a cognition-based functional grammar of Chinese. *Functionalism and Chinese Grammar*, ed. James H-Y. Tai, and Frank F.S. Hsueh, 187-226. Chinese Language Teachers Association Monograph Series no. 1.
- Talmy, Leonard. 1985. Lexicalization patterns: Semantic structure in lexical forms. *Language Typology and Syntactic Descriptio: Grammatical Categories and the Lexicon*, ed. by Timothy Shopen, 3:57-149. Cambridge: Cambridge University Press.
- _____. 1988. Force dynamics in language and cognition. *Cognitive Science* 12:49-100.
- Tang, Ting-chi. 1991. Incorporation in Chinese syntax. *Tsing Hua Journal of Chinese Studies, New Series XXI*, 1-63, 335-376.
- Tsai, Wei-Tien Dylan. 1994. *On Economizing the Theory of A-bar Dependencies*. Cambridge: MIT dissertation.
- Tsao, F. F. 1990. *Sentence and Clause Structure in Chinese: A Functional Perspective*. Taipei: Student Book Co.
- Waldrop, M. Mitchell. 1992. *Complexity: The Emerging Science at the Edge of Order and Choas*. New York: Simon & Schuster.
- Wang, William S-Y. 1969. Competing Changes as a Cause of Residue. *Language* 45:9-25.
- _____. 1991. *Explorations in Language*. Taipei: Pyramid Press.
- _____, and Chinfa Lien. 1993. Bidirectional diffusion in sound change. *Historical Linguistics: Problems and Prospectives*, ed. by Charles Jones, 345-400. London: Longman Group.
- Zhang, Hongming. 1994. The grammaticalization of 'bei' in Chinese. *Chinese Languages and Linguistics* 2:321-360.
- Zhang, Min. 1995. Lexical perspective in syntactic change: Evidence from Mandarin dialects. Paper presented at ICCL-4/NACCL-7. Madison: University of Wisconsin.
- Yue-Hashimoto, Anne. 1993. The lexicon in syntactic change: Lexical diffusion in Chinese syntax. *Journal of Chinese Linguistics* 21.2:213-254.

Hsin-I Hsieh

[Received 1 December 1998; accepted 15 October 1999]

Department of East Asian Languages and Literatures
University of Hawai'i
354 Kekupua St.
Honolulu, HI 96825-2314
U.S.A.
hhsieh@hawaii.edu