What Is Language Complexity?

Ma Qinghua *
(Nanjing University, China)
Wang Xinxin **
(Nanjing Tech University, China)

Abstract: On the basis of complex system theory and the frontier researches of Chinese and foreign linguistics in this field, this paper systematically discusses the conceptual connotation of language complexity, and generalizes the essential characteristics in quality and quantity, which is non-linearity and high organizational depth. The non-linearity is mainly manifested as imbalance, emergence and interactivity, while the high organizational depth as multi-level, multi-dimension, multi-stage, high cardinality, etc. Generally speaking, the more complex the information is, the longer the minimum description length (MDL) of it is, and the higher resource/cost consumption is. In the relation of unity of opposites between complexity and simplicity, language complexity is absolute, while simplicity is relative. Complexity is always in dominance, while simplicity always serves as a lever to balance, control and regulate complexity. They are developing nonlinear adaptive interactions, which leads to the formation and development of linguistic paradigm system (such as grammatical paradigm and rhetorical paradigm). The conclusion provides a theoretical basis for the study of language complexity.

Keywords: language paradigm, complexity, non-linearity, high organizational depth, high cardinality, simplicity

1. Introduction

In syntax, semantics, rhetoric, and even the whole linguistic theory, complexity is often mentioned as an important theoretical concept. Domestically, Lü Shuxiang, at the end part of his representative work On the Analysis of Chinese Grammar (June 1979), proposed sentence complexity as an important grammar topic to be explored and the ultimate goal of grammar research. Guo Shaoyu also regards the complexity of Chinese grammar as an

---

* Dr. Ma Qinghua: Professor, School of Liberal Arts, Nanjing University, Nanjing, Jiangsu, 210093, China. E-mail: maqh86@nju.edu.cn.
** Dr. Wang Xinxin: Assistant Professor, School of Foreign Languages and Literature, Nanjing Tech University, Jiangsu, 211816, China. Email: xiaobai-00@163.com.
important part, as referred in his work *A New Probe into Chinese Grammar Rhetoric* (July 1979). The complexity of language is also a noteworthy topic in foreign linguistics today. Foreign scholars have focused on the complex existence, complex movement and complexity measurement of language from different research fields such as diachronic linguistics, linguistic typology, language contact, stylistics, second language acquisition, children's language development, language psychology, and computational linguistics. Many scholars have achieved some important results, such as Gibson (1998), Tomasello (2003), Kusters (2003), Dahl (2004), Larsen-Freeman (2008), Miestamo, Sinnemäki & Karlsson (2008), Ellis & Larsen-Freeman (2009), Givón (2009), Givón & Shibatani (2009), Sampson, Gil & Trudgill (2009), The Five Graces Group (2009)①, Culicover (2014).

Linguistics today is mostly confined to the perception of partial phenomena of complexity without enough attention paid to a systematic understanding of the nature of language complexity, so it is difficult to look at the various specific complexity phenomena under a unified understanding. Some scholars have complained that “no unified definition (about complexity) exists” (Sinnemäki, 2011:17), and it’s not even clear what exactly complexity means in linguistics (Gong & Coupé, 2011). Some scholars have lamented that complexity, as a nature of the real-world system, is difficult for any formalism to properly capture the whole spectrum of this property (Mikulecky, 2001:344), and it remains a vague concept.

The research status of complexity is not much better in other disciplines. John Horgan (1997:329) listed dozens of different definitions of “complexity” in the general sense or non-linguistic disciplines. Rescher (1998:1, 9) summarized the concept of complexity from a philosophical perspective, believing that complexity is, first and foremost, the number, type, and details of the related structure of the item components and their organization and operation. On this basis, the concept of complexity was classified. Karlsson, Miestamo & Sinnemäki (2008:VIII-IX) introduced this classification into linguistics and explained complexity linguistically. However, Rescher’s idea is not clear, because simplicity is also observed from these aspects. Thus, it does not reveal the nature of complexity.

If you don’t firstly figure out what complexity is, you can’t get the right language complexity measurement or profoundly reveal the intrinsic connection and complexification mechanism among language complexity phenomena. Studies concerning sentence complexity proposed by Lü Shuxiang (1979) cannot be smoothly carried out, either. Therefore, this paper intends to establish a distinctive and comprehensive cognition on the nature of complexity so as to provide an essential theoretical basis for research on

① Constituted by ten linguists including C. Beckner, R. Blythe, J. Bybee, M. H. Christiansen, W. Croft, N. C. Ellis, J. Holland, J. Ke, D. Larsen-Freeman, T. Schoenemann.
language complexity such as sentence complexity, through analysis and argumentation from perspectives of the qualitative and quantitative characteristics of the systematic complexities of language and the dialectical relationship between complexity and simplicity.

2. Qualitative characteristics

The characteristics of language complexity are as follows: uncertainty, incompleteness, hierarchy, sensitivity to initial conditions, dynamicity, non-additivity, non-linearity, fractality, instability, chaoticity, fuzziness, path dependency, situatedness, openness, adaptive, unboundedness, metastability, emergence, catastrophism, provisionality, self-organization, hetero-organization. However, their status is not equal. Non-linearity is the very core and most basic distinguishing one, while others are derived features or specific manifestations of it.

2.1 Distinctive features: Non-linearity

To understand what is complexity, you must first know what is simplicity. Linguists believe that linguistic surface sequence is linear, that is, one-dimensional. It’s a structural feature that is relative to hierarchy. System Theory also has a concept of linearity, yet the meaning of which is completely different.

From the System Theory, a simple system is a linear system, and the whole is the superposition of the subsystems, that is, the sum of parts. Linearity means that the input components are independent of each other during the process of becoming output components, and they are irrelevant. In contrast, complex systems are nonlinear systems. The whole has something that is not included in the sum of parts or in parts, and once it is broken down, its overall properties, state, function, etc., no longer exist, that is, the whole is greater than the sum of the parts (Xu Guozhi, 2000:9, 20-21, 203-205, 298; Auyang, 2002:5, 241).

Complexity depends more on the interaction and feedback between elements of a system, rather than the number of elements (Munné, 2013:176). Thus, any form of integration between the components of output expression can be considered as a contribution to non-linearity. In another dimension, the unpredictability of the output can promote non-linearity. When it is not fully predictable from the input, the output often has to depend on extraneous factors. The core of language complexity is non-linearity, which is a common denominator of various theories involved. This means that almost any combination of expression patterns in languages somewhat deviates from linearity (Dahl,
In the field of syntax, the meaning of construction is not a simple addition of the meaning of parts. Take the Chinese expression for example: the agent-patient relationship in (1a) cannot be considered as an inversion of (1b) by the usual analogy. It can only be understood that the agent-patient relationship of the two sentences has not changed. That is to say, the agent argument of both sentences is ｗǒ ‘I’ and the patient argument is ｎǐ ‘you’. (1a) is an alternative and more emphatic expression of (1b). In the same vein, when you say (2), the object ｓｈｉｔánɡ ‘canteen’ cannot be regarded as a patient argument, but as a place argument.

(1) a. ｎǐ ｘｉǎｎｇ ｓǐ ｗǒ ｌｅ．
   you miss die me PART.
   ‘you (make) me miss to the point of dying.’

   b. ｗǒ ｘｉǎｎｇ ｓǐ ｎǐ ｌｅ．
   I miss die you PART.
   ‘I very much miss you.’

(2) ｗǒ ｃｈī ｓｈｉｔａｎɡ．
   I eat canteen
   ‘I eat in the canteen.’

In fact, such phenomena are more common in the conventional structure, but people are not consciously aware of it. For example, all basic syntactic structures (such as subject-predicate structure, verb-object structure, modifier-head structure) except the coordinate structure respectively corresponds to a group of rather than a single structural semantic relationship. It can be seen that none of the meanings of these structures is the sum of the meanings of local components, otherwise they are all the same with the structure of coordination. However, even the coordinative structure corresponds to many different semantic relations. A coordinate phrase is formed by either conjunction or disjunction, and conjunctive coordination is divided into merger and intersection. Clauses are coordinated with at least 18 kinds of semantic relations, including superposition, mutual explanation, disorderly coexistence, disorderly simultaneity, delay, alternation, exchange, enumeration, contrast, supplement, auxiliary backing, orderly coexistence, orderly simultaneity, weak-successive, weak-progressive, weak-causal, premise, one-way interpretation. In terms of meaning, the first nine are disordered, and the last nine are orderly (Ma Qinghua, 2005:13-23). Similar phenomena are also common in the field of lexicology. For example, the actual meaning of idioms is not equal to its literal meaning. But maybe because of being too abstract or being entangled by another concept of “linearity” which is well known to linguistic scholars as the opposite of “hierarchy”, linguists still rarely ascend to the unified concept of “nonlinearity” to understand the nature of complexity.

2.2 Derived features

The linearity is balanced, reducible, and thus simple. In contrast, the non-linearity is
imbalance, non-reducible and thus complex. The three most important derived features of language’s ‘non-linearity’ are imbalance, emergence, and interactivity.

2.2.1 Imbalance

Imbalance is disequilibrium and asymmetry. Linguistic imbalance is embodied in corresponding relationship, organizational relationship, interactive relationship, and operating processes. At present, the discussion of imbalance or disequilibrium in linguistics is only focused on correspondences.  

A. Imbalance of correspondence. In the corresponding relationship, linguistic items, at the form-meaning-function level, are often manifested in many-to-one or one-to-one relations. Such as: (a) the imbalance of antonymy well-known in the academic world, the asymmetry (Shi Anshi & Zhan Renfeng, 1986; Shen Jiaxuan, 1999) in the process of tokenization or markerization, etc. (b) The intricate correspondence between Chinese word classes and functions (Zhu Dexi, 1985:4-5). (c) ambiguous structure, synonymous transformable structure, “heterogeneous isomorphism” proposed by Zhu Dexi (2010:129). As an example of the heterogeneous isomorphism, (3a) and (3b) are both verb-object structures, which have verb functions as a whole, but only some of the transformation forms, rather than all forms, have corresponding relations.

(3) a. mǎi piào  
   b. xǐhuān kàn  
   ‘buy tickets’    ‘like to watch’

The imbalance of the corresponding relationship deviates from “the idealized mapping method between the form and meaning”. On the one hand, it will increase the length of description of the system, and on the other hand, it will also reduce clarity, thereby increasing complexity (See Kusters, 2003; Dahl, 2004; Miestamo, 2008). Non-correspondence implies the contradiction with expectations, thus increasing complexity. Anderson (2012) believes that people will expect the relationship between different morphological syntactic representations and different vocabulary forms to be one-to-one correspondence. But what we found was not the case. The non-correspondence is an additional morphological complexity.

The imbalance of corresponding relationship does not only indicate its own complexity, but also has important positive significance. It is a prerequisite or the only reliable method for the evolution of complex systems. In the process of evolution, the linguistic complex system often makes corrections to existing systems, rather than creating a completely new
one (Schoenemann, 1999). If there are no unbalanced features of the corresponding relationship between form and meaning, all words will lose the tension of meanings, let alone the generality and elasticity of meanings, or a phonemic variant. The modelization and diversification of all structures (such as synonymous transformation) becomes impossible. Neither can a language develop nor can we know the origin of it.

B. Imbalance of organization. In organizational relations, coordinate structure is a syntactic or logical structure that is balanced and symmetric. This phenomenon coexists with other unbalanced and asymmetric syntactic structures, such as subject-predicate structure, verb-object structure, attributive-head structure, adverbial-head structure, head-complement structure, or logical structures such as various subordinative compound sentences, which are all unbalanced and asymmetric. But even in the coordinate structure, there is some imbalance in its phonetic rhythms, semantic features, and pragmatic features, which affects the word order of this structure (Ma Qinghua, 2006:176-199). The imbalance of dependence relationship is absolute. Balance is relative. Semantic selection and restriction has only one-way dependence under unlimited stimulus mode. For example, when talking about ㄇㄜ ‘baa, bleat’, we naturally think of 陽 ‘sheep’. But when talking about 陽 ‘sheep’, we may think of 肉 ‘meat’, 毛 ‘wool’, 草 ‘grass’, 奶 ‘milk’, 乳 ‘softness’, ㄇㄜ ‘baa, bleat’, etc., which is much more uncertain. Only in the stimulated mode with restrictions, there is a so-called two-way dependence. The onomatopoeia of the 陽 ‘sheep’ can only be ㄇㄜ ‘baa, bleat’. Conversely, when it comes to the subject of ㄇㄜ ‘baa, bleat’, you can only think of the 陽 ‘sheep’.

C. Imbalance of action. In the action relationship, any existence, operation and development in languages are necessary in motivation, well-conditioned in basis, and feasible in strategy. If presented through formula expression, it can be: Necessity/Motivation × Basis/Condition × Strategy/Means → Result (Ma Qinghua, 2012:194). And the motivation, conditions, methods and results are all unbalanced.

D. Imbalance of operation process. In the process of operation, the diachronic development of language is based on adaptability and emergence, and thus is irreversible. There are two important derived features coming from imbalance in the language: ordering and dynamicity.

A. Ordering. The ordering of linguistic system is, in discrete relationship, demonstrated as the various organizational sequences of the structure (for example, the increase of limitations on word order adds to complexity) and the sequence of action relationships. In a continuous relationship, it manifests itself in various continuum of existence and evolutionary continuum. The former includes characteristic continuums (such as ambiguity,
polysemy and vagueness), parts of speech continuum (such as multi-category words), etc., the latter includes shifting continuum, grammaticalization continuum and lexicalization continuum, etc. The continuum can create an unbalanced correspondence of form-meaning, or form-meaning-function, etc., but at the same time, it is conditional on the corresponding imbalance. Among them, the “one” of “many-to-one” becomes the intersection of the continuum, that is, the linked points of multiple items’ connection.

B. Dynamicity. Any variation as a dynamic feature breaks the balance, no matter it is in the language system’s diachronic vertical axis or in the synchronic horizontal axis. The diachronic dynamics of the language, in consequence, derives the characteristics of overall metastability and individual instability, which are characterized by the feature of path-dependence in terms of condition, and by order and continuity in terms of strategy. The synthesis of these derived features or performance characteristics constitutes adaptive features. Language’s adaptability also derives the feature of initial sensitivity. The initial sensitivity, in grammar performance, is manifested in the modifier-head structure where attributes (A) are placed in the front or back of head words (H). As far as SVO language in isolating language is concerned, the differentiation of its word order types is caused by the grammatical perspective differentiation in the initial adaptation process: If the object (O) is preferred to be transformed into the head word (or the core component) in the process of referentialization of a declarative structure, the AH is set to be the initial word order mode (see formula (a) shown below). If it is preferred to transform the subject into the head word (or the core component) in the process of referentialization of a declarative structure, the HA is set to be the initial word order mode (see formula (b) shown below). The structural prototype set by the initial mode becomes the basis of a series of secondary adaptation activities, determining the system differentiation of language types from this node, and stipulating the general direction of self-reproduction of formal systems (Ma Qinghua, 2014a).

(a) If L=SVO, and O→H is preferred in referentialization, then [SV]/[V]→A, and (S)V·O→AH
(b) If L=SVO, and S→H is preferred in referentialization, then [VO]/[V]→A, and S·V(O)→HA

In synchronic dimension, if there is no dynamicity, there is no grammar for wording and phrasing, nor is rhetorical activity consequently. In other words, there is no real language where words or phrases can be organized. The meaning of dynamic language operation is extremely rich. It includes: (a) the process of which can be languagization (Ma Qinghua, 2018a), structuralization, structural complexification, lexicalization, tokenization/markerization (e.g. grammaticalization, idiomization of multimarker cluster), subjectivization, etc.

① “Languagization” is different from “lingualization”. The former focuses on quasi-language evolving into real language, the latter focuses on speech expression, which is synonymous with “verbalization”.

Macrolinguistics (2019) 7
(b) the mechanism of which can be interactions (adaptation, suppression, compensation, variable control, resonance, etc.), integration, etc. (c) the system relations of which can be self-organization and hetero-organization (Ma Qinghua, 2012; 2005; 2006; 2014a; 2014b). Taking the mutual consultation mechanism in the interaction relationship as an example, under the principle of conversational cooperation, sentence legitimacy is often a tentative constant in the communicative contract. When the combination relationship deviates from the general mode, the recipient will not immediately decide it to be grammatically wrong, but will try to figure out the intention of the messenger based on a series of extralinguistic factors like common sense as background, similarity, expressive effect or communicative strategies, etc. from the other dimension (often an irrealis dimension), so as to get the meaning that matches the intent. Only when the efforts of transpositional understanding fails, the recipient will have to regard it as grammatically wrong. The messenger also takes advantage of this to implement the manipulation of language.

Take the combination relationship of modality markers as an example. You can’t say sentence (4) in declarative mood, but in the subjunctive mood which has unrealistic elements, you can say sentence (5). You can also say sentence (6), but under the declarative mood, you can’t say (7), which is acceptable in the counter-question (8) containing anti-reality elements. Subjunctive and counter-question both put the decoding of meaning under unrealistic and even anti-realistic dimensions to acquire acceptance, thus such sentences are in idiosyncracy structures (Ma Qinghua, 2017).

(4) * Tā bù gèng hǎo.
   he not more good
   ‘He is better, no.’

(5) Tā bù gèng hǎo zěnme xíng?
   he not more good how all right
   ‘How can he do it without being better?’

(6) Tā bù néng gàn zhè dāngzi shì.
   he not can do such Quantifier thing
   ‘He can’t do such a thing.’

(7) * Tā néng bù gàn zhè dāngzi shì.
   he can not do such Quantifier thing
   ‘As for not doing it, he can.’

(8) Tā néng bù gàn zhè dāngzi shì?
   he can not do such Quantifier thing
   ‘It is impossible for him not to do it.’

2.2.2 Emergence
Whole emergence refers to the system producing the characteristics of the whole which
isolated parts and the sum of them don’t have, and which are non-additive, unreducible, and unpredictable (Xu Guozhi, 2000:20, 21; Auyang, 2002:5, 182-186). Syntax can be seen as an emerging feature of semantic complexity (Schoenemann, 1999). In a sense, any grammatical construction is actually a complex symbolic structure (Langacker, 2004 [1987]:11). The meaning of the whole is more than the sum of that of parts. When different components make up an indivisible whole, or when there are interdependence, interaction, and inter-retroactive relationship between components or components and the whole, there is complexity in this whole (Morin, 1999:14. Quoted from from Bastardas-Boada, 2013:152).

Syntax is a kind of emerging feature of semantic complexity, which is not contradictory to the relative autonomy of syntax. The internal structure of language is the semantic structure, which indirectly affects the syntax. The grammar and syntactic rules between languages are sometimes similar because they have a universal semantic cognitive basis (Schoenemann, 1999), including universal cognitive content and cognitive style. There are three different levels of emergence in language: The first level is the emergence of hetero-organization of the extralinguistic world in language, including languagization and structuralization, etc. The second level is the self-organization of linguistic system, including grammaticalization, self-reproduction of the grammatical structure system, etc. The third level is both an additional factor of the self-organization system, and a higher level of hetero-organization emergence, including subjective or interlingual contact effects. Deacon (2006:111-150) also expressed similar views.

The decisive factor that forms the whole emergence of language is the incompleteness and situationality of language. Language is the symbolization of the extralinguistic world. A vital part in the semantic structure is relative information, so humans create simple ways to convey this information (Schoenemann, 1999). Situationality, subjective cognition, and efforts-saving are always accompanied with language expression and understanding, so all expressions don’t need to rely entirely on symbols themselves, which means all expressions are incomplete. In this sense, the incompleteness of language is absolute, while completeness is relative. All emergence is manifested in an incomplete form, which is why the whole can be larger than the sum of parts. Gestalt expressions such as omissions, implication, extra-linguistic meanings, entailment, presuppositions, metaphors, etc. depending on contexts and extra-linguistic information are included in varying degrees and ways. In the context of isolation, even if the sentence is complete, for example, (9) also contains the extra-linguistic meaning such as “Zhang San’s digestive system is normal today”. When a bus driver says sentence (10) to the passengers at a bus station, based on the incompleteness and situationality of language, the passengers would not interpret it into “the driver has to abandon the bus to leave”, but would know that the driver means to drive
the bus away, so the whole meaning of the sentence emerges.

(9) Zhāng Sān   jīntiān  zhōngwǔ  chī  le      liǎng  wǎn  mǐfàn.
    ‘Zhang San ate two bowls of rice at noon today.’

(10) Yǒu rén    xià chē   ma?       Méi rén   xià chē   wǒ    zǒu   le.
     ‘Does anyone get off? If No one gets off, I’m leaving.’

Based on the complement of situation or common sense, at the beginning of structuralization, language can still be a way of expression and can be decoded, in spite of the structural indeterminacy due to a low level of patterning. See the following examples of verb-object structure in the ancient Chinese oracle bone inscriptions (Ma Qinghua & Li Weizheng, 2017).

(V, O₁, O₂)

O₁=tool, O₂=object

(11) zhù             Dáyǐ        niú   (祝大乙牛)
     ‘to sacrifice cattle to Da Yi’

O₁=purpose, O₂=tool

(12) yù              sān    lǎo           Bǐ Gēng   (御三牢妣庚)
     ‘to sacrifice three captive cows to Bi Geng’

O₁=tool, O₂=purpose

(13) zhù             Yà Cì    zhì   (祝亚朿彘)
     ‘to sacrifice wild boar for Ya Ci’

O₁=purpose, O₂=object

(14) zòu             yù    qí          fā    (奏玉其伐)
     ‘to sacrifice jade for attacking others’

O₁=tool, O₂=purpose

(15) yù              què        Fù Yǐ  (御雀父乙)
     ‘to offer sacrifice to Fu Yi for Que’

---

① The oracle bone inscriptions, which can be traced back to 3300 years ago, are the earliest unearthed documents in Chinese.
What Is Language Complexity?

(O₁, O₂, V)

O₁=purpose, O₂=object

(16) yuē Fù Shū Mǔ zhū (曰勿鼠母祝)

Particle name god name sacrificial name

‘to offer sacrifice to Mu for Fu Shu’

Early childhood language also has a similar structural indeterminacy (Gui Shichun, 2000:150-151), which does not hinder the recipient’s decoding.

2.2.3 Interactivity

“Interactivity” embodies the original meaning of complexity, which can be traced back to the Latin word, Plexus, meaning “interwoven”. It means the interaction and interdependence between components and rules within a complex system. Language interactivity exists in various relations such as clustering, integration, change, and function. Thus, interactivity and dynamics constitute interactions. The complex system of language is made up of many variables interrelated and interacting. The change of one variable will affect all the other variables, and the result cannot be explained by decomposition, but by an analysis of the dynamics of interaction between all variables. Language’s self-organization covers such things as synergy (dependency, co-occurrence, etc.), conflicts (weighting under competitive relationships, differentiation and reorganization under exclusion, etc.), adaptation, integration, operations and other interactions (Ma Qinghua, 2014a; 2014b; 2006; 2017; 2018b).

Take the dependency as an example. One-way dependence such as concord (the form of its composition depends on the identity of another component) is also an obvious unbalanced relationship. For example, the Swedish words for the concept of “small” change in accordance to modified words.

(17) a. liten flicka (little girl, singular, female, ‘little girl’) b. litet barn (little child, singular, neutral, ‘little Child’

c. det lilla barnet (this little child, singular, neutral, ‘this child’) (Dahl, 2013:29)⁰

Two words with dependency can form direct component relationship or an indirect component relationship in the structure of level, which are both the calculation objects of the dependency distance. Schoenemann (1999) argues that grammar is organized in a hierarchical manner, and that non-hierarchy is not structurally dependent. Simultaneously, the evolution of thinking and semantic units should be prior to syntax and any verbal behavior. Only in this way, is it possible to expect the emergence of “structural dependence”. Because only when there are semantic units in the mind, the sentences can be constructed on the basis of it. Structural dependence limits the fission of semantic units,

⁰ According to Dahl (2013), the morphology changes are more obviously when it comes to “Små flickor (little girls, ‘the little girls’).”
thus limiting the amount of cognitive processing needed to decompose semantic units.

In the language microsystem, any interface must be affected by other interfaces. Take modern Chinese grammar as an example. The main phonological element that acts on it is the rhythm (including pauses, structural lengths (such as sound steps, etc.)). The conflict of rhythm between components can lead to the loss of legitimacy (Lü Shuxiang, 1963), such as:

\[(18)\] a. \([1+1]\) xuǎn kè \\
   select course  \\
   'course selection'  \\

\[(19)\] a. \([1+1]\) sào jīē \\
   sweep street  \\
   'sweep the street'  \\

\[(20)\] a. \([2+2]\) cānjiā duànliàn \\
   participate exercise  \\
   'participate in the exercise'  \\

\[(21)\] a. \([2+2]\) rì yì jīǎnshǎo \\
   day by day decrease  \\
   'decrease day by day'

Sememic grammar uses the two-way semantic selection restriction to reflect the interplay between semantics and grammar. Take the following two relationships as an example.

Joggle Relationship:
\[(22)\] a. chī fàn \\
   eat meal  \\
   'eat a meal'  \\

\[(23)\] a. àn xiàqù \\
   darken down  \\
   ('darken down')  \\

Concord Relationship:
\[(24)\] a. chī sīxiǎng \\
   eat thought  \\
   *'eat a thought'  \\

\[(25)\] a. àn xiàqù \\
   darken down  \\
   ('darken down')

\(\) The given word gives specific requirements for the characteristics of selectional restriction concept as if making a slot or blank and the sememe reflecting the requirement is represented by feminine sememe (marked by \([♀]\)). For example, the feminine sememe for \(chì\) 'eat' being \([♀ +mound]\)); the selectional restriction concept fills the slot or blank in compliance with the requirement and the sememe reflecting the filling requirement is represented by masculine sememe (marked by \([♂]\)), for example, the masculine sememe for \(zuǐ\) 'mouth' being \([♂ +mouth]\). (Ma Qinghua & Huang Yanqun, 2016)
What Is Language Complexity?

't dark down'  
\( {\text{'àn shàngqù}} \)  
darken up \( (\text{'àn: } [♂ +\text{reduction}]; \text{shàngqù: } [♂ -\text{reduction}]) \)

' dark up'  
(24) a. \( {\text{liàng qǐlái}} \)  
brighten up \( (\text{liàng: [♂ +\text{increase}]; qǐlái [♂ +\text{increase}]} \)

' light up'  
\( {\text{liàng xiàlái}} \)  
brighten down \( (\text{liàng: [♂ +\text{increase}]; xiàlái: [♂ -\text{increase}]} \)

' light down'  
Grammar is also affected by the pragmatic interface. Adjectives in an affirmative imperative sentence is of a weak quantity of degree (e.g. (25a)), while those in a negative imperative sentence is of an excessive adjectives (e.g. (26b)), and it’s not the other way around (such as ungrammatical imperatives (25b), (26a)). This is due to the role of the politeness principle. According to this principle, we should minimize the requirements for others and show maximum toleration when limiting others.

(25) a. \( {\text{Rènzhēn yīdiǎn.}} \)  
serious a little  
'Be serious!'  
(26)a. \( {\text{Tài cāícīn.}} \)  
too careless  
' Be too careless'

Grammar is also affected by extra-linguistic factors. For example, with the assistance of common sense, the sentence (27) will not be interpreted as an extraordinary match, so it does not affect normal decoding.

(27) \( {\text{Wǒ míngtiān fēi wūhán.}} \)  
I tomorrow fly Wuhan  
'I will fly to Wuhan tomorrow.'

3. Quantitative characteristics

3.1 Distinctive features: High organizational depth

Some scholars believe that, from a certain perspective, complexity is completely a qualitative concept (e.g. Munné, 2013:176). This view is not entirely appropriate, because quality and quantity are a pair of interdependent contradictions. High organizational depth is the quantitative feature of a complex system that distinguishes a simple system. For language, high organizational depth refers to multi-dimension, multi-level, multi-stage, etc. (Andrason, 2014). "Multi-level is an essential way to organize complex systems." (Xu
Different levels manifest as the emergence of different qualities. Therefore, the level is the emerging grade, and each time it emerges, a new level comes into being. The complex system itself gradually emerges through the integration of a series of intermediate levels. In theory, complex systems should have a structure of level, and the complexity of evolution often appears as the structure of level (Simon, 1987[1982]:188), because complexity needs to evolve from simplicity. Sampson (1978, 1979, 1980) also said that if language evolves from simple phrases to complex structures, then it has the hierarchical organization and the hierarchical structure’s evolution.

The grammatical rules of sentences appear in a hierarchical structure, and the syntax of simple sentences already has hierarchies, so it is also complicated (Givón, 2009:4). In the process of evolution, complex systems usually make corrections to existing systems, rather than creating a completely new system. Therefore, as the complexity of the system increases, the hierarchical approach and the evolution of hierarchical systems have also become more possible. Moreover, because language is fundamentally about meeting the need of communication in real life, it should be able to reflect the reality as accurately as possible, thus showing a similar hierarchy of some other aspects of the world and presenting a certain level of similarity in the language structure itself (Schoenemann, 1999).

Simple systems do not need to be hierarchical, or have only a hierarchy of superimposed relationships from elements to the whole. The emergence relation appears in the simple giant system, but it is only a direct emergence of the overall characteristics, without the need of and the condition for integration at the intermediate level. The whole emergence makes simple systems connect with complex systems to become an evolutionary continuum through simple giant systems. The interjection system is a quasi-language that has the characteristics of a simple giant system and belongs to the first-symbol system. The phonetic elements of an interjection can only be differentiated into three types according to their functions: (a) physiological elements, such as fricative consonants, aspirated consonants or inspiratory consonants which are all conducive to exhalation or intake; (b) physical elements, such as vowels or voiced consonants which ensure the loudness received by recipient; (c) elements that distinguish meanings, such as tones. The phonemes of the former two can be replaced broadly without causing a change in meaning. Unlike the phonetic elements of real language, phonemes of interjection are not the synthesis of physiological, physical and social attributes. (Ma Qinghua, 2011) In grammar, the words and sentences of interjection are indistinguishable, there is only grammatical components without grammatical structures. Interjection is the starting point of languagization and evolutionary continuum, it has a simple hierarchy resulting from evolution, which includes at most two levels. The lower level is emotional interjection based on the direct instinct
What Is Language Complexity?

reflections, such as the exclamation of crying or laughing. The higher level can be attitudinal interjection based on the indirect instinct reflections, such as: (a) praise/disrelish/regret/worry interjection (sucking interjection) ([ts·][ps·][ts·][ps·]), (b) reviling interjection (spit interjection)(pēi, 'pooh'; cuì, 'pshaw'), (c) negative interjection (crying interjection) (rising-falling tone), (d) respond/promise interjection (relied/affirmation/comprehended interjection (comfortable interjection)(low falling tone); or can be a complex emotional interjection formed by simple emotional interjections through composite operation, such as hayaya (strong surprise sounds)<ha (happy sound) + ya (surprise sound). (Ma Qinghua, 2012:4-8, 11-14, 20-22)

3.2 Symbiotic characteristics: High cardinality

A Simple system usually contains few subsystems, but this is only a reference standard (Xu Guozhi, 2000:203-205). A large number of elements alone do not necessarily constitute a complex system. In turn, complex systems necessarily contain a significant number of subsystems. Likewise, for a coordinate structure like a recipe, the increase of items cannot change the nature of its simplest syntactic structure. A relatively high cardinality makes language a typical complex system. High cardinality refers to a relatively high cardinality of the components contained in the system and the relationship or rules between the components (Andrason, 2014).

Language’s openness (such as the number of words in the core word class), unboundedness, fractal (such as recursion), infinite scalability, which are also complex properties related to high cardinality and high organizational depth.

3.3 Quantification of complex properties under epistemological model

The pair of contradictions of quality and quantity can also transform into each other. The complexity of quality must be manifested as a certain amount of representation and reflected through a variety of channels and ways. Two quantitative characterizations under epistemological model can be obtained when describing or applying complex properties.

3.3.1 Iconicity properties: The minimum description length

From the perspective of information theory, the simplicity or complexity of information can be described by the complexity degree of the symbol string that expresses the information. Based on the iconicity relationship, the qualitative complexity can be reflected in the positive correlation from its minimum description length, that is, the more complex

---

① [ ] = round lip, [ ] = inspiratory sound.
② The descriptive complexity (the length of the description necessary to fully describe the system), generative complexity (the length of a set of instructions necessary to provide a prescription for a system output), etc. listed by Rescher (1998: 1, 9), and the self-describing code length, algorithm information content, algorithm complexity, etc. listed by John Hogen (1997: 329) is similar in meaning or principle to the minimum descriptive length in reflecting the qualitative complexity.
the phenomenon is, the longer the minimum description length is. To measure the complexity of grammar, linguists often consider how much information is needed to fully clarify grammar rules (Andrason, 2014), that is, to quantify language complexity by information terminology, and to measure the complexity of the information content through Byte length.

The Byte length is the number of bits of the smallest program that can be accepted directly by the computer. $1 \text{ Byte} = 8 \text{ bits (bit)}$. Bit is a basic unit of information, referring to the choice of two events of equal probability. The minimum description length (hereinafter referred to as “description length”) is determined by Occam’s razor.\(^1\) DeGraff (2001:266) accordingly proposes the “bit complexity” of language. Scholars have expanded its connotations in a metaphorical way, referring to the numerous quantitative characteristics of language. As Mufwene (2013:198-199) pointed out, bit complexity reflects the number of units, and the number of rules in the lexicon, syntax, speech, or other modules in language. The constitutional complexity contains units and rules/constraints a language possesses. Some scholars also hold that the length of an expression is a direct reflection of its complexity (Sinnemäki, 2011). This kind of Bit complexity research is mostly based on the perspective of contrastive linguistics, so it naturally brings about the idea that the greater the number of unit storage of language and corresponding rules is, the higher the complexity is.

Frankly speaking, the length of a natural expression in a language is usually a direct reflection of the constitutional complexity (Sinnemäki, 2011). Because with the addition of each vocabulary item, the corresponding semantic properties are also increased, and any of these elements may interact, making the semantic structure to generate complex meaning relations (Riddle, 2008:141). However, the relationship between the length and complexity of natural language is still a little uncertain. There are at least two reasons: (a) A simple coordinative structure can be extended indefinitely, but a shorter anomaly structure may be more complicated than expected. (b) Natural language expressions have three types of section, i.e. normality, subnormality and super-normality. The dynamic relationship between complexity and simplicity performs fairly differently in different speech sections, which increases the uncertainty of the relationship between length and complexity.

In normal/regular section of expression, the corresponding form-meaning relationship is expressed as synchronized iconicity relationship, which means, the more complex the structure is, the more meaningful it is, and vice versa. This is seen in the expansion of normal/regular syntactic structures, such as $gāoxìng \ ‘happy’\rightarrow (28a)\rightarrow (28b)$. With the

\(^1\) It was proposed by the logician Occam in the 14th century. Its principle is that the simplest and most obvious should be adopted, and doing more is useless.
What Is Language Complexity?

expansion of the modality structure, the meaning has also been expanded accordingly. The iconicity relationship is also a simplifying model.

(28) a. hěn gāoxìng
very happy
b. zhēnde hěn gāoxìng
trully very happy
‘very happy’ ‘truly very happy’

In subnormal section of expression, as stimulated by the economy principle of language, the familiar concept is reduced in length, and the clarity of meaning is also reduced accordingly, maintaining some kind of iconicity relationship. On the other hand, its length can be reduced by special syntactic rules applicable only to names, but its meaning can only be understood with the spontaneous complement of extralinguistic familiar information. For example, the name of (29a) is idiomized from (29b). This kind of unmarked phrase obviously cannot be parsed as the SVO structure according to the Chinese syntactic rules.

(29) a. luóbò shāo ròu
radish cook meat
b. yòng luóbò zài yīqǐ shāo de ròu
with radish at together cook PART. meat
‘meat cooked together with radishes’ ‘meat cooked together with radishes’

In super-normal section of expression, the expressions in the artistic style that deviate various sememe syntactic rules (i.e. conflict with semantic selections and restrictions) have to retreat to the unrealistic dimension, that is, to the imaginary world so as to decode its meaning. Meaning breaks the constraints of concept, restoring into “volatile elements” (see also Beamer, 1982). So, it is possible to significantly increase the connotation of meaning without increasing the length of sentence, such as (30).

(30) Qīngfēng zài tā de jiāyuán lǐ jiāobàn fēnfāng.
autumn wind in her PART. homestead LOC. stir fragrance
‘The autumn wind is stirring the fragrance in her house.’

In contrast, the description of complexity in artificial languages is only carried out in the same section (i.e. section of normality/regularity), and the separation of object language and meta-language is realized. Therefore, the degree of complexity reflected by the minimum description length has a high degree of objectivity and certainty.

The application of minimum description length or bit complexity in the study of language complexity should be controlled within an effective range, that is, the rules in the system should be focused on. “Complete chaotic state is incompressible. Therefore, the description of it requires an infinite length. Information integration under chaotic conditions is not based on any rules. From an intuitive point of view, the complete chaotic state doesn’t have complexity.” (Miestamo, 2013) Patterning and regularization can “enhance the compressibility of an object”, which reduces the length of the description for

Macrolinguistics (2019)
Ma Qinghua & Wang Xinxin

itself and Kolmogorov complexity\(^\text{\textcircled{1}}\) (Dahl, 2013). Effective complexity (also known as “Gell-Mann complexity”) is relative to “Kolmogorov complexity”. It does not measure the length required to describe an object. Instead, it focuses on the length required for the description of a series of rules or structural forms. In language, the structural rules of the language system are reflected in the most compact description length. However, linguistic facts show that rules and chaos are not completely opposite. Rules can change to chaotic state with geometric progression or even exponential scale under some new dimensions and rules, which is indicated by the change in the form-meaning relationship of onomatopoeia in genetic linguistics. It is firstly depleted as onomatopoeic elements. Then, the motivation of relationship between form and meaning becomes concealed and chaotic due to the change of the place of articulation or pronunciation methods (Ma Qinghua, 2012:42-45). The law in-between should be equally tempting for language researchers. For example, the chaotic principles such as initial sensitivity have a profound impact on the formation of certain grammatical rule systems (Ma Qinghua, 2014a; see also Chapt. 2.2.1 “B. Dynamics”). In this sense, chaos can also be a manifestation of high complexity (Briggs & Pitt, 2008[1999]).

3.3.2 Relevant characteristics: Resource/cost consumption

Complex phenomena needs to handled with more resources or at a higher cost accordingly. Computational complexity means the amount of computer resources required to run an algorithm. Generally speaking, the larger the input scale is, the longer the calculation process is, and the more the resources spent on the calculation (time, memory, frequencies of visits to a particular database, etc.) are. If the minimum description length is a measure based on its similarity relationship with complexity, then the resource/cost consumption volume, as another measure of complexity, is acquired on the basis of its causal relationship with complexity. Based on similar ideas, Gibson (1998, 2000) proposed Dependency Locality Theory, which is a theory of computational resource consumption about syntactic parsing, believing that language understanding needs to consume two kinds of computing resources: storage and integration. The longer the component’s dependence distance is, the greater the cost of structural integration and storage is, and the more efforts it takes to process. Therefore, the syntactic complexity also increases. Gibson & Ko (1998), Gibson (2000) and Oya (2011) all proved with empirical evidence that the sentences with shorter dependence distance lead to shorter reading time, and the sentences with longer dependence distance lead to longer reading time, and the number of words in the sentence has an impact on the average dependent distance, with the latter increasing as the number

\(^{\text{1}}\) That is the complexity of algorithm proposed by Kolmogorov et al. in the mid-1960s. It looks for the smallest program or instruction to describe the complexity of a given sequence of symbol strings.
What Is Language Complexity?

of words grows.

Under this indicator, scholars naturally regard difficulty, obscurity and non-essentiality as complex language phenomena. Here are the details: (a) Acquisition difficulty is regarded as a subjective category of language complexity in second language researches (Maitz & Németh, 2014). However, they do not strictly correspond, because the similarity between the target language and the mother tongue also determines the difficulty of acquisition. Dahl (2004:2) and Sinnemäki (2011) and other scholars pointed out that the complexity and difficulty of language should be clearly separated. (b) It is considered that the opacity of morphological form is a form of morphological complexity (Fenk-Oczlon & Fenk, 2008:46). The word motivation is the reason for the form-meaning matching relationship of words. Objectively, it does not change according to time, while subjectively, it can increase cognitive difficulty due to diachronic abrasion. Exploring its lost content requires mobilization of more linguistic or extralinguistic resources. The reading on ancient language also faces the same situation. Anderson (2012) also repeatedly stressed that “it is difficult to understand how people promote language evolution in a morphological manner and showcase this organization to the fullest.” “There is a lack of strong evidence for motivation of morphogenesis.” “For morpheme allomorph, it’s possible to find the ultimate cause of its formation in the characteristics of the expression system.” It is clear that he also considers the obscurity of motivation as a manifestation of increased complexity. (c) Anderson (2012) argues that non-essentiality creates additional complexity. Redundancy and additionality are all the manifestations of non-essentiality. The addition of non-essentiality often means the increase of difficulty. For example, from the performance of patients with aphasia, the greater the number of arguments is, the more difficult it is to process. It is more difficult to process dispensable arguments than obligatory arguments. (Zhou Tongquan, Shu Hua & Ke Youping, 2010) Also, it’s associated with non-essentiality. Some scholars regard the complex features without increasing the effectiveness of communication in a language system as “ornamentally complex” features. For example, be can be used as auxiliary verb have, to form perfect tense, e.g. (31). It’s a way of complexification that does not improve semantics or pragmatics (Kortmann & Szmrecsanyi, 2009).

(31) They’re not left school yet.

Quantitative complexity reflects the number of relevant variables of the language macrosystem and microsystem. In language microsystem, the relevant variables include modules and module subsystems (up to the terminal components, such as thesaurus), rules, etc. Language is an integration of physical, biological and social/cultural factors, and it is therefore seen as an example of a complex system of reality. Similar to other complex systems in the real world, language has a higher cardinality, which makes language a
Ma Qinghua & Wang Xinxin

A typical complex system. “High cardinality” refers to the large number or large scale of sets of components in a system and their relationships or rules (Andrason, 2014). At present, linguists generally believe that the more language structure units, rules and expressive means, the higher the complexity is (Sinnemäki, 2011). Quantitative complexity has clear and specific indicators and connotations. “Degree of complexity” is often used as a broad term, which has a subtle difference from “quantitative complexity”.

The qualitative complexity and quantitative complexity constitute a general interpretation and the most general classification of language complexity. Mufwene (2013:198-199) considered that the “bit complexity” is opposite to “interactional complexity” in language, with the latter referring to the complexity of interactions within or between language units, relationships, rules, or other elements. But this cannot be regarded as sufficiently comprehensive or optimized classification of complexity, because interactivity is only a partial attributive feature of complexity, and bit complexity is an external characterization of quantitative complexity, which can’t replace the internal characterization. The connotation of complexity involves many characteristics, which lie at different levels and have interactivity, indicating that the concept of complexity itself is complex.

4. Dialectical relationship between complexity and simplicity

Further understanding of the nature of complexity can be obtained from the relationship between complexity and simplicity. Many concepts related to language complexity imply this pair of relationships. Complexity and simplicity, as the contradiction of dialectical unity, is a very far-reaching issue (Kortmann & Szmrecsanyi, 2009). The following two aspects of their relationship are most noteworthy.

4.1 Absolute and relative

For language, a complex giant system, complexity is its fundamental attribute, while simplicity is always relative or conditional. The so-called simplicity is only relative to more complex situations, which is often confined to a single interface of language. To put it more clearly, language complexity is absolute, while simplicity is relative.

In terms of complexity, any element in the language system is not a simple item. Each item can be differentiated into several interfaces, such as phonetics, grammar, semantics, pragmatics, etc. Each interface can be decomposed independently. Consequently, the smallest constituent units on the respective interfaces, such as phoneme, morphemes, sememe, etc., are obtained. Then we can first get the simplest and most basic structure by integrating them. Both the smallest unit and the simplest structure have multiple facets. Every phoneme reflects the complex value information under the opposite and complementary relationship. De Saussure (1980[1916]:169, 223) said that language can be
What Is Language Complexity?

regarded as an algebra with only complex items. No matter from which aspect you start the research, you can’t find anything simple; the balance of the elements that are mutually constrained happens anytime and anywhere. He cites the two German nouns *Nacht* ‘(s.) night’ and *Nächte* ‘(p.) night’ as an example. The main vowel of the singular form is *a*, the main vowel of the plural form is *ä*. *Nacht* doesn’t have the ending *e*, and the plural form *Nächte* does. The plural form has an umlaut while the singular one doesn’t. (The pronunciation of *ch* changes from /ʃ/ to /ç/). It shows that they are not isolated symbols, and they create various oppositions in different systemic relationships. Overall, they are not simple items but complex ones. As the simplest syntactic structure, coordinate structure is grammatically balanced and disordered. It always achieves ordering from other interfaces such as semantic, pragmatic, or phonetic interface (Ma Qinghua, 2006:176-199). It can be seen that linguistic simplicity is relative and complexity is absolute. From a technical perspective, complexity can also be relative.\(^1\) It is in this sense that complexification and degree of complexity are proposed. Complexification means an increase in the degree of complexity. Therefore, there are low complexity and high complexity. The structural complexity of the two-word sentences independent of contexts is relatively low. The so-called complex sentences are actually highly complex in nature. In typology, the following three indicators are often used to judge whether the relevant aspects of language are highly complex: (a) over specification, that is, the extreme specialization of the semantic category; (b) structural elaboration, that is, the extreme specialization of the structural rules; (c) Irregularity (McWhorter, 2008:167-168).

4.2 Dynamic acting relationship

For a language, sufficient yet limited paradigms are needed to express infinitely rich content, so complexity and simplicity are always the unity of opposites and unbalanced in status as well. Complexity is absolutely in dominance, while simplicity is always a lever for balancing, controlling, and regulating complexity. In most cases, it seems to play an artful role.

4.2.1 Interaction of different types of complexity

Culicover (2014) distinguished between formal complexity and processing complexity in grammatical complexity. The former is objective complexity, which is directed at the general description of grammar and can be shown from grammar itself. The latter is subjective/cognitive complexity, which aims at the computing resource needed for form-meaning projection of language users. It may not always be manifested itself from grammar. For example, the relational clause extracted the object seems, in terms of form,\(^1\) Miestamo (2008:24), Szmrecsnay & Kortmann (2012:10-12) have also used “absolute complexity” and “relative complexity”, but their actual meanings are objective complexity and subjective complexity. Subjective complexity is the complexity of processing (see Chapt. 4.2.1).
as complex as the relational clause extracted the subject (e.g. (32a-b)), but psychological studies have shown that perhaps due to our limited memory, the former is more difficult to handle (Gordon & Lowder, 2012). The term “formal complexity” is not as good as “paradigm complexity”, because the former is limited to the grammatical static aspect, yet the latter, in addition to this aspect, also includes the static aspects of other modules and the dynamic mode of operation of language. There is actually another type of complexity, that is, the meaning complexity of the content to be expressed. Meaning complexity, paradigm complexity, and processing complexity are direct mappings of semantic triangles (see Figure 1). Only paradigm complexity can reflect the “effective complexity” of language itself (see Chapt. 3.3.1). Therefore, it is the core of linguistic studies. But if you set aside the interaction with the other two kinds of complexity, it is impossible to know the origin of complexity in essence.

(32) a. The senator that the reporter bothered Δ caused a big scandal.
   b. The senator that Δ bothered the reporter caused a big scandal.

Paradigm complexity is the response to the complexity of meaning, and in many cases, the two appear to be in the synchronized iconicity relationship. Mufwene (2013:209) notes that the grammatical complexity of any language is enhanced by the pressure of including more explicit information in the discourse. For example, if a new verb or noun is created with a similar or related noun or verb, it will increase the calculation of language, but from another perspective, it will also reduce the pressure of remembering more different symbols. It can be seen that in order to meet the needs of expressing infinite and complex contents, grammar should be simplified in many ways and rely on a few rules or principles to calculate new meanings.

A. Reduction of meaning complexity

From the effects of paradigm complexity on meaning complexity, fuzziness is the normal state of language complexity, and discarding fuzziness will bring about discontinuity and cause information loss, but fuzziness or vagueness can be reduced through delimitation.

B. Reduction of paradigm complexity

Then look at the interaction of paradigm complexity with meaning complexity or processing complexity. (a) Humans have the ability to use daily accumulated information as a complement to understand language. This makes the words with high frequency and
short length (such as monosyllabic words) which is of low complexity still maintain high complexity in semantics (Fenk-Oczlon & Fenk, 2008:58-60). (b) As we have the ability to understand by insight, the complexity of meaning is relatively easy to handle, especially the closer it is to intuition, the lower the difficulty of understanding is. So, it’s possible to trade the meaning complexity for paradigm simplicity. For example, we often use expressions of multi-role arguments with high semantic complexity to replace expressions of single-role arguments with low semantic complexity, thus reducing the length and redundancy of the entire sentence. For example, sentence (33) with a linear pattern of “NP1 + VP1 + NP2 + VP2” (see Figure 2) is the compression of three expressions (i.e. (34a), (34b), and (34c)). The storage capacity of (33) doesn’t decrease. But the semantic complexity has to be increased to support the reduction of redundancy and sentence length.

Figure 2. Schematic Diagram of Structural Semantic Relationship

(33)  Wǒ  dài   Xiǎo Wáng  shàngjiē.
     I  bring  Xiao Wang   go to the streets
     ‘I will bring Xiao Wang to the streets’

(34) a. Wǒ  dài   Xiǎo Wáng  .   b. Xiǎo Wáng  shàngjiē.  c. Wǒ  shàngjiē.
     I  bring  Xiao Wang      Xiao Wang  go to the streets       I  go to the streets
     ‘I will bring Xiao Wang’     ‘Xiao Wang will go to the streets’   ‘I will go to the streets’

C. Reduction of processing complexity

Some of the paradigm complexity are improved, but instead the processing complexity is reduced. Firstly, the hierarchy, chunking and schemata in information processing, implying an improvement in processing automation, which reduces the efforts required (Givón, 2009:287). In fact, all modelization (including regularization, stylization) can be analogized and can strengthen the compressibility of complex objects and improve processing automation. As a result, processing complexity will be reduced. Taking fractals as an example, fractals in language includes stratification and cross-hierarchy. Stratification creates and adds hierarchies, so it is a complicated operation. In contrast, self-replication in cross-layers is a simplified operation. For example, although the word formation, phrase

---

Paradigm complexity and processing complexity can also have positive correlations, such as irregularities, exceptions, and syntactic idiosyncrasy, which have contributed to a higher degree of paradigm complexity. They are often the reasonable result of the complex effects of multiple rules and forces (such as Chinese clutch words) (Ma Qinghua, 2012:164-189), which reduce the predictability and regularity of grammar, thus improving processing complexity.

Macrolinguistics (2019)
construction rules and sentence structure rules in Chinese belong to different structural levels, they basically adopt the same set of rules (Zhu Dexi, 1985:7-9). Recursion is a kind of dynamic fractal operation. From the perspective of the overlapping of syntactic structures caused by its repeated operations, recursion is objectively a complicating phenomenon which increases the complexity of the paradigm. However, from the perspective of using the same rule and maintaining the same function, recursion is also a simplistic phenomenon. In Givón’s view, hierarchy also reduces the subjective processing complexity. Iconicity and image schemas are also an effective mode to reduce processing complexity. Secondly, markedness is another strategy to reduce processing complexity. According to Givón (2001:330), there are three complexity indicators about markedness: (a) The marked case is more complex; (b) The marked case is less frequent; (c) The marked category is harder to process. Relative to unmarkedness, they are admittedly reasonable. But they fail to reveal the nature of markedness. A marker is a description of a structural relationship or a structural meaning. It is a meta-language. The structural relationship or structural meaning indicated by markers is the object language. The semantic motivation of a marker is often permeated by the folk view of syntactic relations or syntactic meaning (Ma Qinghua. 2007).① The object language and meta-language differentiation caused by markedness and markerization also objectively increase the syntactic item and the level of meaning, which is a complexification phenomenon. But from the perspective of reducing the integration cost of syntactic structure, it is also a simplification of subjective information processing. Again, according to the theory of marked value, the higher the degree of mark of grammatical category is, the more accurate and rapid the decoding is.

4.2.2 Internal interaction of paradigm complexity and complexification of paradigm system

From the internal interaction of paradigm complexity, the following three points can be found. (a) To achieve the simplification of paradigm in one dimension, it is often necessary to be exchanged with the complexification of paradigm in another dimension. For example, repeated use can lead to phonetic fusion and reduction (The Five Graces Group, 2009). Pronouns and omissions are also strategies for adapting referential repetition. Therefore, the elimination of paradigm repetition is achieved by adding other paradigms. Another example is that the complexification of the attributive can reduce the topic conversion, make the sentence structure compact and concise, and the content rich and condensed (Long Youzhen, 2008). It also avoids the loss of meaning rigor due to the omission strategy which leads to over-reliance on context for interpretation, and avoids the excessive

① For example, Chinese coordinative conjunctions usually originate from words with one of the three non-associative meanings: together, equivalent, or adjacent. The folk view of coordinative relationship hided in the three meanings is in line with that of the thought in grammar theory (e.g. “Coordinative phrases are connected equally by components with the same function”). (Ma Qinghua, 2006:69; 2017)
What Is Language Complexity?

redundancy caused by repeating strategies. (b) From the perspective of the system as a whole, there is sometimes a compensation relationship between the simplicity and complexity of paradigm. For example, the morphologically complex language is poor in function words and flexible in order, while the morphologically simple language is abundant in function words and its structure is of high ordering. Obviously, the latter has to use the complexity of the lexical and syntactic structures to compensate or balance the simplification of the morphological and lexical structures and vice versa. Another example is that in ancient Chinese, parts of speech are used flexibly to compensate the simplicity of lexical system and syntactic structures. In modern Chinese, with much richer vocabulary and much more complicated syntactic structure, the flexible use of parts of speech is greatly reduced. (c) As mentioned earlier, simplicity is always the lever to balance, control, and adjust complexity. When the diachronic development of language makes the system more complex and difficult to control, it will lead to different degrees of catastrophe, and stimulate system restructuring so as to reduce the complexity of the system. Anderson (2012) discovered it from the morphological point of view. Wang Li (1980:342) also found that in the middle stage of the development of ancient Chinese with monosyllabic predominance, the phoneme system becomes overly complex, causing the transition to disyllable. “Disyllabilization is one of the means of reducing homonyms”, which is of reciprocal causation with “simplification of sound system”.

In the history of language development, the complexification of the paradigm system is the trend. Language system maturation is a diachronic process that increases the complexity of the system in a nonlinear manner. Overall, any degree of simplification in this process will not surpass the trend of complexification (Dahl, 2004:11). “Syntax is more properly understood as an emergent characteristic of the explosion of semantic complexity that occurred during hominid evolution” (Schoenemann, 1999). In other words, the increase in the complexity of concept understanding for a hominid has led to a significant increase in the need for grammar to realize effective communication, which initiated the process of structuralization, advancing from the chaotic concept of things to clear propositions. The addition of declarative multiple proposition information and surface additional components (such as quantity components, modality components) and deep additional components (optional arguments) led to grammatical advancement and initiated the process of structural complexity. At the supra-sentence or sentence group level, when parataxis is insufficient to be used to organize logical structure, associative marker as an explicit means of connection is constructed. Therefore, from a macro perspective, if every generation naturally acquire the language, then the grammar will remain at a higher level of complexity as time progresses, without the overall simplification (McWhorter, 2008:167-168). On the whole, the process of mother tongue acquisition is also manifested

Macrolinguistics (2019)
Ma Qinghua & Wang Xinxin

as an increase in language complexity.

In the field of rhetoric of the language system, it’s more about the reinforced expression (such as the extraordinary expression of the imaginary and emotional world in an idiosyncratic way), which greatly enhances the complexity of paradigm and meaning. But on the other hand, the rhetorical paradigm (such as the rhetoric) can be modeled, so compared to the extremely complex contents expressed, there is still a simplification. The decline of rhetorical reinforcement leads to grammaticalization and grammatical regularization. It is in this sense that the syntax is seen as an emergent feature of semantic complexity (see Chapt. 2.2.2). Dahl (2004:122) also believes that a mature grammatical paradigm is the result of rhetorical devaluation. Liu Dawei (2010) expressed similar points of views from the perspective of continuum relationship.

5. Conclusion

This paper, based on the principle of complex system and the frontier researches of Chinese and foreign linguistics in this field, systematically discusses the conceptual connotation of language complexity, and generalizes the characteristics of language complexity in terms of quality and quantity, and its main derivative or performance features. Through in-depth discussion of the dialectical relationship and the interactive relationship between complexity and simplicity, as well as the modes and results of the interaction, it further deepened the understanding of the nature of language complexity. Conclusions are as follows: (a) The essential characteristics of the language complex giant system are non-linearity and high organizational depth, with the former mainly manifested as imbalance, emergence and interaction and the latter mainly manifested as multi-level, multi-dimension, multi-stage and high cardinality, etc. Generally speaking, the higher the complexity is, the longer the minimum description length of the information is, and the greater the resource/cost consumption is. (b) In the relation of unity of opposites between language complexity and simplicity, language complexity is absolute while simplicity is relative, and complexity is always in dominance while simplicity is always a lever for balancing, controlling, and adjusting complexity. The two constantly carry out unbalanced adaptive interactions, leading to the formation and development of grammar system. Other language paradigms such as the rhetorical paradigm also conform to similar principles. The conclusions of this paper provide a solid theoretical basis for researches concerning language complexity and help to solve the mystery of how language uses the lever of simplicity to control complexity under the principle of opposing unity.

References
Anderson, S. R. 2012. Dimensions of Morphological Complexity [A]. In M. Baerman, D. Brown & G. G.
What Is Language Complexity?

Corbett (eds.). 2015. Understanding and Measuring Morphological Complexity [C]. Oxford: Oxford University Press, pp. 11-28.

Andrason, A. 2014. Language Complexity: An insight from complex-system theory [J]. International Journal of Language and Linguistics, 2(2):74-89.

Auyang, S. Y. 2002. Foundations of Complex System Theories [M]. Chinese Translation. Shanghai: Shanghai Technology and Education Press.

Bastardas-Boada, A. 2013. General Linguistics and Communication Sciences: Sociocomplexity as an integrative perspective [A]. In A. Massip-Bonet, A. Bastardas-Boada (eds.). Understanding Complex System Complexity Perspective on Language, Communication and Society [C]. Heidelberg: Springer-Verlag, 11:151-174.

Beamer, M. 1982. The Representation of the Spirit and the Expression of the Inexpressible [A]. Aesthetics [C]. Vol.2. Beijing: China Social Science Press.

Briggs, J. & Peat, F. D. 2008 [1999]. Seven Life Lessons of Chaos [M]. Chinese Translation. Shanghai: Shanghai Century Publishing (Group) Co., Ltd.

Culicover, P. 2014. Construction, Complexity, and Word Order Variation [A]. In F. J. Newmeyer & L. B. Preston (eds.). Measuring Grammatical Complexity [C]. Oxford: Oxford University Press, 148-178.

Dahl, Ö. 2004. The Growth and Maintenance of Linguistic Complexity [M]. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Dahl, Ö. 2013. Lectures on Linguistic Complexity [R]. Stockholm University.

De Saussure, F. 1980 [1916]. Course in General Linguistics [M]. Chinese Translation. Beijing: The Commercial Press.

Deacon, T. W. 2006. Emergence: The hole at the wheel’s hub [A]. In P. Clayton, P. Davies (eds.). The Re-Emergence of Emergence: The emergentist hypothesis from science to religion [C]. Oxford: Oxford University Press, 111-150.

DeGraff, M. 2001. On the Origin of Creoles: A Cartesian critique of neo-Darwinian linguistics [J]. Linguistic Typology, 5(2/3):213–310.

Ellis, N. C. & D. Larsen-Freeman. 2009. Constructing a Second Language: Analyses and computational simulations of the emergence of linguistic constructions from usage language as a complex system [J]. Language Learning, 59(s1):90-125.

Fenk-Oczlon, G. & A. Fenk. 2008. Complexity Trade-Offs between the Subsystems of Language [A]. In M. Miestamo, K. Simmennäki & F. Karlsson (eds.). Language Complexity: Typology, contact, change [C]. Amsterdam/Philadelphia: John Benjamins Publishing Company, 43-65.

Gibson, E. & K. Ko. 1998. A Integration-Based Theory of Computational Resources in Sentence Comprehension [G]. Paper presented at the Fourth Architectures and Mechanisms in Language Processing Conference, University of Freiburg, Germany.

Gibson, E. 1998. Linguistic Complexity: Locality of syntactic dependencies [J]. Cognition, 68(1):1-76.

Gibson, E. 2000. The Dependence Locality Theory: A Distance-based theory of linguistic complexity [A]. In Y. Miyashita, A. P. Marantz & W. O’Neil (eds.). Image, Language, Brain [C]. Cambridge, MA: MIT Press, 95-126.

Givón, T. & M. Shibatani (eds.). 2009. Syntactic Complexity: Diachrony, acquisition, neuro-cognition, evolution [C]. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Givón, T. 2001. Syntax: An introduction [M]. Vol.1. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Givón, T. 2009. The Genesis of Syntactic Complexity: Diachrony, ontogeny, neuro cognition, evolution [M]. John Benjamins.

Gong, T. & C. Coupé. 2011. A Report on the Workshop on Complexity in Language: Developmental and evolutionary perspectives [J]. Biolinguistics, 5(4):370-380.

Gordon, P. C. & M. W. Lowder. 2012. Complex Sentence Processing: A review of theoretical perspectives on the comprehension of relative clauses [J]. Language and Linguistics Compass, 6/7:403-415.

Gui Shichun. 2000. Psycholinguistics [M]. Shanghai: Shanghai Foreign Language Education Press.

Guo Shaoyu. 1979. New Exploration on Chinese Grammar and Rhetoric [M]. Beijing: The Commercial Press.
Horgan, J. 1997. *The End of Science* [M]. Chinese Translation. Hohhot: Yuanfang Press.
Karlsson, F., M. Miestamo & K. Sinnemäki. 2008. Introduction: The problem of language complexity [A]. In M. Miestamo, K. Sinnemäki, F. Karlsson (eds.). *Language Complexity: Typology, contact, change* [C]. Amsterdam/Philadelphia: John Benjamins Publishing Company, VIII-IX
Kortmann, B. & B. Szmarcianski. 2009. World Englishes between Simplification and Complexification [G]. In *World Englishes-Problems, Properties and Prospects*. Selected papers from the 13th IAWE conference, 265-285.
Kusters, C. W. 2003. *Linguistic Complexity*: The influence of social change on verbal inflection [D]. PhD Dissertation, University of Leiden.
Larsen-Freeman, D. 2008. *Complex Systems and Applied Linguistics* [M]. Oxford: Oxford University Press.
Liu Dawei. 2010. From Grammatical Construction to Rhetoric Construction [J]. *Contemporary Rhetoric*, 3/4.
Long Youzhen. 2008. The Generative Forms of Attribution [J]. *Journal of Hubei Normal University (Philosophy and Social Science)*, 2(5):55-58.
Lü Shuxiang. 1963. Preliminary Study on the Question of Chinese Monosyllable and Bi-syllable [J]. *Study of Chinese Language*, (1):10-22.
Lü Shuxiang. 1979. *The Problems of Chinese Grammar Analysis* [M]. Beijing: The Commercial Press.
Ma Qinghua. 2005. On Self-organization of Coordinate Structure [M]. Shanghai: Fudan University Press.
Ma Qinghua. 2006. *Multidimensional Research on Semantics* [M]. Beijing: Language & Culture Press.
Ma Qinghua. 2007. The Path of Grammaticalization for Coordinate Constructions and its Universality [J]. *Macrolinguistics*, 1(1):82-104.
Ma Qinghua. 2011. A Study toward the Primitiveness of the Relationship between Interjection's Form and Meaning [J]. *Science of Language*, 10(5):482-496.
Ma Qinghua. 2012. *Linguistics Topics under System Theories* [M]. Shanghai: Shanghai People's Publishing House.
Ma Qinghua. 2014a. Self-reproduction of Syntax System from the Perspective of Adaptive Principles [J]. *Linguistic Researches*, 7:1-14.
Ma Qinghua. 2014b. The Integration of Grammatical Structure under the Principle of Identity [J]. *Journal of Shanxi University (Philosophy & Social Science)*, 37(4):50-59.
Ma Qinghua & Huang Yanqun. 2016. The Determinacy of Lexical Selection Restriction and Its Basis for Existence [J]. *Macrolinguistics*, 4(1):17-36.
Ma Qinghua. 2017. The Integration and Variation of Modality Governance Structure in Chinese [J]. *Journal of Shanxi University (Philosophy & Social Science)*, 40(1):74-84.
Ma Qinghua & Li Weizheng. 2017. On the Modeling and Development of Verb-object Construction from the Oracle Bone Inscriptions to Jinwen Shangshu [J]. *Macrolinguistics*, 5(2):1-37.
Ma Qinghua. 2018a. The Significance of Onomatopoeia in Languagization: From the perspective of sound-meaning relationship under dynamic system principle [J]. *Macrolinguistics*, 6(2):43-62.
Ma Qinghua. 2018b. Synergy of Chinese Modality and Its Complexity [J]. *Academic exchanges*, 6(6):126-131.
Maitz, P. & A. Németh. 2014. Language Contact and Morphosyntactic Complexity: Evidence from German [J]. *Journal of Germanic Linguistics*, 26(1):1-29.
McWhorter, J. 2008. Why Does a Language Undress? Strange cases in Indonesia [A]. In M. Miestamo, K. Sinnemäki & F. Karlsson (eds.). *Language Complexity: Typology, contact, change* [C]. Amsterdam/Philadelphia: John Benjamins Publishing Company, 167-190.
Miestamo, M. 2008. Grammatical complexity in a cross-linguistic perspective [A]. In M. Miestamo, K. Sinnemäki & F. Karlsson (eds.). *Language Complexity: Typology, contact, change* [C]. Amsterdam/Philadelphia: John Benjamins Publishing Company, 23-42.
Miestamo, M. 2013. *Language Complexity: Typological perspectives, Guest Lecture* [R]. Stockholm University, ZeFIS Kolloquium, Wuppertal, Jan 30.
Mikulecky, D. C. 2001. The Emergence of Complexity: Science coming of age or science growing old [J].
What Is Language Complexity?

Computers & Chemistry, 25(4):341-348.
Mithun, M. 2009. Re(e)volving Complexity Adding Intonation [A]. In T. Givón (eds.). Syntactic Complexity Diachrony, Acquisition, Neuro-Cognition, Evolution [C]. Amsterdam/Philadelphia: John Benjamins Publishing Company, 53-80.
Mufwene, S. S. 2013. The Emergence of Complexity in Language: An evolutionary perspective [A]. In A. Massip-Bonet & A. Bastardas-Boada (eds.). 2013. Complexity Perspective on Language, Communication and Society [C]. Heidelberg: Springer-Verlag, 13:197-218.
Munné, F. 2013. The Fuzzy Complexity of Language [A]. In A. Massip-Bonet, A. Bastardas-Boada (eds.). Understanding Complex System Complexity Perspective on Language, Communication and Society [C]. Heidelberg: Springer-Verlag, 175-196.
Oya, M. 2011. Syntactic Dependency Distance as Sentence Complexity Measure [A]. In Proceedings of the 16th Conference of Pan-Pacific Association of Applied Linguistics [C]. 313-316.
Rescher, N. 1998. Complexity. A Philosophical Overview [M]. London: Transaction Publishers.
Riddle, E. M. 2008. Complexity in isolating languages: Lexical elaboration versus grammatical economy [A]. In M. Miestamo, K. Sinnemäki & F. Karlsson (eds.). Language Complexity: Typology, contact, change [C]. Amsterdam/Philadelphia: John Benjamins Publishing Company, 133-151.
Sampson, G., D. Gil & P. Trudgill (eds.). 2009. Language Complexity as an Evolving Variable [C]. Oxford Studies in the Evolution of Language, 13. Oxford: Oxford University Press.
Schoenemann, P. T. 1999. Syntax as an Emergent Characteristic of the Evolution of Semantic Complexity [J]. Minds and Machines, 9(3):309-346.
Shen Jiaxuan. 1999. Asymmetry and Markedness Theory [M]. Nanchang: Jiangxi Education Press.
Shi Anshi & Zhan Renfeng. 1986. The Commonality, Classification and Asymmetry of Antonyms [A]. In Yang Shengchu (eds.) Collection of Essays on Chinese Semantics [C]. Changsha: Hunan People’s Press, 395-411.
Simon, H. A. 1987. The Science of the Artificial [M]. Chinese Translation. Beijing: The Commercial Press.
Sinnemäki, K. 2011. Language Universals and Linguistic Complexity —Three case studies in core argument marking [D]. Ph.D. dissertation, University of Helsinki.
Szmrecsanyi, B. & B. Kortmann. Introduction: Linguistic complexity —Second language acquisition, indigenization, contact [A]. In Kortmann & Szmrecsanyi (eds.). 2012. Linguistic complexity —Second language acquisition, indigenization, contact [C]. Berlin: Walter de Gruyter, 6-34.
The Five Graces Group. 2009. Language is a Complex Adaptive System [J]. Language Learning, 59(s1):1-26.
Tomasetto, M. 2003. Constructing a Language: A usage-based approach to language acquisition [M]. Cambridge, MA: Harvard University Press.
Wang Li. 1980. History of Chinese Language [M]. Beijing: Zhonghua Book Company.
Xu Guozhi. 2000. Systems Science [M]. Shanghai: Shanghai Technology and Education Press.
Zhou Tongquan, Shu Hua & Ke Youping, 2010. The Complexity of Verbs and its Effects on Speech Processing [J]. Wuhan University Journal (Humanity Science), 63(6):767-773.

(Translated by Qi Shan, School of Foreign Studies, Nanjing University)