
# Proceedings of The 29th North American Conference on Chinese Linguistics (NACCL-29)

Volume 1

Conference Organizer: Richard VanNess Simmons (Rutgers University)
Proceedings Editor: Lan Zhang (University of Memphis)

Proceedings of the 29th North American Conference on Chinese Linguistics (NACCL-29)

Volume 1/edited by Lan Zhang

*Published by*:

University of Memphis
Memphis, TN USA

*Distributed by*:

NACCL Proceedings Online
Department of East Asian Languages and Literatures
398 Hagerty Hall
1775 College Road
The Ohio State University
Columbus, Ohio USA

URL: http://naccl.osu.edu

CONTENTS OF VOLUME 1

## CONFERENCE ORGANIZER'S REMARKS

## AND ACKNOWLEDGEMENTS

The idea of Rutgers serving as host of a NACCL meeting was first raised when Marjorie Chan approached me during NACCL-25 at the University of Michigan in 2013 (held June 21-23). At the time her proposal was for us to host a few years into the future and I quickly accepted, thinking that there was plenty of preparation time. After some discussion in the months that followed, we agreed that Rutgers would host in 2017 for the 29th meeting. The subsequent few years passed quickly, however, and when it came time to actually move forward on the conference preparations, suddenly time seemed far less plentiful. But I was fortunate to have an enthusiastic team of colleagues and students who willingly joined in to assist. Things proceeded apace from the call for papers, through the review of abstracts, the scheduling of space and presentations, and the preparation of the program.

The conference came to a successful conclusion after a busy and fast-paced two-and-a-half days. But in the bustle of conference planning and hosting, I had not given thought to the preparation of a conference proceedings volume. I was also quickly distracted by many other obligations and activities that came to the fore after NACCL-29 ended. So when Lan Zhang wrote to me a few weeks later in July to inquire about plans for a volume of conference proceedings, I was at first at a loss as to how to reply—until I noticed that her inquiry included an offer to assist. After a bit of further discussion and consultation with Marjorie Chan, I was delighted that we were able to parlay Lan Zhang's offer of assistance into full editorship of the proceedings. She readily took the reins and we sent out a call for submissions soon after. Just one-and-a-half years later, the result of her editorship is the two volumes now before us. They reveal that we could not have identified a better editor.

I think that all contributors and readers will be equally and greatly impressed by the results of Lan Zhang's detailed and painstaking effort. She has put together a fine record and worthy representative of the NACCL-29 conference. Contained herein is a collection of high quality papers that reflect the many facets of the conference theme, *Perspectives on the History, Geography, and Sociolinguistics of Chinese and Chinese Dialects*, led by the generous and significant contributions of two of the keynote speakers at the conference. The quality and range of the selection of papers in these pages will certainly ensure that these proceedings will have significant prestige and scholarly impact in the field and stand among the finest that NACCL has offered over its three decades of activity.

I wish here to express my sincerest thanks to Lan Zhang for all the time and effort she put into these NACCL-29 proceedings. I am most deeply indebted to her for so generously taking up the editorship and for seeing it through to completion, thus bringing the NACCL-29 endeavor to such a fine and high-quality conclusion. These proceedings would likely not have happened were it not for her freely offered yet assiduous stewardship of the project. I would also like to thank David Prager Branner and Marinus Van Den Berg, the two contributing keynote speakers, for their articles. I am grateful and honored that they have thereby acknowledged the value of this venue to represent their important work. I am of course also grateful for the contributions from

all the other scholar authors represented in these proceedings. These volumes would have no content if these scholars had not done the research and writing that we can now see within their pages. In addition, I wish to acknowledge once again the work and organizational contributions of the NACCL-29 Organizing Committee at Rutgers in 2016-2017 (alphabetically): Luca Iacoponi, John Phan, Jenny Yang, Wei Yang, and Yu Lou. Thanks are also due to Marjorie Chan for giving me the opportunity to host NACCL-29 at Rutgers and for cheering us on from planning through production. Finally, many thanks are due to the Program Coordinators at the Rutgers University Office of Continuing Professional Education that year, who facilitated logistics for NACCL-29, Dalynn Knigge and Bianca Kovalenko. They made it possible for the conference to actually happen, and, together with all the others mentioned above, thus set up the *raison d'être* for these proceedings.

Richard VanNess Simmons

April 2, 2019

# PREFACE

The 29th North American Conference on Chinese Linguistics (NACCL-29) was held at Rutgers University on June 16-18, 2017.

After the conference ended, the presenters were invited to submit their revised papers for publication in the Proceedings of the 29th North American Conference on Chinese Linguistics. Thirty-six of the presented papers were submitted and included in the proceedings. The papers have been divided into nine parts and placed in two volumes: Volume 1 is composed of Parts 1 through 7, and Volume 2 includes Parts 8 and 9. In Volume 1, Part 1 contains the papers of two invited speakers, Dr. David Branner and Dr. Marinus Van Den Berg. Dr. Branner dedicates his paper to Professor Jerry Norman. Part 2 comprises papers on phonetics and phonology; Part 3 pertains to discourse analysis; Part 4 is on historical linguistics research; Part 5 concerns language acquisition; Part 6 deals with lexicology; and Part 7 is on language contact. In Volume 2, Part 8 contains papers on syntax and semantics, and Part 9 is on sociolinguistics and regional varieties.

Proceedings of the 29th North American Conference on Chinese Linguistics (NACCL-29). 2017. Volume 1.
Edited by Lan Zhang. University of Memphis, Memphis, TN. Pages 1-84.



# Wenyan Syntax as Context-Free Formal Grammar

David Prager Branner
*New York, NY*

*Remembering Jerry Norman (1936–2012)*

An essential feature of the syntax of premodern written Chinese (Wenyan 文言) is its lack of morphology, resulting in overwhelming simplicity. Here I describe Wenyan using a context-free grammar, removing semantics to reduce the rules of syntax to just three and the parts of speech to two (plus some particles). All eight combinations of two elements and three rules are well attested and meaningful, and many more complex analyses can be reduced to them. I also illustrate this analysis by comparing Wenyan to a functional programming language.

Contents



---

[1.] This paper was presented as a plenary lecture at the 29th North American Conference on Chinese Linguistics (NACCL-29) on 17 June, 2017. The first public presentation of the ideas in this paper was in Branner 2015.

## 1. Philosophy and Key Ideas

This paper concerns the syntax of Wenyan (*wényán* 文言) — premodern Chinese written language. Many goals are possible in the study of a large corpus of an ancient written language. Mine here is to develop a formal model to aid reading, given Wenyan's signature features. It is a symbolic system, in the sense described by Yuen Ren Chao:[2]

[2] Chao 1968a:202. "This would at first sight seem to be a kind of intellectual perversity. But that is exactly the procedure of much of modern mathematics."

> A symbolic system is built up in which the terms and relations do not refer to anything concrete and are defined implicitly by the set of their behaviour in the system.

I rely here on two useful ideas from Computer Science: *separation of concerns* and *context-free grammar*.[3] I introduce them in this first section, after an outline of the key ideas in the grammar. Although this section is highly formal, the rest of the grammar is quite simple and not formal at all — Wenyan possesses the terse elegance of a symbolic system, and these two Computer Science concepts are useful for giving precision to that idea. But the reader may prefer to skip the philosophy overview and start with Section 2.

Section 2 discusses the key principle in Wenyan grammar: the indeterminacy of parts of speech. Section 3 presents an actual context-free grammar of Wenyan and discusses its semantic implications. I list conventions of notation at the very end of the paper

### 1.1. Outline of the Model

The main points of the grammar described here are six:

(a) All regular semantic categories reduce to the functions verb (V) and noun (N), which are the two parts of speech in this model. Words whose semantic categories are not regular do not reduce in this way and are treated as grammar particles. Grammar particles generally disambiguate relationships between the regular parts of speech.

(b) Part of speech is merely function, and is not intrinsic to verbs and nouns, nor are these categories truly distinct: verbs can be treated as nouns on the fly, and vice versa. But in reading it is useful to distinguish noun from verb function, and to be able to state how words acting in the two functions are related to one another. That is the basis of this grammar.

(c) All four combinations of verb and/or noun in two-word combinations are well attested in writing: $N_1$ $N_2$, $V_1$ $V_2$, V N, and N V.

(d) Two words of either part of speech, when juxtaposed, can be related in two ways: coordinately (in sequence) or subordinately (the first modifying the second).

[3] In Computer Science literature these phrases are sometimes seen abbreviated to SoC and CFG.

(e) Once coordinate and subordinate relations between words are taken into account, eight combinations of verb and/or noun are possible, and all eight are well attested in texts. Fig. 2, later in the text, shows the eight combinations, and Section 3 details the semantic interpretation of each, with examples.

(f) Part of speech is fundamentally ambiguous. But, when necessary, combinations of verb and/or noun can be reduced through recursive application of the rules above. Every pair of words can be reduced to V or N function, and every group of three or more V or N can be reduced to two.

### 1.2. Separation of Concerns

To separate concerns means to minimize the complexity of a system, by assigning different kinds of functionality to discrete components of that system. The phrase is due to Edsger W. Dijkstra (1930–2002), who advocates the independence of tools at different levels, in a hierarchy of components of a program or collection of programs:[4]

> Whenever a piece of mathematical reasoning appeals to a theorem, the only thing that matters on that level is *what* the theorem asserts and on that level it is equally immaterial how that theorem has been proved (elsewhere). … We can — and should — apply the same principle in program composition, where it is rewarding to separate for each program component clearly "what it does" and "how it works." … It is unwise to take the internal structure of a program component into account on the level where it is used on account of what it does.

That is, tools should be usable as black boxes: unanalyzed wholes. Separation of concerns became widely implemented in the philosophy underlying the UNIX operating system. Programmers encounter it most often today in the form of input-output redirection and piping.[5]

As written, Dijkstra's statement may seem hard to connect to analyzing language. As applied to Wenyan syntax, separation of concerns chiefly means excluding considerations of semantics and etymology as much as possible, as they are the domain of lexicon rather than syntax. That is an unusual choice; it is more common to mix semantics and other ideas with syntax — as, for instance, when we distinguish pronouns from common nouns from proper nouns, or animate from inanimate nouns, or adverbs of manner from adverbs of extent, or negatives from other kinds of words.

---

[4.] Dijkstra 1969:10; cf. 1974.

[5.] See Kernighan & Pike, 1984:29–33.

Meaning plainly contributes to how we settle syntactic relationships, and full separation is impossible, but segregation is not. My intention is that syntax be limited to describing relationships between parts of speech.

We are aware of historical considerations when we read and analyze Wenyan — of trends in diction and style, of the history of phonology as it affects rhyming and prosodic organization, and so forth. But if possible, syntax itself should be *achronic* — free of the considerations of time — again as a matter of separation of concerns. That is, unlike the free-wheeling practice of reading, when you consider syntax you ideally work under a single system of analysis.

### 1.3. Context-Free Grammar

I represent Wenyan syntax abstractly using context-free grammar, a form of notation that is efficient and permits systematic analysis.

Context-free grammar seems to have been conceived twice. N. Chomsky published first, in 1956. His version consists of formal *rules* for describing the relationships between elements of natural language. He introduced it, under the name "phrase-structure grammar," as a foil against which to argue for his more sophisticated Transformational Grammar model for natural languages.

Independently, in research beginning the year before Chomsky's paper but not published until afterwards, the Gesellschaft für Angewandte Mathematik und Mechanik (GAMM), later joined by the Association for Computing Machinery, developed proposals for something whose various stages of development were termed a "universal programming language," a "common formula language," and an "international algebraic language."[6] It was introduced as an "International Algebraic Language" in 1958,[7] and revised as the "reference language" for ALGOL 1960.[8] The papers reporting the GAMM/ALGOL research make no mention of Chomsky's 1956 paper, and there is no indication that one influenced the other.

In Computer Science, "grammar" generally means context-free grammar, without further specification. That is different from Linguistics, where (even only in reference to Chomsky) grammar is a general term and we can make no assumptions about which of countless models is intended. Credit for "context-free grammar" is hard to assign. The name is certainly due to Chomsky,[9] but there exist two different traditions of notation for it (see below). Chomsky's notation uses syntax trees prominently, as illustrated in the

---

6. Perlis & Samelson 1960:268–71.

7. Backus 1959.

8. Naur 1960.

9. Chomsky 1959:393.

next section. But trees are not original to Chomsky — they were already in use in the mid-19th century in both linguistics and mathematics.[10] And it is hard to say just where the innovation of context-free grammar lies. Perhaps there actually is no stark innovation in this form of systematic notation, only accrued best practice.

### 1.3.1. The Two Forms of Context-free Grammar

Conventional context-free grammar applies to the parts of a classical *syntax tree*. A syntax tree is made up of nodes, branching downward from a *stem* that represents a whole sentences or other analyzable unit (Fig. 1).[11]

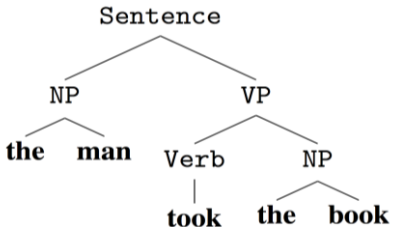


Fig. 1: Syntax tree (Chomsky 1956:117)

A context-free grammar offers two advantages over a tree. First, it is modular, so if the model is complex we can *compose* it (build it up out of small parts) or study it incrementally, rather than as a whole. Second, we can generalize into a single entity all the patterns expressed in many different trees. The name context-free refers to the fact that each rule can be applied independently of any others.

People tend to find syntax trees intuitive to grasp. By contrast, below consider a context-free grammar representing the tree of Fig. 1. The system distinguishes *terminal strings* (called *leaf nodes* in a tree: nodes with no further children, and bold-faced in Fig. 1) from *non-terminal strings* (called *inner, non-leaf nodes* in a tree). Non-terminal strings are subject to *rewrite rules*, by which they are reinterpreted or replaced; terminals are

---

[10] On the history of tree-models in analyzing syntax, see Brittain 1973:67–72 on D.B. Chamberlin and 1973:138–45 on Charles Gauss and B.T. Hodge, Percival 1976, Mazziota 2016, and Mazziota & Kahane 2017. Percival traces analysis by phrasal constituents to the fifteenth century. As for mathematics, Donald Knuth shows that tree abstractions were already being analyzed mathematically in the mid-19th century (Knuth 1968:405–6, revised in 1997:406–7; Sec. 2.3.4.6).

[11] The traditional term is *root*. Knuth suggests *apex*, and reviews the linguistic problems that arise from using *root* to mean the beginning of a tree that apparently grows downward (1997:311; the whole of section 2.3 is revised from 1968:305ff). But I prefer *stem*, since it seems to me to apply to a plant growing in any direction.

special in that they are not subject to those rules. Non-terminal strings are dominated by noun phrases (NP) and verb phrases (VP) — a noun phrase has all the same semantic characteristics of a noun; it differs only in that it can be further analyzed into constituent parts. In the four examples below, Chomsky's 1956 notation appears on the left; on the right is the same statement in Backus-Naur Form. Each is followed by a plain English statement of its meaning:

| | *Chomsky's 1956 notation* | *Backus-Naur Form* |
|---|---|---|
| (1) | Rule: Sentence → NP VP | ⟨Sentence⟩ ::= ⟨NP⟩ ⟨VP⟩ |

Rule (1) means: *Sentence* can be *rewritten* as *NP VP*, noun-phrase followed by verb-phrase; all three elements are non-terminals.

| | | |
|---|---|---|
| (2) | Rule: VP → Verb NP | ⟨VP⟩ ::= ⟨Verb⟩ ⟨NP⟩ |

Rule (2) means: any verb-phrase *VP* in this tree can be rewritten as *Verb* followed by a noun-phrase *NP*; all three elements are non-terminals.

| | | |
|---|---|---|
| (3) | Rule: NP → 'the man', 'the book' | ⟨NP⟩ ::= 'the man' \| 'the book' |

Rule (3) means: any noun-phrase *NP* in this tree can be rewritten as one element from the collection "the man" and "the book," both of which are terminals.

| | | |
|---|---|---|
| (4) | Rule: Verb → 'took' | ⟨Verb⟩ ::= 'took' |

Rule (4) means: the element *Verb* in this tree can be rewritten as "took," a terminal.

Most of the differences between the notations look superficial:

(5) In Backus-Naur Form, non-terminals are marked by enclosure within angle brackets ⟨…⟩, where Chomsky uses no explicit demarcation.

(6) The disjunction ("*or*") relationship is represented with a pipe | where Chomsky uses a comma.

(7) For Chomsky's arrow, Backus-Naur Form uses a complex symbol composed of two colons and an equality sign, ::=.

These differences in notation are superficial but betoken three important conceptual distinctions.

First, there is a distinction in how the models are to meant be applied. Backus-Naur Form consists of "metalinguistic formulas" that embody "syntactic definitions," against Chomsky's "rules" that "produce derivations."[12] Chomsky's notation and terminology suggest cause-effect relationships and the passage of time, whereas the GAMM/ALGOL language is more frankly axiomatic and achronic.

Chomsky implies the passage of time not only with the word "derivations." As a matter of fact, he is working in the context of finite state languages, for which "rules" are achronic and do not imply the passage of time at all. But he is also offering a "theory of the structure of a [natural] language" producing valid (and only valid) utterances; he defines *theory* explicitly:[13]

> Linguistic theory attempts to explain the ability of a speaker to produce and understand new sentences, and to reject as ungrammatical other new sequences, on the basis of his limited linguistic experience.

In other words, Chomsky is modeling the actual production (and understanding) of language by an actual competent speaker.[14] Chomsky's version of a context-free grammar describes the way abstract linguistic entities produce the literal components of natural language, in their observed sequence. In the context of Computer Science, where parsing is the main application of a grammar, there is no need for the notion of a "speaker" with the "ability to produce" valid utterances. Instead, the grammar defines the fundamental types of which more complex, composite types or entities are made up; the application is analytical.

The *connective* symbol that connects the left and right parts of each rule or formula is characteristic of the difference between the two conceptions. Chomsky uses a rightward arrow, suggesting production. The arrow has long been the traditional connective for "implies" in propositional logic, and was in use as "stands for" and "the sign of definition" by the generation of mathematical logicians before Chomsky.[15] Backus-Naur Form and its immediate predecessors use, variously:

∼ :≣ ::=

---

[12] Backus 1959:14, Naur 1960:301, Chomsky 1956:117–8. Backus uses "formulas"; Naur substitutes "formulæ" and adds "syntactic definitions."

[13] Chomsky 1956:113.

[14] Even though he says elsewhere that "grammars take a completely neutral point of view," favoring neither speaker nor hearer (Chomsky 1959:137n1).

[15] See Church 1940:58, Church & Quine 1952:179.

— all suggesting equivalence.[16] For my purposes, in keeping with the principle of achronic formal analysis, equivalence is more parsimonious than production.

A second conceptual difference between the Chomsky and GAMM/ALGOL notations involves the completeness of the context-free grammar. Chomsky treats context-free grammar as the foundation on which to add two kinds of information to his fully elaborated grammar: morphophonemic rules and "transformations."

My view is that reading a purely written language is more like parsing a programming language than like modeling the linguistic behavior of a competent native speaker. Consider what the utility of morphophonemics or transformation might be in application to reading Wenyan. That assumption may well misrepresent how fluent readers, and especially those fluent readers who are native speakers, actually read. But it is not a goal of this paper to model linguistic behavior by humans.

### 1.3.2. The Purpose of Using Context-free Grammar

The model described in this paper uses context-free grammar to describe the relationships between parts of speech, in order to make those relationships precise. Precision in syntax helps pin down the meanings of words and phrases, which are sometimes extremely fuzzy Yuen Ren Chao (1896–1986) often used *looseness* to describe syntactic relationships in Mandarin, and Wenyan is much the same way.[17]

Morphophonemics — semantically motivated phonological changes to word structure — is supposed to allow for changes like making the English past tense from the present. The word *rule* is misleading here, since we may understand it to imply a regular pattern. But the set of rules for English includes various unique cases for which there is no pattern: *have* corresponds without pattern to *had*, and *take* without pattern to *took*, alongside the regular *walk* to *walked*. In reality, *rule* just means one element of a *mapping*: any abstraction on the left of the connective symbol can be replaced with any of the elements on the right.

But there are no morphophonemic rules acting on the received Chinese script (the written form of Wenyan that this syntactic model describes) and it appears they never have. The script represents sound — even ancient sound, to say nothing of modern — far too coarsely to indicate any derivational morphology. Wenyan in its received form is a written language made up of inert symbols. Morphophonemic changes also occur only very minimally in how written Wenyan is read aloud today. Under the logical razor of separation of concerns, morphophonemics is outside of my purview.

---

16. The initial report of 1958 (Perlis & Samelson 1960:173) uses a large tilde ～; Backus 1959 uses :≡, Naur 1960 uses ::=, which remains customary today.

17. Chao 1968b, Sections 2.4.1, 2.10.7, 5.2.6, and many other places.

Transformations are the signature linguistic concept of the Chomsky school; their purpose is to predict valid sentences that may not yet be attested, and to validate different attested syntactic patterns against one another. Chomsky's core example is making an active verb passive: if sentence (8), below, is valid, then sentence (9) is also valid:

(8) *The man ate the food.*

(9) *The food was eaten by the man.*

In English, even transformations may themselves also require application of morphophonemics — forming the English passive *was eaten* from *was* — whereas Wenyan verbs are normally ambiguous as to grammatical voice and show no change in form, even if we generally turn to active or passive expressions in translation:

(10)

| | |
|---|---|
| *zhìrén* 治人 | ⟨VP⟩ "for ruling to occur, affecting others" ☞ 'to rule others' |
| *zhì yú rén* 治於人 | ⟨VP⟩ "for ruling to occur vis-à-vis others" ☞ 'to be ruled by others' |
| *shírén* 食人 | ⟨VP⟩ "for feeding to occur, affecting others" ☞ 'to feed others' |
| *shí yú rén* 食於人 | ⟨VP⟩ "for feeding to occur vis-à-vis others" ☞ 'to be fed by others' |

治於人者食人、治人者食於人[18]
Those who are ruled feed others; those who rule are fed by others.

Note that between literal and idiomatic glosses I use a pointing index finger (☞), avoiding the arrow (→) used elsewhere.

(11)

| | |
|---|---|
| *yìwù* 役物 | ⟨VP⟩ "for putting into service to occur, affecting things" ☞ 'to put things into service' |
| *yì yú wù* 役於物 | ⟨VP⟩ "for putting into service to occur vis-à-vis things" ☞ 'to be put into service by things' |

傳曰、君子役物、小人役於物[19]
The saying is: The gentleman makes people and things his tools; the petty person is made a tool by people and things.

[18] *Mèngzǐ*, "Téng Wén gōng, shàng" 《孟子》滕文公上.
[19] *Xúnzǐ*, "Xiūshēn" 《荀子》脩身.

(12) *zhìrén* 制人 ⟨VP⟩ "for control to occur, affecting others"
*zhì yú rén* 制於人 ⟨VP⟩ "for control to occur vis-à-vis others"

通者常制人、窮者常制於人[20]
The successful usually control others; the unsuccessful are usually controlled by others.

However, even without a difference in the form of the graphs 治 or 食 or 役, Chomsky's larger point is that the syntactic pattern V 於 N implies the existence in the corpus of another pattern V N, and vice versa. Transformation predicts the existence of one sentence from the existence of the other, validating both.

Apart from morphophonemics, there are two reasons that a model like transformationalism is unnecessary for a study of Wenyan syntax. One is separation of concerns: the distinction between active and passive voice, and even the idea of grammatical voice itself, are all semantic and hence out of scope. The other is that this syntax does not model the linguistic behavior of a native speaker; it merely interprets the relationships between different parts of speech. Those relationships guide the semantic interpretation of a closed corpus by making precise the meanings of words and phrases, which are uniquely imprecise in Chinese.

A model of a native speaker's linguistic behavior would have to do much more than that. It would have to predict valid forms not yet attested. That, however, is different from my goal, which is to parse existing forms in a way that stays reasonably close to the original character of the language. For my purpose, a much simpler model is better, lest it generate invalid sentences. As Knuth remarks, "it is quite difficult to fathom the significance of a [formal] language defined by productions."[21]

### 1.3.3. Ambiguity: Competing Explanations for an Expression

Chomsky considers context-free grammar inadequate because some expressions may match more than one syntactic pattern, and the grammar provides no guidance for choosing the correct one. He calls such cases "nonequivalent derivations" and "constructional homonymity" (sic),[22] preferring a model of grammar with more decision-making power.[23]

---

[20] *Xúnzǐ*, "Xiūshēn" 《荀子》榮辱.

[21] Knuth 1964:735.

[22] Chomsky 1956:121, 124. As of this writing, *homonymity* has not yet been collected into the Oxford English dictionary, which prefers *homonymy* for this usage.

[23] Chomsky 1956:118, 123.

Here, however, context-free grammar is a tool for guiding the reader to find plausible semantics when reading. Ambiguity is useful, because each different syntactic structure that can we can apply to an expression brings with it distinctive semantic interpretations. Considering those distinctive interpretations improves our chances of finding a plausible way to interpret the text. Ambiguous relationships may be an obstacle to predicting a native speaker's linguistic behavior, but not to analyzing a corpus. In the latter, there does not have to be a single "correct" pattern.

### 1.4. Basic Assumptions

Seven basic assumptions in this paper are as follows (some others are discussed in Section 4.1):

(13) Premise: This is a model for analyzing premodern Chinese written language. It does not promise to produce correct sentences, and it is not meant apply to oral language.

(14) Premise: The model uses only two parts of speech — noun and verb. All other familiar parts of speech can be reduced to these two and all other constructions are, overall, either nominal or verbal.

(15) Premise: Part of speech is relative, not intrinsic. Context makes noun function as verb, and vice versa.

Premise 15 is the key principle of Wenyan syntax and it appears to apply to all words in these two parts of speech.

Aside: perhaps it seems strange to be using "context-free" grammar to describe a language in which context is everything. Context-free, however, describes the individual rules of the grammar, and means that each of them is independent of any others. There is no inconsistency with the rhetorical sense of "context" in Premise 15.

(16) Premise: *Syntax* here means the structure of relationships among parts of speech.

(17) Premise: Words that do not class together as parts of speech are called *particles*.

The two assumptions discussed earlier in the present section are:

(18) Premise: Relationships among parts of speech are expressed using *context-free grammar* as an efficient representation.

(19) Premise: I strive to apply this idea under the model of *separation of concerns*, without involving other aspects of grammar, especially semantics, but also evidence of the scholia.

I discuss the assumptions other than those two, and their implications, in Section 2.

## 2. The Key Principle of Wenyan Syntax and Its Implications

### 2.1. Wenyan as the Object of Study

I do not define exactly the corpus to which the model applies, but in the main it comprises the received written texts of the millennium covering the Eastern Zhōu 周, Warring States (Zhànguó 戰國), and Hàn 漢 eras (770 B.C.E.–C.E. 220), with a long tail of texts composed up to modern times by people who had intimate knowledge of that core corpus. There are other corpora, ancient and more recent, but I do not address them here. I think it less useful to delimit the corpus in the abstract than to seek a system of syntax that is coherent and covers the core corpora reasonably well. Using single system of analysis highlights exceptional passages in the core corpus, as well as usage in older or newer texts that the rules of the system do not cover.

My working assumption is that these texts represent language that was always, even in the Warring States, fundamentally written, so that it usually embodied a high diglossic register, rather than genuine oral language (of any register). Some received ancient texts purport to record actual dialogue, and perhaps they do; but the important thing is that they also appear to fit within the model described here.

The corpus of Western Zhōu and Warring States bronze inscriptions fit this model, too.[24] And the model can apply to much of the lexicon of modern Mandarin, but not to all of it, and in any case Mandarin is not the target object.

### 2.2. On Reducing Other Parts of Speech to Noun and Verb

Most of the familiar parts of speech — things like pronouns and prepositions, and so forth — are semantic categories; their syntactic roles, however, are only either *verbal* (verb: V) or *nominal* (noun: N). Nouns are a large category:

(20) Definition: What is normally meant by *pronouns and proper nouns* are fundamentally nouns.

Semantically, pronouns and proper nouns are indeed distinct from common nouns, but (under separation of concerns) semantics is not at issue in syntax. They do not share all of

[24.] Branner 2018.

the possible behavior of common nouns (pronouns, for instance, do not take attributive modifiers) but that would affect rules about production — how we generate correct Wenyan — not how we read or analyze it.

(21) Definition: What is normally meant by *prepositions* or *coverbs* are fundamentally verbs.

In Mandarin and for prescriptive purposes, it is useful to distinguish the coverb from other verbs, but in describing Wenyan the distinction is unnecessary.[25]

(22) Definition: What is normally meant by *adverbs* is something that modifies verbs, usually as to *manner or extent*. (Manner and extent are semantic matters, however.) Both verbs and nouns can modify a verb; there is no place for a separate adverb category.

(23) Definition: What is normally meant by *adjectives* are *verbs of state or quality*. They may appear distinct in meaning from verbs of action, but that is a detail of semantics, not syntax.

In regard to their syntactic behavior, adjectives and verbs are not distinct in Wényán, but they are distinct in modern spoken languages, such as Mandarin, and that can cast a haze over our clear understanding of Wenyan syntax. The issue depends in part on how we define intensive adverbs, but the adverb is not a necessary category in Wényán.

### 2.3. The Interchange of Noun and Verb

The verb-noun distinction is semantic in nature. For Wenyan syntax, it is a convenient fiction because it ramifies syntactic relationships; those relationships are then useful for specifying the semantics of words in combination.

---

[25] The word *co-verb* (hyphenated) apparently originates with George Kennedy (1938:v–vi); the first statement in English that Chinese prepositions are fundamentally verbs also seems to be from Kennedy's hand (1939:15):

> 5a) Chinese has no original class of words that can be called prepositions. All of the functions of prepositions in European languages are performed in Chinese by verbs.

Coverb, normally *jiècí* 介詞 'preposition', is sometimes rendered literally as *gòngdòngcí* 共動詞 in Chinese.

The idea of a starkly reduced set of Chinese parts of speech has been in the air, in North American sinology, from the mid-twentieth century onward. Western students of language certainly noticed and praised the minimalist beauty of Wenyan long before that. Wilhelm von Humboldt (1767–1835), for one, described it as *reiner Gedanke vermittelst der Sprache* [pure thought mediated by language].[26]

Learning to re-conceive of a noun as a verb and a verb as a noun is one of the basic skills of solving syntactic problems when reading Wenyan. Cultivating this skill, explicitly, appears to have been part of the highly minimalistic pedagogical tradition of Peter A. Boodberg (1903–1972). Boodberg did not teach that there were only verbs and nouns; his *Syllabus* also names adjectives and adverbs as parts of speech.[27] Even more suggestive is the work of George Kennedy (1901–60), who advocates three parts of speech: nouns, verbs, and *ambs* (from "ambivalent"), meaning verb-like words that take no object and can serve as adjunct to a noun.[28] Other hands, such as John A. Cikoski, have also proposed highly restricted inventories of parts of speech and interchangeable functions among them.[29] Christoph Harbsmeier describes "hermaphroditic verbo-nominal hybrid functions."[30]

The difference between nouns and verbs is semantic because verbs name actions or states; nouns name objects material or abstract, including persons and places. This single semantic contrast is the basis of the verb-noun part-of-speech distinction, and the distinction is the whole basis of syntax. Overall, the system of syntax in this paper outputs semantics but it does not take semantics as its input, apart from this one noun-verb contrast.

Below I introduce some examples of the verb-noun relationship. The first are cases where the noun is apparently primary and the verb apparently derived. (I discuss the terms primary and derived in Section 2.4, below.)

---

[26.] Humboldt 1907[1836]:164.

[27.] Boodberg 1943: "Introduction".

[28.] Kennedy 1964:370–77. He called parts of speech *word classes*. Kennedy also tried to subclassify his three parts of speech as to what the present paper calls semantic function: a noun *N* functioning as the object of a verb is symbolized *No* but a noun functioning adverbially is symbolized *Nd*; an amb *A* serving adjectivally is symbolized *Aj*; and so on.

[29.] See Cikoski 1978. Cikoski feels that "the working principles of [Wenyan] grammar … cannot possibly lie entirely in the structure of its sentences and nothing else" (1978:39–40), an outlook obviously at variance with the model in the present paper.

[30.] Harbsmeier 1998:142.

### 2.3.1. The Verb Usage of a Noun

Since the nouns in the following cases are concrete, I illustrate only the corresponding verbal usages. Literal renderings are in double quotes ("…") and more idiomatic renderings in single ('…') or none. If both appear in the same gloss, the symbol ☞ connects them.

(24) *rì* 日 — N 'sun; day'[31]
V 'to last for a day'

Here is an example of the verbal usage:

(25) *búrì* 不日 — ⟨VP⟩ 'not to pass a day's duration'

不日成之[32]
"to accomplish it such that one does not pass a day's duration"
☞ 'to get it done in under a day'

Here are three further examples:

(26) *běn* 本 — N 'root or trunk of tree; fundamental issue'
V 'to treat as the fundamental issue; to found upon'

(27) *běntǒng* 本統 — ⟨VP⟩ 'to found upon general principles'

未本仁義之統也[33]
[These rulers] did not found their government on Confucian notions of goodness and righteousness.

(28) *bì* 畢 — N 'kind of hunting net with long handles'
V 'to go hunting with a 畢-net' (also written 罼)

---

[31] In keeping with the idea that part of speech is not intrinsic in Wenyan, I want to understand these glosses as *alternants*, so I do not assign them separate citation numbers.
[32] *Mèngzǐ*, "Liáng Huì wáng, shàng" 《孟子》梁惠王上.
[33] *Hàn shū*, "Xíngfǎ zhì" 《漢書》刑法志.

(29) *bìzhī* 畢之 ⟨VP⟩ ‘to hunt them with nets’

鴛鴦于飛、畢之羅之[34]
The mandarin ducks are flying off —
net them with handled nets, net them with thrown nets.

(30) *chǔ* 楚 N ‘the state of Srhah’
V ‘to behave as people do in the state of Srhah’

(31) *qí chǔ* 其楚 ⟨NP⟩ ‘their behaving as appropriate to the state of Srhah’

日撻而求其楚[35]
to beat [your children] daily and hope they behave as though in Srhah

The expression *qí chǔ* 其楚 is a noun phrase here, but within it *chǔ* has its verbal meaning.

### 2.3.2. The Nominal Usage of Verbs

In contrast, here are examples of verbs — meaning those words that we easily conceive of as verbs — that also have noun behavior attested. Below I list the verbal usage first.

(32) *biàn* 變 V ‘to change, be changed’
N ‘a change; esp. major/sudden change’

(33) *biàntài* 變態 ⟨NP⟩ ‘changing circumstances’

並遇變態而不窮、審之禮也[36]
He doesn’t run out of ideas, even if he repeatedly runs into changing circumstances — that is because he carefully considers them against the touchstone of how a person lives up to social relationships.

In the noun phrase *biàntài* 變態, *biàn* 變 is verbal.

---

34. *Shījīng*, “Xiǎoyǎ, Sānghù zhī shí, Yuānyāng” 《詩經》小雅·桑扈之什·鴛鴦.
35. *Mèngzǐ*, “Téng Wén gōng, xià” 《孟子》滕文公下.
36. *Xúnzǐ*, “Jūndào” 《荀子》君道.

(34) *sānbiàn* 三變 ⟨NP⟩ ‘three changes’

君子有三變、望之儼然、即之也溫、聽其言也厲[37]
The gentleman undergoes three changes in how he appears to others. To view him at a distance, he seems imposing. When you approach him, he is warm. When you hear what he says, he seems stern.

(35) *chǎn* 產 V ‘to give birth, be born’
N ‘product, production; property that produces income’

(36) *dúchǎn* 獨產 ⟨VP⟩ ‘to be produced singly’

士貴故孤興、物貴故獨產[38]
A scholar is worthy and so appears in this world without peers; an object is valuable and so is produced as the only one of its kind.

(37) *héngchǎn* 恆產 ⟨NP⟩ ‘constant income’

無恆產而有恆心者、惟士為能[39]
If it is a question of maintaining a constant state of mind without a constant income, only a scholar can manage it.

(38) *wàng* 望 V ‘to hope for, gaze at’
N ‘aspiration; prestige; the full, “gazed-at” moon’

(39) *wàngzhī* 望之 ⟨VP⟩ ‘to gaze at them’

誅國之民望之若父母[40]
The people of a state vanquished in a punitive war gaze at them with admiration as though their parents.

37. *Lúnyǔ*, “Zǐzhāng” 《論語》子張.
38. *Lùnhéng*, “Zìjì” 《論衡》自紀.
39. *Mèngzǐ*, “Liáng Huì wáng, shàng” 《孟子》梁惠王上.
40. *Lǚshì chūnqiū*, “Mèngqiū jì, Huái chóng” 《呂氏春秋》孟秋紀·懷寵.

(40) 夫世子者、國之基也、而百姓之望也[41]
The Crown Prince is the foundation of the state. That being so, he is the hope of the common people.

## 2.4. The Relativity of Part of Speech

In the examples above, it may seem obvious to the reader that one behavior is more basic than the other. In gloss 41 it is easy to conclude that a word is intrinsically nominal and that the verbal sense is derived:

(41) *lèi* 類 N ‘category’
V “to ‘category’ ” ☞ ‘to fall into the same category’

The noun definition seems more concrete, and Axel Schuessler cites cognates in Tibeto-Burman languages and transcriptions into Chinese, all nouns.[42] But the verbal sense is well attested in various ancient or conservative texts in the corpus; for instance:

(42) *búlèi* 不類 ⟨VP⟩ ‘not to fall into a category’

自底不類[43]
I brought myself to the point that I was not among [the good].

(43) *zhīlèi* 知類 ⟨VP⟩ ‘to understand “categorying” ’

九年知類通達[44]
In their ninth year, pupils understand how things fall into categories and they possess judgment.

The notion that, of a verbal and nominal usage, one is primary and the other derived is a convenience in reading. As a matter of separation of concerns, a syntactic analysis should not depend on the etymological question of whether a usage is actually primary or derived. Neither of these parts of speech is really independent of the other; it is their contrast that makes each meaningful in the overall system. Almost every content word can behave as noun or as verb in some context. Certainly every word, even a particle, can function nominally when one is referring to it as a word or graph. For instance:

---

41. *Shuōyuàn*, “Jiànběn” 《說苑》建本.
42. Schuessler 2007:347; 2009:314.
43. *Shàng shū*, “Shāng shū, Tàijiǎ, zhōng” 《尚書》商書·太甲中.
44. *Lǐ jì*, “Xué jì” 《禮記》學記.

(44) *qǐ zhī shǔ* 豈之屬 ⟨NP⟩ 'things that belong to the graph 豈'
*cóngqǐ* 从豈 ⟨VP⟩ 'to follow the graph 豈'

凡豈之屬皆从豈[45]
Everything that belongs to the graph 豈 "follows" (is placed after, or is derived from) the graph 豈.

Here "the graph 豈" is a noun, even though as a word *qǐ* 豈 is a particle. And we can understand essentially every utterance used alone as functioning verbally — as long we can cast it as declaring some action or state. Under Premise 15, that part of speech is not intrinsic, there is no need to make assumptions about which sense is later historically than the other.

The flexibility of parts of speech is not limited to Wenyan; it is common in most forms of Chinese — in Mandarin, for instance. Below are three common cases of nouns used as adjectives. (Mandarin adjectives are a subcategory of verbs with distinctive syntax, and parts of speech here are for Mandarin.):

(45) *péngyou* 朋友 ⟨NP⟩ 'friend'

(46) *bù péngyou* 不朋友 ⟨VP⟩ "not to "friend' "
☞ 'not to behave as a friend should'

(47) *jūnzǐ* 君子 ⟨NP⟩ 'gentleman'

(48) *hěn jūnzǐ* 很君子 ⟨VP⟩ "very much to 'gentleman' "
☞ 'gentlemanly'

(49) *zhōngguó* 中國 ⟨NP⟩ 'China'

(50) *fēicháng zhōngguó* 非常中國 ⟨VP⟩ "very 'China' "
☞ 'highly characteristic of China'

There are countless such examples. Here is the same type of behavior in Taiwanese, another modern spoken language: *hoan* 番 has noun sense and means "non-Hàn person"; it is also an adjective.[46] Here are noun usages of the word or morpheme:

---

[45] *Shuōwén* 說文 6.

[46] Examples from Russell Sprinkle *et al.* (1976:245-6 etc.), exactly as in the original. The romanization is Church romanization, derived originally from Carstairs Douglas 1873; the asterisk rep-

(51) *hoan-á* 生番; 土人; 外國人 aborigine; foreigner (rude expression)

(52) *chhe*-hoan* 生番; 野蠻人 aborigines of Taiwan; an unruly fellow'

The morphological suffix *-á* in example 51 is diminutive or nominalizing in various senses; and in example 52 *hoan* is a bound form. Here are adjectival (that is, fundamentally verbal) usages: first as a free form and then subjected to intensive reduplication (a kind of adjectival morphology productive in Taiwanese):[47]

(53) *i chin hoan* 他很不講理 He is very barbarous (unreasonable).

(54) *hoan-hoan* 不講理 rude, ignorant, and unreasonable

Are these examples fundamentally nouns or verbs? Both usages are attested, but at least in the Mandarin examples we feel by instinct that their primary meanings are semantically nominal as a matter of common usage, and that the adjectival meanings are derived special cases. That is natural, because parts of speech in both Mandarin and Taiwanese are more rigid than in Wenyan. Yet even in these languages, examples of this sort are numerous and widespread; interchange of verb and noun remains productive in many forms of oral Chinese. It is a common typological feature across Sinitic languages.

Some words appear completely ambiguous as to part of speech: numbers, and also times, dates, and seasons. It is useful to think of them as the names of states: as states they are functionally verbal, but as names they are also functionally nominal. Again, this is a semantic problem, since verbs and nouns as parts of speech are interchangeable. But which is it more useful to think of as the primary sense?

Here is an example of a number, *sì* 四 'four', translated first as a noun:

(55) 不得終禮、廢者幾、孔子曰、四[48]
When it is not possible to complete the ritual, how many tries does it take for the ritual to be abandoned? Confucius said, "Four of them."

---

resents a nasalized final. There is no direct representation of the Taiwanese in characters in Sprinkle, and parts of speech are not marked.

47. Taiwanese also has intensive retriplication (for instance, *hoan-hoan-hoan* 'extremely uncivilized', with otherwise unique tone sandhi) but examples are not usually reflected in dictionaries, and even the distinctive phonetics of this construction often fails to make its way into formal descriptions of the language.

48. *Lǐ jì*, "Zēngzǐ wèn" 《禮記》曾子問.

Now consider it as a verb:

(55′) … “It takes four [tries].” or “[One does it] four times.”

How meaningful is the difference between 55 and 55′? It may seem to be picking nits, but I think a number is semantically ambiguous as a verb or a noun. It can be a noun because we manipulate it as though it were an object, but it can also be a verb because a quantity is a state: how many of some object we are talking about is part of the object’s state.

Seasons and dates most often function to modify verb phrases, a context where any semantic difference between nouns and verbs is erased, but it is possible to render them explicitly (if awkwardly) as verbs:

(56)

| | |
|---|---|
| *sānnián* 三年 | ⟨VP⟩ ‘to occur during the third year’ |
| *chūn* 春 | ⟨VP⟩ ‘to occur during the Spring’ |
| *wáng èryuè* 王二月 | ⟨VP⟩ ‘to occur during the second month as rectified by the King’ |
| *jǐsì* 己巳 | ⟨VP⟩ ‘to occur on day 6 of the cycle’ |
| *rì yǒu shízhī* 日有食之 | “During the appearance of the sun, there occurred ‘eating it’.” ☞ ‘There was an eclipse.’ |

三年、春、王二月、己巳、日有食之[49]
It was the third year and it was the Spring, during the second moon under the King’s rectified calendar, during day 6 of the cycle. There was an eclipse.

Or as nouns:

(56′)

| | |
|---|---|
| *sānnián* 三年 | ⟨NP⟩ ‘the third year’ |
| *chūn* 春 | ⟨NP⟩ ‘the Spring’ |
| *wáng èryuè* 王二月 | ⟨NP⟩ ‘the second month as rectified by the King’ |
| *jǐsì* 己巳 | ⟨NP⟩ ‘day 6 of the cycle’ |

三年、春、王二月、己巳、日有食之
Spring of the third year. The second moon under the King and day 6 of the cycle. There was an eclipse.

The difference is merely trivial in English translation. But in Chinese it is arbitrary, which is more significant than being trivial.

[49] *Zuǒ zhuàn* quoting *Chūnqiū*, “Executive Yǐn, year 3” 《左傳》隱公三年.

English speakers may consider colors indeterminate as to part of speech. But in Wenyan they function verbally, with regularity, because color, like number, can be a state rather than a thing.

(57) *chì* 赤 V ‘to be red’

病氣疝、客於膀胱、難於前後溲、而溺赤[50]
The overall state of his illness was “hilling up” and fullness in the belly, and it took up residence in his bladder. He had difficulty voiding waste, either in front or in back, and his urine was red.

(58) *hēi* 黑 V ‘to be black’

其氣清寒而黑[51]
Its overall state is clear-cold, and the [dominant] color is black.

Colors are indeed both nominal and adjectival in English, but that should not be at issue in deciding what part of speech to consider primary when untangling a problem passage in Wenyan.

### 2.5. Syntax as the Structure of Part-of-Speech Relationships

If part of speech is not intrinsic, then what is the point of treating a word as noun or verb at all? Could we have a syntactic system with no parts of speech?

Changes of semantics feel real to most of us, but it is more effective to think of a Chinese word as possessing a cloud of meaning, without rigid internal divisions. The distinction between Wenyan noun and verb is a fiction, but it is a useful one because it allows us to examine the collection of syntactic relationships between words. Without at least a single distinction (N vs. V), it is hard to imagine how to specify those relationships. Of course, much reading is carried out by the brain, impressionistically and without

[50] *Shǐ jì*, “Biǎn Què Cāng gōng lièzhuàn” 《史記》扁鵲倉公列傳. I render *shàn* 疝 abstractly, following *Shuōwén* “腹痛也 [It is when the belly hurts]” and *Shìmíng* 釋名 “心痛曰疝、疝詵也、氣詵詵然上而痛也 [When the heart hurts it is called 疝. It means 詵 “spreading rumor rampantly’. It means having pain when the overall state of illness involves rising after massing together].” The *Guǎngyùn* lists two different readings for the graph 疝: {sran-2a} → **shān* and {sranH-2a} → *shàn*. Let us agree to avoid rendering *qì* 氣 anachronistically as “energy.”

[51] *Chūnqiū fánlù, “Zhìshuǐ wǔxíng*” 《春秋繁露》治水五行.

conscious analysis.[52] It is when we have have to apply our minds to figuring out a difficult sentence that we slow down and proceed rationally, holding the larger meaning of the various words in mind and trying ‘‘fit’’ them, in their appropriate parts of speech, into a gestalt.

Since it is useful to have somewhere to start, even if we have to revise our views later, we try to assign the parts of speech a category N or V by “basic meaning” of the word. The process of settling on that basic meaning is not trivial. We must consider the word's most familiar usages, its behavior as documented in reference works, the structure of the graphs used to write it, the relative times when it is attested in different senses, and so on.

All four combinations of N and V in two positions are well attested and meaningful. In addition there are two ways that the two positions can be related. They can be *coordinate*, in which their meanings are sequential and of approximately equal semantic *weight* (relative importance). Or they can be *subordinate*, in which case the one of the elements has greater weight than the other; the element with greater weight is the *head* and the other an *adjunct* and modifies it. Figure 2 shows these relationships in tabular form.

coordinate relationship

| | | second syllable | |
|---|---|---|---|
| | | N | V |
| first syllable | N | N N | N V |
| | V | V N | V V |

subordinate relationship

| | | second syllable | |
|---|---|---|---|
| | | N | V |
| first syllable | N | (N)-N | (N)-V |
| | V | (V)-N | (V)-V |

Fig. 2: Possible combinations of V and N
in coordinate and subordinate relationships

The actual semantic interpretations of each combination appear in Section 3, below.

[52] Whether or not the brains of fluent Wenyan readers are really dealing cognitively with only two parts of speech, or whether their reading is mediated by the grammar of various spoken languages, is an interesting question that I do not think has been broached experimentally.

### 2.6. Particles

Particles are a large topic. By definition full of diverse behavior, they are mostly out of scope in this presentation, but here I offer brief discussion.

Particles embody syntactic function. At times they determine meaning decisively outright. But they also very often serve to disambiguate sentences that would otherwise be difficult to understand (even if still grammatical). Disambiguation usually means disambiguating the syntax, but at times it only means clarifying semantics. In the latter case, the syntax of the disambiguated sentence may not be the same as that of the original, ambiguous sentence.

However, there are times when a particle is semantically determinative, so it is not an infallible tool for disambiguating confusing syntax. Its presence may give a phrase a meaning different from what that phrase would normally mean without it. An example is:

(59) *zhōu mèng* 周夢 N V "Tiw dreams"

莊周夢為胡蝶[53]
Tiw, of clan Tsrang, dreamt of being a butterfly.

(60) *zhōu zhī mèng* 周之夢 ⟨NP⟩ 'the fact or situation of Tiw dreaming'

周之夢為胡蝶
"the situation of Tiw dreaming of being of a butterfly"
☞ Tiw's dream of being a butterfly…

Example 59 is a whole sentence semantically; example 60 is explicitly a noun phrase because *zhī* 之 appears between topic *zhōu* 周 and comment *mèng wéi húdié* 夢為胡蝶. See Sections 3.2.4, 3.2.5, and 3.3.3 on topic-comment structure.

Another example is:

(61) *zàocháo* 造朝[54] ⟨VP⟩ "to go to court" ☞ 'to attend court'

(62) *zào yú cháo* 造於朝 ⟨VP⟩ "to perform the act of attending, vis-à-vis court" ☞ 'to attend court'

where *zào* is most naturally understood in example 61 as transitive, even though it is semantically ambiguous as to transitivity; *zào* is intransitive in example 62 because *yú* 於 appears between it and the noun *cháo* 朝. It is simplest to translate both examples as "to

53. Both examples from *Zhuāngzǐ*, "Qíwù lùn" 《莊子》齊物論.
54. Both examples from *Mèngzǐ*, "Gōngsūn Chǒu, xià" 《孟子》公孫丑下.

attend court," but their syntax is not identical. (See also discussion on the objects of verbs of motion in examples 128–132.)

Particles do not make up a uniform and well-populated class, all showing the same syntactic function. The only well-populated classes are words that can behave as verb and words that can behave as noun — those are sets with, at least, many thousands of members. No function associated with any particle has cardinality anything like that, and some of the important particles — such as *zhě* 者, *suǒ* 所, *yǐ* 矣 — are essentially unique in their behavior. We might call particles a distinct part of speech in their own right, but given their failure to "category," it is better to treat them as something outside the framework of parts of speech.

For example, exclamations are best treated as particles, unless they can be shown to have verbal or nominal behavior. *Wū* 惡 and *wūhū* 惡乎 "Oh no! Alas!" (but also "how?") may be such cases.

Verb and noun are traditionally called *shízì* 實字 "substantial words, solid words, content-words', and words not of that category are called *xūzì* 虛字 "insubstantial words, empty words' [particles]. Particles are a tentative category, dominated by grammaticalized content-words. Whatever their meaning before grammaticalization, in their behavior as particles they tend to supply functionality rather than concrete semantics, and to disambiguate syntax. Under separation of concerns, I consider only their behavior, consigning examination of their origins to the distinct subject of etymology.

Particles contribute to syntactic patterns that are part of the larger picture of Wenyan grammar.

(63) Rule of Thumb: *bù* 不 normally implies the following expression is verbal. See the use of *búrì* 不日 and *búlèi* 不類 in examples 25 and 42.

(64) Rule of Thumb: *suǒ* 所 implies following expression is verbal and transitive; the combination 所 V is normally a noun phrase, meaning the object of the verb.

(65) Rule of Thumb: *ér* 而 implies that the preceding and following expressions are verbal. A fine example is 40, where *guó zhī jī* 國之基 'foundation of the state' and *bǎixìng zhī wàng* 百姓之望 'hope of the common people' are noun phrases, but the pattern N 也 is verbal: 'to be foundation of the state' and 'to be the hope of the common people'.

(66) Rule of Thumb: *wú* 無 'there is not; not to have' normally implies the following expression is nominal; it also has a semantically imperative meaning as an auxiliary verb 'do not…', but it is useful to treat semantic

auxiliary verbs as taking their "main" verbs as objects and therefore as nouns.

(67) Rule of Thumb: *yú* 於 has diverse functionality, but its most common usages imply the expression preceding it is verbal and the one following it is nominal. See examples 10–12, and 62.

(68) Rule of Thumb: *zhī* 之 also has diverse functionality, including regular uses as verb and as noun. But its most common use as a particle implies that the following expression is nominal.

(69) Rule of Thumb: *yě* 也 often implies that the preceding expression is nominal, although that is not true uniformly. Also, a nominal expression suffixed by *yě* 也 is normally verbal.

But I describe syntax in the present essay based on relationships between content words only, as represented by the part-of-speech types of their meanings. Particles contribute to Wenyan syntax, and they add to the inventory of syntactic relationships, though not to the inventory of relationships at the level of the three rules in Section 3. The behavior of a particle is best described using *patterns*, in which it appears mixed with content words whose relationships it mediates. For instance,

(70) ⟨VP⟩ 者 ⟨NP⟩ 'one who ⟨VP⟩; that which ⟨VP⟩; the action/state of ⟨VP⟩'

(71) ⟨NP⟩ 也 ⟨VP⟩ 'to be ⟨NP⟩; to be a case of ⟨NP⟩'

(72) 非 ⟨NP⟩ 也 ⟨VP⟩ "to negate or deny ⟨NP⟩"
☞ 'not to be ⟨NP⟩; not to be a case of ⟨NP⟩'

Note that the particles in these patterns are not parts of speech but particular words. If another particular word can replace one of them, that fact has to be documented in detail — particles do not fall into classes the way ⟨NP⟩ and ⟨VP⟩ do.

## 3. Wenyan Context-Free Grammar of Nouns and Verbs

The context-free grammar is limited to interactions between nouns and verbs, unmediated by particles. Notational conventions of the grammar now follow.

(73) Convention: Each *definition* consists of three parts: a non-terminal, the connective ::=, and one or more elements separated by a pipe | .

(74) Convention: Non-terminals are set in angle-brackets: for example, ⟨NP⟩.

(75) Convention: Almost all elements to the left of the connective are themselves made up of non-terminals because the system of definitions is almost completely recursive: ⟨VP⟩ can consist of a terminal V and ⟨NP⟩ can consist of a terminal N.

(76) Convention: The topmost element of the system is the abstraction *S*, which stands for the word *sentence* and would represent the *stem* (or *apex*, or *root*) in a traditional syntax tree; S is not a part of speech but represents a whole structure, present recursively in this system.

(77) Convention: The structure written *(x)-y* is subordinate, so that *y* is the primary element and *x* modifies it. In contrast, *x y* means that *x* and *y* are coordinate or interpreted in sequence.

(78) Convention: Subscript numbers are used to show that two of the same element appearing together do not represent identical elements: for instance, ⟨$NP_1$⟩ ⟨$NP_2$⟩ means two different ⟨NP⟩ elements, not one ⟨NP⟩ element repeated.

The grammar itself consists of three rules:

(79) Rule: ⟨S⟩ ::= ⟨VP⟩ | ⟨NP⟩

(80) Rule: ⟨NP⟩ ::= ⟨$NP_1$⟩ ⟨$NP_2$⟩ | ⟨NP⟩ ⟨VP⟩ | (⟨$NP_1$⟩)-⟨$NP_2$⟩ | (⟨VP⟩)-⟨NP⟩ | ⟨VP⟩ | N

(81) Rule: ⟨VP⟩ ::= ⟨$VP_1$⟩ ⟨$VP_2$⟩ | ⟨VP⟩ ⟨NP⟩ | (⟨$VP_1$⟩)-⟨$VP_2$⟩ | ⟨NP⟩ ⟨VP⟩ | (⟨NP⟩)-⟨VP⟩ | ⟨NP⟩ | V

That is the whole syntax of Wenyan. It is deceptively precise, because it treats verbs and nouns as distinct, when they are actually interchangeable. They are only hazarded as distinct in order to inflate the whole system.

Note that this conception of "syntax" governs only the relations between individual words, mostly only those directly adjacent to one another. Other hands have done much valuable work on sentence level discourse, but I consider that a semantic issue, beyond syntax.[55]

[55] I am thinking especially of the work of Barbara Meisterernst, e.g., Meisterernst 2016.

I keep the overall system simple by limiting syntax to the interactions of noun, verb, and particles, by excluding semantics, and by making use of recursive definitions. Note that ⟨NP⟩ ⟨VP⟩ appears under both Rule 80 and Rule 81, as discussed in Sections 3.2.4 and 3.3.3.

This set of definitions is dense and hard to apply, so I now list common semantic interpretations of each component, with examples and discussion.

### 3.1. Definitions of ⟨S⟩ and Notation

The presence of ⟨S⟩, a sentence, as a non-terminal with two replacements, covering the other two non-terminals in the grammar, is a formality that allows us to being applying the rules to whole pieces of text. We begin at some sentence and replace it according to the first rule.

Because Wenyan did not usually have punctuation in traditional times, the matter of how to delimit a sentence ⟨S⟩ is circular: we try to parse the text so that all words are accounted for in valid structures. In practice today, punctuation is done manually, and at times there are differences of opinion. An exception is passages whose patterning strongly implies breaks between phrases and sentences — through rhyming, line-length, or "counterbalanced" diction.[56] Overall, a "sentence" is a very lax concept in Wenyan.

Any verb phrase can potentially serve as the whole expression of ⟨S⟩; any of the seven definitions of ⟨VP⟩ (discussed in Section 3.2) may serve here, or indeed anywhere else that ⟨VP⟩ appears in the grammar.

Similarly, any noun phrase can potentially serve as the whole expression of ⟨S⟩. There is no absolute cleavage between ⟨NP⟩ and ⟨VP⟩, so any ⟨VP⟩ structure not listed as a replacement for ⟨NP⟩ is nevertheless reachable by the single step of replacing ⟨NP⟩ with ⟨VP⟩. But our interest is in finding relatively short paths to a plausible interpretation.

Let me digress to explain the notation in strings such as:

(82) $\langle S\rangle \rightarrow \langle NP\rangle\ \langle VP\rangle \rightarrow \langle VP_1\rangle\ \langle VP_2\rangle \rightarrow V_1\ N_1\ V_2\ N_2$

which may be confusing. (It is taken from example 106.) It reads as follows, left to right:

| | | |
|---|---|---|
| (82′) | ⟨S⟩ | The whole expression, at a high level, is a sentence. |
| | → ⟨NP⟩ ⟨VP⟩ | Reinterpret the high-level verb phrase as composed syntactically of a noun phrase followed by a lower-level verb phrase. |

[56] Counterbalancing (*duìchèn* 對稱) is sometimes translated as "parallelism" or "antithesis," although the phenomenon is broader than covered by those terms.

→ ⟨$VP_1$⟩ ⟨$VP_2$⟩ Reinterpret this ⟨NP⟩ ⟨VP⟩ structure semantically as two verb phrases, closer to the apparent meaning of the whole expression, but ⟨NP⟩ ⟨VP⟩ better describes its high-level syntactic structure.

→ $V_1$ $N_1$ $V_2$ $N_2$ Reinterpret the low-level categories as two verb-noun pairs.

The further to the right an element is, the more semantically explicit it is, and the closer to the actual glosses of the words involved. The direction of the "production" arrow (→) may seem counterintuitive. Think of it instead in the Neo-Platonic sense: to the left of the arrow is a more abstract or ideal form, and the right is the manifestation of the ideal form.

Note that this representation is not organized as a rigorous formal system; it is just intended to guide the elucidation of expressions.

### 3.2. Definitions of ⟨NP⟩ and Their Semantics

#### 3.2.1. ⟨NP⟩ ::= ⟨$NP_1$⟩ ⟨$NP_2$⟩

The usual semantic application of this structure is two coordinate nouns joined by conjunction ("and") or disjunction ("or"):

(83) Semantic Application: nouns joined by conjunction or disjunction
⟨$NP_1$⟩ and ⟨$NP_2$⟩
⟨$NP_1$⟩ or ⟨$NP_2$⟩

Examples are numerous and unchallenging (apart from figurative or idiomatic cases):

(84) *zhāoxì* 朝夕 ⟨NP⟩ → $N_1$ $N_2$ 'day and night'

朝夕必時、上貢必適、祭祀必敬[57]
Day and night must always come at their appropriate times; offering tribute must always be fitting [to the status of both parties]; making sacrifice to gods and ancestors must always be done with respect.

(85) *ěrmù* 耳目 ⟨NP⟩ → $N_1$ $N_2$ "ear and eye" ☞ 'hearing and seeing'

[57] *Lǔshì chūnqiū*, "Jìqiū jì, Shùnmín" 《呂氏春秋》季秋紀·順民.

> 孝子之養老也、樂其心不違其志、樂其耳目、安其寢處[58]
> When a filial son serves his elderly parents, he pleases their hearts and does not overstep their intentions. He gives them pleasant things to see and hear, and makes their bedroom and dwelling place comfortable.

### 3.2.2. ⟨NP⟩ ::= (⟨$NP_1$⟩)-⟨$NP_2$⟩

There is one common semantic application of this subordinate structure:

| | | |
|---|---|---|
| (86) | Semantic Application: | ⟨$NP_2$⟩ that is part of or belongs to ⟨$NP_1$⟩<br>⟨$NP_2$⟩ characterized by or associated with ⟨$NP_1$⟩ |

In this syntax, whole precedes part. So clan names precede the given names of individuals and larger regions precede the smaller regions located within them:

| | | |
|---|---|---|
| (87) | *dǒng zhòngshū* 董仲舒 | ⟨NP⟩ → ($N_1$)-$NP_2$ ‘Middleson Hla (仲舒) of clan Tongh (董)’[59] |
| (88) | *wǔlíng lǔzhōng* 武陵漊中 | ⟨NP⟩ → ($NP_1$)-$NP_2$ ‘Middle-Roh (漊中) District in Mrah’reng (武陵) Commandery’[60] |

Possessor precedes possessed:

| | | |
|---|---|---|
| (89) | *jūnjū* 君車 | ⟨NP⟩ → ($N_1$)-$N_2$ ‘the ruler’s carriage’[61] |
| (90) | *wúyán* 吾言 | ⟨NP⟩ → ($N_1$)-$N_2$ “my words” ☞ ‘what I say’ |
| (91) | *hújūn* 胡君 | ⟨NP⟩ → ($N_1$)-$N_2$ ‘the ruler of the Ga savages’ |

And in general, attribute precedes its bearer.

In passing, note that constructions like *tiānshàng* 天上 ‘in the Heavens’, are also semantically (N)-N; that analysis contradicts Kennedy’s claim about coverbs (footnote 25).

---

58. *Lǐjì*, “Nèizé” 《禮記》內則.
59. *Hàn shū*, “Dǒng Zhòngshū zhuàn” 《漢書》董仲舒傳.
60. *Hòu Hàn shū*, “Sù zōng Xiàozhāng dì jì” 《後漢書》肅宗孝章帝紀.
61. All three examples are from *Hán Fēizǐ*, “Shuōnán” 《韓非子》說難.

When a noun phrase is counted, the number and an suitable classifier generally appear after the noun phrase. It is also reasonable to think of ⟨$NP_2$⟩ as head and ⟨$NP_1$⟩ as adjunct, as in section 3.2.2:

(92) *dàichén sānrén* 大臣三人 ⟨NP⟩ → ⟨$NP_1$⟩ ⟨$NP_2$⟩ → ($V_1$)-$N_1$ ($V_2$)-$N_2$ "three persons who were great ministers" ☞ 'three important officials'

一朝而殺大臣三人[62]
He killed three important officials during the course of one morning.

(93) *kūndì wǔrén* 昆弟五人 ⟨NP⟩ → (⟨$NP_1$⟩)-⟨$NP_2$⟩ → ($V_1$)-$N_1$ ($V_2$)-$N_2$ "five persons who were eldest brother and younger brothers" ☞ 'five brothers'

魯有昆弟五人者、其父死[63]
There were five brothers in the state of Rah, and their father died.

We can also interpret ⟨$NP_2$⟩ in this structure as syntactically ⟨VP⟩ — describing the state of ⟨$NP_1$⟩ — so that the overall structure is topic-comment, ⟨NP⟩ → ⟨NP⟩ ⟨VP⟩ (see Section 3.2.4):

(92′) *dàichén sānrén* 大臣三人 ⟨NP⟩ → ⟨NP⟩ ⟨VP⟩ → ⟨$NP_1$⟩ ⟨$NP_2$⟩ → ($V_1$)-$N_1$ ($V_2$)-$N_2$ "the fact that as for great ministers, they are three persons" ☞ 'three important officials'

(93′) *kūndì wǔrén* 昆弟五人 ⟨NP⟩ → ⟨NP⟩ ⟨VP⟩ → ⟨$NP_1$⟩ ⟨$NP_2$⟩ → ($V_1$)-$N_1$ ($V_2$)-$N_2$ "the fact that as for eldest brother and younger brothers, they are five persons" ☞ 'five brothers'

Semantically, these expressions appear to be noun phrases, and so it is reasonable for their high-level structure to be ⟨NP⟩ → ⟨NP⟩ ⟨VP⟩.

62. *Lǚshì chūnqiū*, "Wángdào" 《呂氏春秋》王道.
63. *Shǐ jì*, "Xià běnjì" 《史記》夏本紀.

### 3.2.3. ⟨NP⟩ ::= (⟨VP⟩)-⟨NP⟩

There is one common semantic application of this subordinate structure:

(94) Semantic Application: ⟨NP⟩ characterized by ⟨VP⟩

The word *adjective* is often used to describe a modifying ⟨VP⟩; a better term is *attributive* 'naming attributes'. Semantically, this ⟨VP⟩ may be a verb of state or of action — it may be a simple V or something more complex. There are already examples of this common structure in 92–93′, so I limit myself to a few simple cases:

(95) *sīlì* 私利 ⟨NP⟩ → (V)-N 'private advantage'[64]

(96) *guìrén* 貴人 ⟨NP⟩ → (V)-N 'person of high rank'

(97) *tónglèi* 同類 ⟨NP⟩ → (V)-N 'the same kind'

(98) *nìlín* 逆鱗 ⟨NP⟩ → (V)-N 'reversed fish-scales'

### 3.2.4. ⟨NP⟩ ::= ⟨NP⟩ ⟨VP⟩

The ⟨NP⟩ ⟨VP⟩ structure is most common as a whole ⟨S⟩ of topic-comment or subject-predicate structure, but such a whole ⟨S⟩ can also serve as either ⟨NP⟩ or ⟨VP⟩, depending on context. As ⟨VP⟩ it is discussed in Section 3.3.3. A whole topic-comment structure as ⟨NP⟩ has been mentioned briefly, in connection with examples 92′–93′, but a more general semantic application is:

(99) Semantic Application: object of some ⟨VP⟩

In example 100, a whole topic-comment unit can serve as the object of a verb:

(100) X若Y ⟨S⟩ → ⟨NP⟩ ⟨VP⟩ 'X is like Y'
*shì* ... *ruò* ... 視X若Y ⟨VP⟩ → ⟨VP⟩ ⟨NP⟩ → V ⟨NP⟩ ⟨VP⟩ 'to see that X is like Y'

視子弟與臣若其身、惡施不慈[65]
If you regard your juniors or your ministers as being like your own self, how would you impose any lack of kindness on them?

64. All four examples are from *Hán Fēizǐ*, "Shuōnán" 《韓非子》說難.
65. *Mòzǐ*, "Jiān'ài, shàng" 《墨子》兼愛上. Reading *wū* 惡 'how?'.

It is simplest to see X若Y as a whole sentence, embedded intact as the object of *shì* 視.

The most common application of this structure is the so-called *pivot structure*. A pivot noun is one that is simultaneously the object of one verb and the subject of another. We observe only a small number of specific verbs introducing this structure, and there are too few of them to gather them into their own syntactic category. The main ones are:

| | | |
|---|---|---|
| (101) | *lìng* 令 | V ‘to command’ |
| | 令⟨NP⟩⟨VP⟩ | “to have/let/make ⟨NP⟩ ⟨VP⟩” |
| | *lìng wángzǐ jū* 令王子居 | ⟨VP⟩ → $V_1$ $NP_{1=2}$ $V_2$ ‘to have the Prince live at’ |

願令王子居於堂上[66]
I would like to have the Prince stay in the hall.

In 101, *wángzǐ* 王子 is semantically the object of *lìng* but also the subject of *jū* 居; for the time being, I symbolize that dual role by writing $NP_{1=2}$.[67] Similarly, in 102 *gōngzǐ* 公子 is the object of *shǐ* and the subject of *sǐ* 死.

| | | |
|---|---|---|
| (102) | *shǐ* 使 | V ‘to send’ |
| | 使⟨NP⟩⟨VP⟩ | “to have/let/make ⟨NP⟩ ⟨VP⟩” |
| | *shǐ gōngzǐ sǐ* 使公子死 | ⟨VP⟩ → $V_1$ $NP_{1=2}$ $V_2$ ‘to have the Executive’s son die alone’ |

吾不能生而使公子獨死矣[68]
I can’t let the Executive’s son die alone while I stay alive.

If the third person pronoun is the pivot noun, it appears as *zhī* 之, which is used for semantic direct and indirect objects:

| | | |
|---|---|---|
| (103) | 令之⟨VP⟩ | “to have/let/make him/her/etc. ⟨VP⟩” |
| | *lìng zhī sǐ* 令之死 | ⟨VP⟩ → ⟨VP⟩ ⟨NP⟩ → $V_1$ $N_{1=2}$ $V_2$ ‘to have him/her/etc. die’ |

66. *Lǚshì chūnqiū*, “Xiàoxìng lǎn, Shǒushí” 《呂氏春秋》孝行覽·首時.
67. Traditionally, the graph 令 in the sense ‘command; to command’ represented a word pronounced differently from the word ‘to have someone do something’ written with the same character and presumably related historically: *lìng* vs. **líng*; the latter reading is no longer current.
68. *Hán shī wàizhuàn*, 9 《韓詩外傳》九.

文王聞之、喟然而嘆、故拘之牖里之庫百日、欲令之死[69]
When the Cultured King heard that, he felt weary and sighed. And so they held him for a hundred days in the keep of Enticement Village, hoping to have him die.

(104) 使之⟨VP⟩ "to have/let/make him/her/etc. ⟨VP⟩"
*shǐ zhī jū* 令之居 ⟨VP⟩ → ⟨VP⟩ ⟨NP⟩ → $V_1$ $N_{1=2}$ $V_2$ 'to have him/her/etc. dwell at…'

子謂薛居州善士也、使之居於王所[70]
You say Sngat Kratu is a fine person, and would have him take up residence in the King's chamber.

That is in keeping with the fact that the pivot noun is a direct object.

But it is misleading to write $V_1$ $N_{1=2}$ $V_2$ because it makes N seem ambiguous. This structure is better written ⟨VP⟩ ⟨NP⟩ → ⟨$VP_1$⟩ N ⟨$VP_2$⟩, where N ⟨$VP_2$⟩ is the object of ⟨$VP_1$⟩.

In some cases there is no explicit pivot noun, and the second verb itself is the object of the first, by ⟨NP⟩ ::= ⟨VP⟩:

(105) 令⟨VP⟩ "to cause ⟨VP⟩-ing"
*lìng jū wàizhái* 令居外宅 ⟨VP⟩ → ⟨VP⟩ ⟨NP⟩ → ⟨$VP_1$⟩ ⟨$VP_2$⟩ "to cause dwelling in an outer residence" ☞ 'to have [him/etc.] live in a residence outside'

王奇孝材能、佩之王印、號曰將軍、令居外宅、多給金錢[71]
The King was amazed at Qhruhs's talent and bestowed the royal seal on him, gave him the title General, had him live in a dacha outside the city walls, and supplied him with much gold and copper coin.

### 3.2.5. ⟨NP⟩ ::= ⟨VP⟩

This is the formal statement of the fact that what appears to be a verb may also function as a noun. Examples, especially with single-word verbs, appear earlier in the paper. However, semantic verb phrases are often longer than semantic noun phrases, and it is

69. *Shǐ jì*, "Lǔ Zhònglián Zōu Yáng lièzhuàn" 《史記》魯仲連鄒陽列傳.
70. *Mèngzǐ*, "Téng wén gōng, xià" 《孟子》滕文公下.
71. *Shǐ jì*, "Huáinán Héngshān lièzhuàn" 《史記》淮南衡山列傳.

common for them to serve as topics in topic-comment structure, which means that it is effective to treat them as noun-phrases syntactically. Some examples follow.

In 106, 事親 'to serve one's parents' and 為大 'to consider great' are both V N, and each reduces to ⟨VP⟩. But the whole expression 事親為大 is semantically ⟨NP⟩ ⟨VP⟩ rather than ⟨$VP_1$⟩ ⟨$VP_2$⟩, because "serving one's parents," as the name of an action, functions semantically as a noun phrase.

| | | |
|---|---|---|
| (106) | *shìqīn* 事親 | ⟨VP⟩ → V N 'to serve one's parents' |
| | *wéi dà* 為大 | ⟨VP⟩ → ⟨$VP_1$⟩ ⟨$VP_2$⟩ → V N 'to consider great' |
| | *shìqīn wéi dà* 事親為大 | ⟨S⟩ → ⟨NP⟩ ⟨VP⟩ → ⟨$VP_1$⟩ ⟨$VP_2$⟩ → $V_1$ $N_1$ $V_2$ $N_2$ "as for serving one's parents, [we] consider it greatest" ☞ 'to serve one's parents is the greatest act' |

事孰為大、事親為大、守孰為大、守身為大[72]
"Serving whom do we consider the greatest? Serving one's parents we consider the greatest. Taking care of whom do we consider the greatest? Taking care of oneself we consider the greatest." ☞ Who is it the most important to serve? Your parents. Who is it the most important to watch carefully? Your own self.

When an expression with clear verbal character serves as the ⟨NP⟩ element, the mnemonic translation "when ⟨NP⟩ takes place, ⟨VP⟩ takes place" may be useful. For instance:

| | | |
|---|---|---|
| (107) | *shí* 食 | V 'to eat' |
| | *bùyǔ* 不語 | ⟨VP⟩ → 不 V 'not to converse' |
| | *shí bùyǔ* 食不語 | ⟨S⟩ → ⟨NP⟩ ⟨VP⟩ → ⟨$VP_1$⟩ ⟨$VP_2$⟩ → $V_1$ 不 $V_2$ "when eating takes place, not conversing takes place" ☞ 'not to converse while eating' |

食不語、寢不言、…… 寢不尸、居不容[73]
He would not converse while eating or speak while lying in bed. … While lying in bed, he was not corpselike; while hanging around the house, he would not bother about how he looked.

---

72. *Mèngzǐ*, "Lí Lóu, shàng" 《孟子》離婁上.
73. *Lúnyǔ*, "Xiāngdǎng" 《論語》鄉黨.

Similarly:

| | | |
|---|---|---|
| (108) | *shàngjiāo* 上交 | ⟨VP⟩ → $(V_1)$-$V_2$ 'to interact upwardly' |
| | *bùchǎn* 不諂 | ⟨VP⟩ 'not to be sycophantic' |
| | *shàngjiāo bùchǎn* 上交不諂 | ⟨S⟩ → ⟨NP⟩ ⟨VP⟩ → ⟨$VP_1$⟩ ⟨$VP_2$⟩ → $(V_1)$-$V_2$ 不 $V_3$ 'when interacting upwardly takes place, not being sycophantic should take place' |

上交不諂、下交不驕、則可以有為矣[74]
If you avoid being sycophantic when you have dealings with your superiors, and being arrogant when you have dealings with your subordinates, you may be able to accomplish something.

When the noun phrase is semantically verbal, the following verb phrase (the second element) may also be a complement, so called because it completes the first element, by describing the extent of the action or state of the first element, or a consequence of first element. Continuing the mnemonic model above, it may be useful to express this as "when ⟨NP⟩ takes place, its extent or consequence is ⟨VP⟩."

| | | |
|---|---|---|
| (109) | *shǒusǐ* 守死 | ⟨S⟩ → ⟨NP⟩ ⟨VP⟩ → ⟨$VP_1$⟩ ⟨$VP_2$⟩ → $V_1$ $V_2$ "when holding fast takes place, the extent is dying" ☞ 'to hold fast until death' |

守死善道[75]
to hold fast until death, and to improve the course of moral advancement

| | | |
|---|---|---|
| (110) | 不舍晝夜[76] | ⟨S⟩ → ⟨NP⟩ ⟨VP⟩ → ⟨VP⟩ ⟨NP⟩ → ⟨VP⟩ $N_1$ $N_2$ "when not casting [it] aside, the extent is day and night" ☞ 'it is not cast aside, day or night' |

Example 109, despite being V N, does not mean "to hold death fast"; nor does example 110 mean "not to cast day or night aside." "Day or night" is also not an adverbial expres-

74. Yáng Xióng, *Fǎyán*, "Xiūshēn" 《揚子法言》修身. *Shàngjiāo* 上交 and *xiàjiāo* 下交 could also be (N)-V: 'to interact in an upward/downward direction'.
75. *Lúnyǔ*, "Tàibó" 《論語》泰伯.
76. *Lúnyǔ*, "Zǐ hǎn" 《論語》子罕.

sion, even though it appears to be that in the final translation, It would have to be in the syntactic position of adjunct (preceding the verb phrase) to have adverbial semantics.

| | | |
|---|---|---|
| (111) | *juéjǐng* 掘井 | ⟨NP⟩ 'well-digging' → ⟨VP⟩ → V N 'to dig a well' |
| | *jiǔrèn* 九軔 | ⟨VP⟩ 'to be nine *nerses*' → ⟨NP⟩ → $N_1$ $N_2$ "nine *nerses*" |
| | 掘井九軔 | ⟨S⟩ → ⟨NP⟩ ⟨VP⟩ → ⟨VP⟩ ⟨NP⟩ → V $N_1$ $N_2$ $N_3$ "when digging a well, its extent is nine *nerses*" ☞ 'to dig a well nine *nerses* deep' |

掘井九軔而不及泉、猶為棄井也[77]
If you don't reach a spring when you dig a well nine *nerses* deep, you might just as well be abandoning the well.

| | | |
|---|---|---|
| (112) | *wénzhī* 聞之 | ⟨$VP_1$⟩ → V N 'to hear it' |
| | *gānxīn* 甘心 | ⟨$VP_2$⟩ → V N 'to make sweet the heart' |
| | *wénzhī gānxīn* 聞之甘心 | ⟨VP⟩ → ⟨NP⟩ ⟨VP⟩ → ⟨$VP_1$⟩ ⟨$VP_2$⟩ → $V_1$ $N_1$ $V_2$ $N_2$ "on hearing it, the result is to make glad the heart" ☞ 'to be delighted to hear it' |

天子既好宛馬、聞之甘心[78]
Since the Son of Heaven loved horses from Ionian Fergana, he was delighted at the news.

### 3.2.6. ⟨NP⟩ ::= N

A non-terminal noun may consist of nothing but a terminal noun, which cannot be analyzed further. How can we be confident the solitary noun serves as a whole utterance? The easiest way is when it appears in quoted text:

(113) 春秋曰、荊[79]
"The *Springs and Autumns* says 'Brambles'."

---

77. *Mèngzǐ*, "Jìnxīn, shàng" 《孟子》盡心上. One *ners*, Mandarin *rèn* 軔%仞, is a measure of depth some 80 "inches" [*cùn* 寸] long in the Zhōu.
78. *Hàn shū*, "Xíngfǎ zhì" 《漢書》刑法志.
79. *Chūnqiū fánlù*, "Juéguó" 《春秋繁露》爵國.

☞ The *Springs and Autumns* uses the name Brambles [ — meaning the state of Srhah — in describing an action by Srhah].

(114) 何謂六水、曰、河水赤水遼水黑水江水淮水[80]
What does Six Rivers refer to? It means the C'gaj, T'qhrak, Rew, Hmek, Krong, and Gwrij Rivers.

(115) 王曰、叟[81]
The King said, "Venerable sir!"

(116) 人何以知道、曰、心[82]
How can people understand the Way? [Through] the heart.

As usual, it is also possible to understand these nouns as functioning verbally; for instance:

(115′) 王曰、叟
… "You are a venerable gentleman."

(116′) 人何以知道、曰、心[83]
… [It is the action of] the heart.

### 3.3. Definitions of ⟨VP⟩ and Their Semantics

#### 3.3.1. ⟨VP⟩ ::= ⟨$VP_1$⟩ ⟨$VP_2$⟩

Semantically, this coordinate structure is one of conjunction. It means the actions or states of the verb phrases occur (or at least are named) in series:

(117) Semantic Application: ⟨$VP_1$⟩ and ⟨$VP_2$⟩

The following example illustrates it; the actions *háo* 號 'to wail' and *qì* 泣 'to weep' both take place, presumably in that order.

80. *Huáinán zǐ*, "Dìxíng xùn" 《淮南子》墜形訓.
81. *Mèngzǐ*, "Liáng Huì wáng, shàng" 《孟子》梁惠王上.
82. *Xúnzǐ*, "Jiěbì" 《荀子》解蔽.
83. *Xúnzǐ*, "Jiěbì" 《荀子》解蔽.

(118) *háoqì* 號泣 ⟨VP⟩ → $V_1$ $V_2$ ‘to wail and to weep’

趙武號泣[84]
Mrah of clan Drewh wailed and wept.

### 3.3.2. ⟨VP⟩ ::= ⟨VP⟩ ⟨NP⟩

Semantically, this structure is straightforward:

(119) Semantic Application: verb-object

The only catch is that both nouns and verbs may serve as the ⟨NP⟩ object. Here *object* is a lazy choice of words. Semantically, it may be a direct object — in which case it receives or bears the action of the verb phrase. Or, again semantically, it may be an indirect object, in which case the action of the verb phrase is passed to it. The difference between these two situations is only important when both are explicitly present in the same phrase, or when a disambiguating particle (*yǔ* 與, *yǐ* 以) appears.

Examples of verb plus direct object are common. As for indirect objects:

(120) *yǔzhī* 與之 ⟨VP⟩ → V N ‘to give [it] to him/her’

今乞所以養母、是說夫子之義也、必與之[85]
Now he is begging for something to serve his mother with. That is a righteous act that would please your Master, and you must by all means give it to him.

(121) *sǐjūn* 死君 ⟨VP⟩ → V N ‘to die for one’s ruler’
*sǐfù* 死父 ⟨VP⟩ → V N ‘to die for one’s father’

臣死君而眾人死父[86]
Ministers die for their lords and the masses die for their fathers.

Semantics and context are what allow us to call the nouns *zhī* 之, *jūn* 君, and *fù* 父 indirect rather than direct objects; the syntax does not specify that semantic relationship.

When there are two nouns after the verb, the order is usually indirect object first and direct object second, as shown in Figure 3:

---

84. *Xīnxù*, “Jiéshì” 《新序》節士.
85. *Lǔshì chūnqiū*, “Jìdōng jì, shìjié”《呂氏春秋》季冬紀·士節. Reading 說％悅 *yuè* ‘pleased; to please’
86. *Chūnqiū fánlù*, “Wǔxíng xiāng shēng” 《春秋繁露》五刑相生.

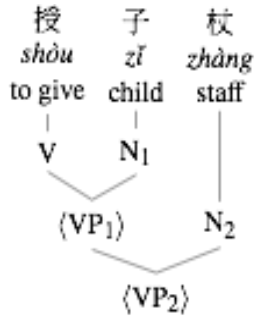


Fig. 3: Composition of Two Cases of Operation ⟨VP⟩ → ⟨VP⟩ ⟨NP⟩

(122) *shòu zǐ* 授子 ⟨$VP_1$⟩ → V $N_1$ 'to give to sons'
*shòu zǐ zhàng* 授子杖 ⟨$VP_2$⟩ → ⟨$VP_1$⟩ $N_2$ 'to give ceremonial staffs to sons'

三日授子杖、五日授大夫杖、七日授士杖[87]
On the third day, you give the sons the ceremonial staffs; on the fifth day, you give the grandees the staffs; on the seventh day, you give scholars the staffs.

We can understand this syntax as *composition* or *chaining* (successive application; see Sec. 4.2.3) of binary operations. As shown in Figure 3, diagramming *shòuzǐ zhàng* 授子杖 V $N_1$ $N_2$ of example 122, the first binary operation combines *shòu* 授 and *zǐ* 子 V $N_1$ as a verb phrase ⟨$VP_1$⟩; the second binary operation combines 授子 ⟨$VP_1$⟩ $N_2$ into a more complex ⟨$VP_2$⟩.

The semantic direction of the verb's action is not necessarily toward the first noun:

(123) *duó zhī* 奪之 ⟨$VP_1$⟩ → V $N_1$ 'to take from them'
*duó zhī shí* 奪之食 ⟨$VP_2$⟩ → ⟨$VP_1$⟩ $N_2$ 'to take food from them'

毀其備、散其積、奪之食[88]
Destroy their preparations; scatter what they have stored up; take from them their food.

Sometimes the noun phrase is not a true semantic object at all, but merely an associated noun, with which the action or state of the verb has a relationship.

87. *Lǐ jì*, "Sāngfú sìzhì" 《禮記》喪服四制.
88. *Guǎnzǐ*, "Xiǎowèn" 《管子》小問.

(124) *ānrén* 安仁 ⟨VP⟩ → V N ‘to be tranquil vis-à-vis humaneness’

仁者安仁[89]
The humane are tranquil in their humaneness.

(125) *xiān* 先N ⟨VP⟩ → V N ‘to be ahead of N’
*dòngxīn* 動心 ⟨VP⟩ → V N ‘to move the mind’
*xiān* 先⟨NP⟩⟨VP⟩ ⟨VP⟩ → ⟨VP⟩ ⟨NP⟩ → ⟨$VP_1$⟩ ⟨$VP_2$⟩ → $V_1$ $N_1$ $V_2$ $N_2$ “to have priority over $N_1$ in respect to $V_2$ $N_2$” ☞ ‘to $V_2$ $N_2$ earlier than $N_1$’

是不難、告子先我不動心[90]
That isn’t difficult. Sir Kuk learned to keep his mind from moving before I did.

This is once again a matter of semantics, not syntax. Wenyan makes no syntactic distinction among the different semantic functions of a noun phrase following a verb phrase.

It is sometimes said that verbs of motion have a different syntactic relationship to nouns of place than other verbs, but I think that is a semantic impression; it seems to me that the syntax is indistinguishable — even without invoking the notion of the obliquely-related noun, the place is the target of the action of the verb, as any other noun would be:

(126) *qùqí* 去齊 ⟨VP⟩ → V N ‘to leave Dzei’

孟子去齊[91]
Mencius left Dzei.

(127) *shìwèi* 適衛 ⟨VP⟩ → V N ‘to go to Gwrats’

子適衛[92]
The Master went to Gwrats.

89. *Lúnyǔ*, “Lǐrén” 《論語》里仁.
90. *Mèngzǐ*, “Gōngsūn Chǒu, shàng” 《孟子》公孫丑上.
91. *Mèngzǐ*, “Gōngsūn Chǒu, xià” 《孟子》公孫丑下.
92. *Lúnyǔ*, “Zǐlù” 《論語》子路.

The object of a verb of motion can be an abstract destination or thing abandoned, too, rather than a place:

(128) *shì shān'è* 適善惡 ⟨VP⟩ → V ⟨NP⟩ 'to move toward good or bad'

氣也者、所以適善惡之馬也與[93]
As for a someone's overall state of being, isn't it a "horse" by which that person moves toward good or bad?

(129) *qiānshàn* 遷善 ⟨VP⟩ → V N 'to move toward the good'

民日遷善而不知為之者[94]
The masses move toward the good every day, yet they don't know that they are doing so.

(130) *yíwù* 遺物 ⟨VP⟩ → V N 'to leave things behind'
*lírén* 離人 ⟨VP⟩ → V N 'to get away from people'

似遺物離人而立於獨也[95]
It was as though you were standing in a solitary state, having put aside all the things and people of the world.

Special applications of verb-object syntax, expressing the semantics of what are commonly called passive and auxiliary verbs, appear in Sections 4.3.2 and 4.3.3, below.

### 3.3.3. ⟨VP⟩ ::= ⟨NP⟩ ⟨VP⟩

This is a topic-comment structure, the commonest combination of elements at the level of the whole sentence. Topic-comment is semantically quite flexible, but it is what we usually associate with subject-predicate sentences in conventional grammars. This structure appears in Section 3.2.4 as ⟨NP⟩, but here it has a distinct semantic application that is verbal:

(131) Semantic Application: topic-comment, with object as topic

The conditions under which topic-comment functions verbally is when the topic is semantically the object, meaning that the verb acts on it. It means essentially the same thing

93. Yáng Xióng, *Fǎyán*, "Xiūshēn" 《揚子法言》修身.
94. *Mèngzǐ*, "Jìnxīn, shàng" 《孟子》盡心上.
95. *Zhuāngzǐ*, "Tián Zǐfāng" 《莊子》田子方.

as if the noun and verb were reversed, as in ⟨VP⟩ → ⟨VP⟩ ⟨NP⟩ (Section 3.3.2). Topic-comment semantics places focus (semantic emphasis) on whatever appears in the topic position.

(132)

| | |
|---|---|
| *gùjiù* 故舊 | ⟨NP⟩ → ⟨VP⟩ → $V_1$ $V_2$ ‘old friends, people one has known in the past’ |
| *bùyí* 不遺 | ⟨VP⟩ ‘for abandonment not to occur’ |
| *gùjiù bùyí* 故舊不遺 | ⟨VP⟩ → ⟨NP⟩ ⟨VP⟩ “as for those one has known in the past, abandonment does not occur” |

故舊不遺、則民不偷[96]
If you do not abandon your old allies and colleagues, then the people will not grow lazy in their interpersonal relationships.

I use the awkward translation “for abandonment not to occur” in order to approximate what I have described above as the default “middle voice” of Chinese verbs.

(133)

| | |
|---|---|
| *měishì* 每事 | ⟨NP⟩ ‘every matter’ |
| *wèn* 問 | V ‘to inquire’ |
| *měishì wèn* 每事問 | ⟨NP⟩ ⟨VP⟩ “as for every matter, to inquire” ☞ ‘to inquire about everything’ |

入大廟、每事問[97]
When he entered the great temple, he asked about everything.

This semantic application takes a form that many instructors treat as a special case: when a verb is negated and the noun is a one-syllable pronoun or question word, this pattern is common.

(134)

| | |
|---|---|
| *bù wǒ zhī* 不我知 | ⟨VP⟩ ‘not to know me’ |

無曰不我知也[98]
Don’t go around saying “No one knows about me!”

96. *Lúnyǔ*, “Tàibó” 《論語》泰伯.
97. *Lúnyǔ*, “Xiāngdǎng” 《論語》鄉黨.
98. *Dà Dài Lǐ jì*, “Zēngzǐ zhìyán, shàng” 《大戴禮記》曾子制言上.

(135) *wèi zhī néng yì* 未之能易 ⟨VP⟩ ‘not to be able to change it’

我未之能易也[99]
I have not been able to change it.

But pronouns and question words are not distinct parts of speech in terms of their gross syntax, and in the context of other ⟨NP⟩ ⟨VP⟩ expressions, this behavior is not exceptional. Compounds such as those in examples 136–137 appear to be lexicalized ⟨NP⟩ ⟨VP⟩ topic-comment expressions of this kind:

(136) *héyǐ* 何以 ⟨VP⟩ “as for what? by means of it” ☞ ‘how?’

(137) *shìyòng* 是用 ⟨VP⟩ “as for this, because of it” ☞ ‘for this reason’

A ⟨NP⟩ ⟨VP⟩ topic-comment expression that has become fully fossilized as a grammar pattern, with the noun now lexicalized as a particle, is:

(138) *suǒ* 所 V ⟨NP⟩ “(the object of V)” ☞ ‘that which is V-ed’, ‘the person who is V-ed’, (etc.)

*Suǒ* 所 literally means ‘place’, and that allows us to consider 所 V as ⟨NP⟩ ⟨VP⟩ with topic-comment semantics. If “place,” standing for the object of the verb, is the origin of the expression, then it has been grammaticalized. As mentioned in Rule of Thumb 64, the meaning of 所 V is a noun phrase.

### 3.3.4. ⟨VP⟩ ::= (⟨$VP_1$⟩)-⟨$VP_2$⟩

Here the relationship between ⟨$VP_1$⟩ and ⟨$VP_2$⟩ is subordinate. The semantic application has diverse senses:

(139) Semantic Application: ⟨$VP_1$⟩ modifies the main element ⟨$VP_2$⟩:
⟨$VP_2$⟩ in a ⟨$VP_1$⟩ manner or way
⟨$VP_2$⟩ under the circumstances of ⟨$VP_1$⟩
given that / granted that ⟨$VP_1$⟩, ⟨$VP_2$⟩
because/though/despite/when/if ⟨$VP_1$⟩, ⟨$VP_2$⟩

Examples are straightforward. Here is a coordinate pair of subordinate ($V_1$)-$V_2$ expressions:

99. *Lǐ jì*, “Tángōng, shàng” 《禮記》檀弓上.

(140) *fàng* 放 V ‘to let loose’
*fàn* 飯 V ‘to eat; to feed’
*liú* 流 V ‘to flow’
*chuò* 歠 V ‘to sip, drink’

*fàngfàn liúchuò* 放飯流歠 ⟨VP⟩ → ⟨$VP_1$⟩ ⟨$VP_2$⟩ → ($V_1$)-$V_2$ ($V_3$)-$V_4$
“to eat in a ‘letting go’ way and to drink like a torrent flowing” ☞ ‘to eat and drink with abandon’

放飯流歠而問無齒決[100]
While eating and drinking with abandon, they ask about the saying “don’t cut off a piece of food with your teeth.”

### 3.3.5. ⟨VP⟩ ::= (⟨NP⟩)-⟨VP⟩

Semantically, ⟨NP⟩ modifies the action or state of ⟨VP⟩ in this subordinate structure:

(141) Semantic Application: ⟨NP⟩ modifies the action or action of ⟨VP⟩:
to ⟨VP⟩ in a manner characterized by ⟨NP⟩
to ⟨VP⟩ like ⟨NP⟩ does
to ⟨VP⟩ using ⟨NP⟩

(142) *bǐshuō* 筆說 ⟨VP⟩ → (N)-V “to argue with pen”
→ ‘to make written arguments’

安危之際、文人不與、無能建功之驗、徒能筆說之效也[101]
At moments of danger to the state, intellectuals do not take part; they lack the proven ability carry out acts of merit, and are able only to produce results by written argument.

(143) *lìxíng* 力行 ⟨VP⟩ → (N)-V “to act using strength”
→ ‘to exert strength or effort’

好學近乎知、力行近乎仁、知恥近乎勇[102]
Love of learning is near to wisdom; exerting effort is near to goodness; having a sense of shame is near to bravery.

---

100. *Mèngzǐ*, “Jìnxīn, shàng” 《孟子》盡心上.
101. *Lùnhéng*, “Chāoqí” 《論衡》超奇.
102. *Lǐ jì*, “Zhōngyōng” 《禮記》中庸. Or arguably *lìxìng* V N ‘to aply strength to one’s actions’.

(144) *ròushí* 肉食 ⟨VP⟩ → (N)-V “to eat in a way characterized by meat” ☞ ‘to eat meat’

Let me digress on this last case, a set phrase with an idiomatic figurative use:

(144′) *ròushí zhě* 肉食者 ⟨NP⟩ “meat-eater” ☞ ‘wealthy or powerful official’

肉食者謀之、又何間焉[103]
If the Meat-Eaters plan something, how can anyone else intervene?

Examples like 144 can usually also be understood in a number of different ways, syntactically. One alternate syntax is N V (Section 3.3.3). Consider how that analysis alters 144:

(144″) *ròushí* 肉食 ⟨VP⟩ → N V “as for meat, to eat it” ☞ ‘to eat meat’ or ‘meat is eaten’

We can also read it as $(N_1)$-$N_2$, semantically a noun phrase (Section 3.2.2):

(144‴) *ròushí* 肉食 ⟨NP⟩ → $(N_1)$-$N_2$ “food characterized by flesh.” ☞ ‘meat as food’

We can even read it as verb-object semantically (Section 3.3.2):

(144⁗) *ròushí* 肉食 ⟨VP⟩ → V N “to have meat as one’s food” ☞ ‘to eat meat’

Here is an example of (N)-V that is interesting because it shows how a construction of this sort is negated.

(145) *chǐjué* 齒決 ⟨VP⟩ → (N)-V “to sever using the teeth” ☞ ‘to bite off’

濡肉齒決、乾肉不齒決[104]
Marinated meat you bite off; dried meat you do not.

We are not obliged to use one passage to validate our interpretation of another; part of the gift of context-free grammar is the liberty to embrace ambiguity— with Harbsmeier, we

103. *Zuǒ zhuàn*, “Executive Zhuāng, year 10” 《左傳》莊公十年.
104. *Lǐ jì*, “Qǔlǐ, shàng” 《禮記》曲禮上.

should see Chinese as structurally "supple" or "indeterminate," not "loose" or "vague."[105] As in example 145, we sometimes find examples like these negated as whole V phrases:

(146) 不擇賢以托其身、不力行以自定[106]
They don't select wise advisors on whom to rely; they don't exert their strength to fix their own goals.

(147) 哀公射而中稷、其口疾不肉食[107]
The Sorrowing Executive hit the idol of the harvest god when shooting. The Executive's mouth became so afflicted that he could not eat meat.

Adverbial functionality — modifying the action or state of a verb — is associated with the structure (⟨$VP_1$⟩)-⟨$VP_2$⟩, as well as (⟨NP⟩)-⟨VP⟩. ⟨VP⟩ and ⟨NP⟩ are not distinct when they are adjuncts to a verb. They differ only in how close each appears to be, semantically, to noun sense as against verb sense overall, and the question boils down to which is easier for us to apply the meanings of the words in question to. In practice, it is usually not difficult to decide whether to treat a subordinate expression as noun or verb.

There is, however, a different problem with the (N)-V structure — as mentioned in connection with examples 144″–144‴, it is often unclear whether the noun is subordinate or a topic. The reason for this ambiguity is that Chinese verbs are indeterminate as to agency and voice.

Below is a fuller example of the ambiguity. First, reading the nouns as subordinate to the verbs that follow them:

(148)

| | |
|---|---|
| *mù* 木 | N 'tree' |
| *yǒng* 雍 | V (%壅) 'to corral' |
| *qiāng* 槍 | N 'spear; sharpened stake' |
| *lěi* 纍 | V 'to accumulate, build up' |

| | |
|---|---|
| *mùyǒng qiānglěi* 木雍槍纍 | ⟨VP⟩ → ⟨$VP_1$⟩ ⟨$VP_2$⟩ → ($N_1$)-$V_1$ ($N_2$)-$V_2$<br>'to corral [wild animals] with saplings and build up [a fence of] sharpened stakes' |

105. Harbsmeier 1998:123–173.
106. *Kǒngzǐ jiāyǔ*, "Wǔyí jiě" 《孔子家語》五儀解.
107. *Shuōyuàn*, "Biànwù" 《說苑》辨物.

> 搤熊羆、攣豪豬、木雍槍纍、以為儲胥、此天下之窮覽極觀也[108]
> They catch bears and tie up wild boars; they corral them in with saplings and build fences around them with sharpened stakes, keeping them as domestic animals. This is the most extraordinary sight in all the world.

Now, reading the nouns as topics of the verbs:

(148′) *mùyǒng qiāngléi* 木雍槍纍 ⟨VP⟩ → ⟨$VP_1$⟩ ⟨$VP_2$⟩ → $N_1$ $V_1$ $N_2$ $V_2$
'for saplings to corral [wild animals]
and for sharpened stakes to be built up'

> 木雍槍纍
> Saplings corral them; around them sharpened stakes are built up.

Fortunately, syntactic ambiguity is not fatal to this manner of reading Wenyan; on the contrary, it liberates the reader to explore semantics.

### 3.3.6. ⟨VP⟩ ::= ⟨NP⟩

This is the formal statement of the fact that what appears to be a noun may also function as a verb. However, there are two other, more interesting cases, in which a noun phrase appears to be functioning verbally, because it is semantically the comment of a topic-comment structure or even takes, itself, an object:

(149) Semantic Application: topic-comment (equation):
to be ⟨NP⟩
to be the same as, non other than, actually ⟨NP⟩
to happen to the extent ⟨NP⟩
to make into, treat as, view as ⟨NP⟩

(150) Semantic Application: transitive verb
to make into, treat as, view as ⟨NP⟩

Noun phrases in series are not always conjoined, as in Section 3.2.1; sometimes they have topic-comment structure, as in Section 3.2.5, where the NP appears to function as VP. In contrast, the structure described in 149 identifies one noun phrase as being the same as (or having the extent of) another noun phrase, rather than using a semantic verb phrase to give information about it. But because the second noun phrase describes state, it is useful to analyze is as functioning as a verb phrase syntactically.

---

[108.] Yáng Xióng, "Chángyáng fù" 揚雄《長楊賦》.

(151) *gāozǔ* 高祖 ⟨NP⟩ 'the high founder'
*zhōngyáng lǐ rén* 中陽里人 ⟨NP⟩ 'a person of Mid-Sunlit Town'
*X Y rén* Ｘ Ｙ人 ⟨$NP_1$⟩ ⟨$NP_2$⟩ → ⟨NP⟩ ⟨VP⟩ 'X is a person of Y.'

高祖、沛豐邑中陽里人[109]
The High Founder was a person of Mid-Sunlit Town in Phekspfrong Town.

(152) *liáng xiàowáng* 梁孝王 ⟨NP⟩ 'the filial king of Rang'
*jǐngdì mǔ dì* 景帝母弟 ⟨NP⟩ 'the younger brother of the Esteemed Emperor's mother'
*X Y dì* Ｘ Ｙ弟 ⟨$NP_1$⟩ ⟨$NP_2$⟩ → ⟨NP⟩ ⟨VP⟩ 'X is younger brother to Y.'

梁孝王、景帝母弟[110]
The King titled "The Filial," of Rang, was the younger brother of the Esteemed Emperor's mother.

The first noun phrase in ⟨$NP_1$⟩ ⟨$NP_2$⟩ may also be marked with the topic-marker *zhě* 者:

(153) *tiānxià* 天下 ⟨NP⟩ → ($N_1$)-$N_2$ "the below-part vis-à-vis Heaven" ☞ 'the realm under Heaven, the whole world'

天下者、高祖天下.[111]
The Realm under Heaven is our Founder's Realm under Heaven.

But such examples are relatively rare; this meaning is normally expressed with a particle marking the second noun:

(154) ⟨$NP_1$⟩ ⟨$NP_2$⟩ 也 '⟨$NP_1$⟩ is ⟨$NP_2$⟩'

(155) ⟨$NP_1$⟩ 乃 ⟨$NP_2$⟩ '⟨$NP_1$⟩ is ⟨$NP_2$⟩'

109. *Shǐ jì*, "Gāo zǔ běnjì" 《史記》高祖本紀.
110. *Shǐ jì*, "Hán Chángrú lièzhuàn" 《史記》韓長孺列傳.
111. *Shǐ jì*, "Wèi Qí Wǔ Ān hóu lièzhuàn" 《史記》魏其武安侯列傳.

But we should not assume that the particles are fundamental and merely elided in sentences such as 151 and 152; each syntactic pattern must stand on its own, without being equated to some other pattern in order to make sense of it. Interestingly, making this meaning negative requires a verb — in other words, the syntax has to be ⟨NP⟩ ⟨VP⟩:

(156) ⟨$NP_1$⟩ 非 ⟨$NP_2$⟩ ⟨NP⟩ ⟨VP⟩ "⟨$NP_1$⟩ negates ⟨$NP_2$⟩"
☞ '⟨$NP_1$⟩ is not ⟨$NP_2$⟩'

The ⟨VP⟩ in this ⟨NP⟩ ⟨VP⟩ example is composed of a verb (normally *fēi* 非) and a noun, and is parallel to a different positive form:

(158) ⟨$NP_1$⟩ 為 ⟨$NP_2$⟩ ⟨NP⟩ ⟨VP⟩ "⟨$NP_1$⟩ acts or serves as ⟨$NP_2$⟩"
☞ '⟨$NP_1$⟩ is ⟨$NP_2$⟩'

It is not impossible that the nouns describing place of origin and kinship in examples 151 and 152 are functioning as verbs in these sentences, rather than nouns:

(159) X 人 ⟨VP⟩ "to be in the role of a person of X"
☞ 'to be a person of X'

(160) X 弟 ⟨VP⟩ "to be in the role of younger brother to X"
☞ 'to be younger brother to X'

Why? Because there are some examples where the second ⟨NP⟩ consists of two apparent ⟨NP⟩ joined by *ér* 而, as shown in examples 161–162. *Ér* normally appears between verbal expressions, not nouns, so if *zǐ* 子 and *dì* 弟 are functioning verbally there, when they are the heads of paired phrases sandwiching *ér*, they could also be functioning verbally even when they appear in a single phrase, as in 151–152. The translation need not change, but our understanding of syntax must. In fact, there are examples of *dì* 弟 in the corpus that are negated with *bù* 不 and therefore decisively verbal ("decisively," again, within the context of the noun-verb contrast being a useful fiction).

(161) *zhōu wǔwáng zǐ* 周武王子 ⟨VP⟩ 'to be son to the Martial King of Tiw'

| | | |
|---|---|---|
| | *chéngwáng dì* 成王弟 | ⟨VP⟩ 'to be younger brother to the Accomplished King' |

晉唐叔虞者、周武王子而成王弟[112]
Youngson Ngwra, of clan Nrhang of Tsins, was son to the Martial King of Tiw and they were younger brothers to the Accomplished King.

| | | |
|---|---|---|
| (162) | *zhōu wénwáng zǐ* 周文王子 | ⟨VP⟩ 'to be son to the Cultured King of Tiw' |
| | *wénwáng dì* 武王弟 | ⟨VP⟩ 'to be younger brother to the Martial King' |

管叔鮮蔡叔度者、周文王子而武王弟也
Youngson Sar, of clan Kwanh, and Youngson Dak, of clan Srhats, were sons to the Cultured King of Tiw and they were younger brothers to the Martial King.

One curious structure has a verb phrase functioning as a noun phrase and a noun phrase functioning as a verb phrase. Semantically, it appears to be comment-topic — that is, ⟨VP⟩ ⟨NP⟩, with the comment given semantic focus because it appears at the start of the sentence. But I suggest that topic-comment is a preferable way to understand it.

| | | |
|---|---|---|
| (163) | ⟨VP⟩ ⟨NP⟩ | "⟨NP⟩ ⟨VP⟩!" |
| | 勉勉 | ⟨VP⟩ 'diligent and motivated' |
| | 我王 | ⟨NP⟩ → $N_1$ $N_2$ 'our king' |

勉勉我王、綱紀四方[113]
"How diligent and motivated! — our king. He brings order to all the various parts of the kingdom." ☞ Truly our king is diligent and motivated. …

The ⟨NP⟩ topic generally takes on verbal quality; I schematize such a reading mnemonically as: "⟨VP⟩, and it is ⟨NP⟩ to which ⟨VP⟩ applies." This effect is heightened by the frequent use of *yě* 也 after the ⟨NP⟩ (although here I have chosen examples without *yě*).

This structure is most common with a disambiguating particle between ⟨VP⟩ and ⟨NP⟩:

112. *Shǐ jì*, "Jìn shìjiā" 《史記》晉世家.
113. *Shījīng*, "Dàyǎ, Yùpǔ"《詩經》大雅·棫樸.

(164) ⟨VP⟩ 哉 ⟨NP⟩ "⟨NP⟩ ⟨VP⟩!"

君子哉若人[114]
"He is a gentleman! — this man." ☞ What a gentleman this man is.

(165) ⟨VP⟩ 矣 ⟨NP⟩ "⟨NP⟩ ⟨VP⟩!"

久矣吾不復夢見周公[115]
"Long it has been! — my not having seen the Executive of Tiw again in dreams." ☞ I haven't seen the Executive Tiw again in my dreams in a very long time.

The meaning of the noun can also function transitively (Semantic Application 150), taking another noun as object, in what Yuen Ren Chao called a putative verbal sense:[116]

(166) *rén* 人 N 'person'; V 'to treat or view as a person'

接練時、錄母之變、始人之也、不言氏姓、貶之也 [117]
For the text to record the mother's untoward act, at the Receiving the Soul and the White Silk sacrifices, is to begin treating her correctly as a person. Not naming her clan is to demean her.

Putative semantics is the same as other transitive or causative verbal usage. For instance:

(167) *huǒ* 火 N 'fire'; V 'to apply fire to'
*lú* 盧 N (%盧) 'shack, cottage'; V 'to make into a dwelling'

人其人、火其書、盧其居[118]
Turn [the Sangha's] members into normal people; burn their books; turn their residences into normal homes.

---

114. *Lúnyǔ*, "Gōngyě Cháng" 《論語》公冶長.
115. *Lúnyǔ*, "Shù ér" 《論語》述而.
116. Chao 1968b, sec. 8.1.3[2]: "Many adjectives are also transitive verbs in a causative sense, that is, they tell how something is to be made so; and a few in a putative sense, that is, they indicate that something is thought or found to be so." See also sec. 8.1.5(2) and 8.1.7(Vc 6).
117. *Gǔliáng zhuàn*, "Executive Zhuāng, year 1"《穀梁傳》莊公元年.
118. Hán Yù "Yuán dào" 韓愈《原道》.

The semantic application is important, but it is not necessary to restrict that application to the putative alone.

### 3.3.7. ⟨VP⟩ ::= V

A non-terminal verb phrase may also consist of nothing but a terminal verb, which cannot be analyzed further. It is easiest to be confident that the solitary verb serves as a whole utterance when it appears in quoted text:

(168) *rán* 然 V "to be thus" ☞ 'Yes. That is so; that is correct.'

對曰、然[119]
He replied, "Yes."

(169) *nuò* 諾 V "to assent" ☞ 'Okay, I agree to your demand.'

平原君笑應曰、諾[120]
The Lord of the Level Plains laughed and replied, "All right, then."

(170) *yǒu* 有 V "to exist; there is" ☞ 'Yes, such a thing does exist or happen.'

孟嘗君曰、有[121]
Lord Eldest-son C'dang said, "Yes, there is."

(171) *fǒu* 否 V "to deny" ☞ 'No, I disagree.'

使者曰、否[122]
The messenger said, "No."

119. *Shuōyuàn*, "Jìngshèn" 《說苑》敬慎.
120. *Shǐ jì*, "Píngyuán jūn Yú Qīng lièzhuàn" 《史記》平原君虞卿列傳.
121. *Zhànguó cè*, "Qí 4, Mèng Cháng jūn zhú yú Qí ér fùfǎn" 《戰國策》齊四·孟嘗君逐於齊而復反.
122. *Hán shī wàizhuàn*, 10 《韓詩外傳》十.

## 4. Reflections

### 4.1. Premises of this Model

The premises on which this model stands are not usual among readers of Wenyan and may seem paradoxical. One is that it radically demotes the role of semantics; another is the need for what I will call "cloud-of-meaning" thinking.

#### 4.1.1. Limiting Semantics in Order to Explain Semantics

Part of speech is not intrinsic in Chinese words; it is determined by relationships between words. And yet, if we hazard a starkly primitive part-of-speech system, admitting no more than one semantic contrast, between verbs and nouns, it then becomes easy to explain a great deal of other semantic behavior, without having to assert any further semantic premises.

Ideas we can generally do without if we limit the application of semantics beyond a verb-noun contrast include:

- whether a verb's voice is necessarily active or passive;
- whether a verb is intrinsically transitive or intransitive;
- whether verbs of motion behave distinctly from other verbs;
- the distinction between verbs of state and of action;
- the distinction between a noun as subject or object of a verb;
- any regularly occurring parts of speech other than verbs and nouns;
- whether the verbal sense of a word is primary and the nominal sense derived, or the other way around;
- whether a particle is decisive or not in disambiguating confusing relationships between words.

The paradox is that by admitting no more than a tiny semantic contrast, we are then able to explain a great deal more semantic behavior.

It bears repeating that my initial premise of a distinction between nouns and verbs is a fiction. That means this entire construction is not falsifiable in the narrow sense. But I submit that, granted the initial fiction, it is a highly useful construction.

#### 4.1.2. The Cloud of Meaning

Another premise of this model — and like the contrast between noun and verb, it is a fiction — is that we view a character as representing one and only one particular word. All usage and senses of the character seem drawn from manipulation of that single word, and those senses together appear to fill out a mass of related meanings. I like to think of that mass of related meanings as a cloud: it has volume, in that the relationships among

senses are usually too complex to think of on a simple linear or planar model, and at the same time its boundaries are hazy.

This idea is not the same as relying on etymology to identify different historical senses of a word and choosing among them. Etymology attempts to represent historical fact and so is necessarily temporal. In Computer Science terms, it embodies a directed acyclic graph. It is a graph because it is made up of connected nodes (words and senses); the graph is directed because time progresses in only one direction as we proceed from node to node, and it is acyclic because, given the direction of time, we cannot return to any particular node after we have visited it on our progress through the graph. But a directed acyclic graph is a poor model for exposing the cloud of meaning in Chinese, particularly because in Chinese as I have described it, cycles are possible: when a verb is used as a noun in Wenyan, that does not mean the verb is historically prior to the noun. Writers seem to have felt able to assign parts of speech on the fly or even to leave them ambiguous. And representing that ambiguity is the great strength of the model described in this paper.

This premise has implications for lexicography. Just how to express the relationships among meanings is a problem without an ideal solution. Chinese native dictionaries have long tended to offer synonyms for any word that could not be defined explicity. Western-language dictionaries of Chinese cannot turn to synonyms but have tended to offer translations rather than definitions. Both groups of dictionaries have avoided specifying parts of speech.[123] Many Western works in the missionary-colonial period listed compounds haphazardly, but it is no longer remembered that the editors included those compounds as a way of delimiting the whole range of meanings of individual characters, rather than as a poor effort to document the whole vocabulary of compounds.[124]

Modern character dictionaries attempt to arrange definitions in order so that, as Paul Kroll puts it, they[125]

> may suggest a certain development of meanings or understandable progression from a basic sense to various derived meanings.

Kroll takes care to say that his dictionary — a sort of thesaurus of English translations of Chinese character-meanings, rather than a traditional dictionary — is not etymological,

123. Branner & Meng 2010a, 2010b, 2010c, 2013.

124. Branner & Meng 2013:559–61. Frederick W. Baller (1852–1922) and Robert Henry Mathews (1877–1970) both made this point explicitly in the prefaces of their works.

125. For instance, *Dà cídiǎn* 1985, *Hànyǔ dà zìdiǎn* 1986–90, Wáng Lì et al. 2000, Kroll 2015. The quotation is from Kroll 2015:x.

but his word *progression* describes the etymological model closely. And that is a different thing from a cloud of meaning, because a progression attempts to model actual history.

Development means a temporal sequence beginning from, normally, a single ancestor. That idea embodies a fallacy: that there is an ultimate, original meaning of every word, like some psychological motive hidden in a human being's personality, and whenever we use the word we call occultly on that original meaning. Take English *check* as an example. Essentially all living usage of *check* derives ultimately from the Persian word *šāh* 'king': from the name of an attack on the king in the game of chess, to the name *chess* itself, to the pattern of alternating squares seen on chessboards, to square boxes in which people confirm items completed or inspections passed, to the verb meaning "to inspect and confirm," and so forth. *Šāh* is indisputably the origin of *check*; whenever we say *check*, in a sense we are using a living avatar of ancestral *šah*. And yet the ways we normally use *check* are not altered a jot if we're not aware of its history.

The fallacy, then, is claiming we gain necessary information about usage by examining etymology. What seems lacking to me in Wenyan dictionaries is not semantic detail or even usage, but a sense of the interconvertibility of each morpheme's various syntactic functions. Supporting the achronic cloud of meaning requires a kind of organization different from a temporal progression.

Some hands — especially Lukaš Zádrapa — catalogue shifts in part of speech in various texts, defining narrowly the nature of the semantic changes involved. It seems to me, however, that precision here invites unnecessary prescriptiveness.[126]

The cloud of meaning is useful fiction, just like the contrast between noun and verb. It cannot really be the case that a character really always represents only a single word with varying senses, because at a minimum we know — from comparing different versions of the same text — that there is such a thing as a *loangraph*: a character that normally represents some known word, but sometimes represents an entirely different word, one more often written with some different character. Or we may see a character with multiple meanings that we cannot, even with great generosity, persuade ourselves all originate in a single ancestor. If one character can represent unrelated words, then we may easily fool ourselves into thinking we can connect etymologically the different uses of a single word, when what we are looking at is actually more than one different word written with the same character.

Exactly which characters represent which words is not the result of a natural evolution. Scribal and editorial custom played a part, during the early production and transmission of texts, as did official standardization later on (beginning no later than the Hàn). So

126. Zádrapa 2016.

it is a fallacy to think that the cloud of meaning we perceive in the usage of a given word is based on that word's natural development. Rather, we are looking at the result of conscious effort by literate editors to normalize Wenyan as an isolating language. We try to connect different uses of a single character as though that character were a single word. That is our best model for reflecting the semantic flexibility and the interchangeability of parts of speech that are the veritable signature of a stringently isolating language. In Wenyan we find that signature affixed almost everywhere.

### 4.2. Likeness to a Functional Programming Language

It may be surprising to think that we can compare Wenyan to a programming language, but that becomes possible and interesting once we have describe it as a symbolic system adhering to a context-free grammar. The comparison is interesting because it vivifies some of the abstract beauty of this natural language against a totally different product of people's minds.

What is the actual utility of representing Wenyan as a context-free grammar? Primarily to highlight its extraordinary simplicity, which I think has been obscured in previous studies. To recapitulate, the language has only two parts of speech, plus some special grammar words whose behavior is unique and therefore outside of the general rules of syntax. The syntactic relationships I describe in this paper are *atomic*: we perceive no composition of finer subprocesses to produce them. The key unary process is for a word to change its state from verb to noun or noun to verb. There are only two binary processes for combining parts of speech: in sequence or by subordinating the first to the second. All eight resulting combinations are meaningful and well attested in the literature, and all can be reduced to either verbal or nominal function overall.

I propose comparing Wenyan to a *functional programming language*, such as LISP or Haskell, because typologically those languages tend to be close to symbolic systems. They tend to have only a few distinct operations, which likens them to Wenyan as I have described it. Other characteristics, which require some explanation, are that functions are first-class entities, that state is generally not retained, and emphasis on the composition of operations.

#### 4.2.1. Functions as First-Class Entities (and Data as Function)

For the purposes of comparison, verbs are like functions, and nouns are like data. Because Wenyan verbs can always behave as nouns, they possess something called *first-class* status in the functions of a programming language: a function is not restricted from any of the behaviors of data, and can be handled just as data is.

A noun behaving verbally (⟨VP⟩ ::= ⟨NP⟩) has no easy parallel in functional languages, because data does not normally behave directly as a function. But the idea is

worth pursuing a little because it deepens the likeness between the two conceptual systems.

It is sometimes said that a hash table is an example of data behaving as a function; but a hash table is really a collection of arbitrary functions that include data as output. Better still is to view it a a single surjective function, not data per se.

Rather than comparing a noun behaving verbally to a hash table, more fitting parallel is the programming strategy in which data is always accessed as the return value of a function, in order to prevent the uncertainties that may arise when mutable objects are passed by reference. That is something like the way a noun behaves when it functions as a semantically intransitive verb, as in examples 92–93′ and 159–162: in example 161, we could say that *dì* 弟 functions verbally to "return" the semantic noun "younger brother."

Noun-as-verb behavior can also include semantic transitivity, as in examples 166–167, which seems to me a semantic version of type-coercion in programming languages: converting one type of data into another (in either its character or its representation). When Example 167 says *huǒ qí shū, lú qí jū* 火其書、廬其居 " 'fire' their books and 'cottage' their residences," it is changing the character of books and residences to fire and cottages. When Example 166 says *rén zhī* 人之 "treat her as a person," it is changing the character of the mother.

### 4.2.2. No Retention of State

As for not retaining state, it is a characteristic of functional programming languages that when complex process produces intermediate results, those results are not retained but are passed directly to the next steps in the process. *State* is the name of such intermediate results.

Wenyan words in a multi-word expression tend contribute to the overall meaning more or less in sequence, and generally without long gaps between the appearance of a word and the next reference to it. In particular, the language avoids the long embedded clauses of some styles of modern Chinese.

Note that Wenyan's complete lack of inflection actually prevents it from retaining very much state. Inflected languages, on the other hand, are able to support a long distance between words that are closely connected in building up the meaning of longer utterances.

### 4.2.3. Composition of Operations

Functional programming languages tend to have small numbers of operations, which are composed into more complex functions. The operations in Wenyan are unary or binary, but a group of more than two words is normally composed of smaller operations, so

that the group as a whole is a ⟨NP⟩ or ⟨VP⟩. This point is already illustrated in many places above, but there are a few interesting sidelights to be shone with more examples.

#### 4.2.3.1 Composing Three Elements

Even though Chinese prosody favors pairs, there are cases of triplets of things, especially when objects or qualities are listed:

(172) *niányuèrì* 年月日 ⟨NP⟩ → $N_1$ $N_2$ $N_3$ 'year, month, and day'

> 凡紀事言年月日者、詳悉重之也[127]
> Whenever it mentions the year, month, and day, in recording events, it explicitly always means the author considered it important.

When things are in series, we can think of them as an example of *nesting* of subordinate structures, or the *chaining* (successive accumulation) of binary operations, as illustrated in connection with Figure 3, above. That way, the atomic operation is still fundamentally binary, but it is carried out twice instead of once. See Figure 4:

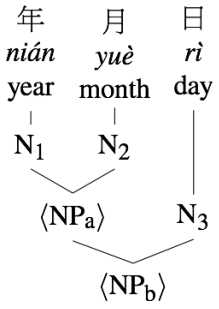


Fig. 4: Two Chainings of Binary Operation ⟨NP⟩ → ⟨$NP_1$⟩ ⟨$NP_2$⟩
(figure uses subscript *a* and *b* to avoid confusion with ⟨$NP_1$⟩ ⟨$NP_2$⟩)

Example 172′ revises the formula for *niányuèrì* 年月日 to match Figure 4:

(172′) *niányuèrì* 年月日 ⟨$NP_b$⟩ → ⟨$NP_a$⟩ ⟨$N_3$⟩ → $N_1$ $N_2$ $N_3$ 'year, month, and day'

Why not analyze this expression as in 172″?

---

127. *Lùnhéng*, "Zhèngshuō" 《論橫》正說.

(172″) *niányuèrì* 年月日 ⟨$NP_b$⟩ → ⟨$N_1$⟩ ⟨$NP_a$⟩ → $N_1$ $N_2$ $N_3$ ‘year, month, and day’

Because there is a tendency to order objects semantically from large to small, from more to less important, or in order of temporal occurrence.

Even though Chinese prosody favors pairs and quartets, triplets are known in Mandarin. For instance:[128]

(173) *gāodàshàng* 高大上 ⟨Adj⟩ (contraction of *gāoduān dàqì shàng dàngcì* 高端、大气、上檔次 ‘high-end, grand, and of superior grade’; said mainly of goods)

(174) *cōngjiāngsuàn* 蔥薑蒜 ⟨NP⟩ ‘scallions, ginger, and garlic’

(175) *dǎzáqiǎng* 打砸搶 ⟨VT⟩ ‘to beat, smash, and grab: to go on the rampage, to loot and pillage’

These triplets are regular lexical items and their constituents follow an idiomatic order, without an obvious large-to-small or temporal pattern.

Verbal triplets are also not unheard of in Wenyan, either with single morphemes or compounds:

(176) *tiāndìrén* 天地人 ⟨NP⟩ → $N_1$ $N_2$ $N_3$ ‘Heaven, Earth, and humankind’

> 通天地人曰儒[129]
> Being able to link the principles of Heaven, Earth, and humankind is called being a scholar.

*Tiāndìrén* is probably also in large-to-small semantic order.

128. The parts of speech in examples 173–175 are for Mandarin, not Wenyan.

129. Yáng Xióng, *Fǎyán*, “Jūnzǐ” 揚雄《法言》君子.

(177) *chuándào* 傳道 ⟨VP⟩ → V N ‘to transmit the truth’
*shòuyè* 受業 ⟨VP⟩ → V N ‘to receive [training in] a trade’
*jiěhuò* 解惑 ⟨VP⟩ → V N ‘to relieve ignorance’

師者所以傳道受業解惑也[130]
A teacher is the means for transmitting the truth, learning a trade, and dispelling ignorance.

Conjunction is the operation most commonly chained, but subordination is also seen, for instance:

(178) *qí* 齊 N ‘Dzei (a place)’
*huán* 桓 N ‘pillar’

*qíhuán* 齊桓[131]
[the person called] “The Pillar,” of the state of Dzei

(179) *gōng* 公 N ‘Executive (title)’

*qíhuán gōng* 齊桓公[132]
the Executive called “The Pillar,” of the state of Dzei

Figure 5 shows the structure of the expression as a tree:

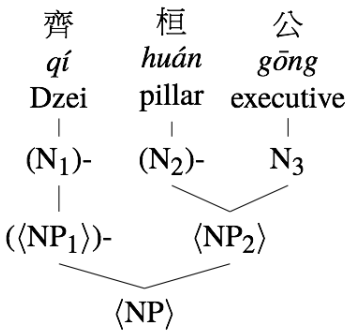


Fig. 5: Two Chainings of Binary Operation ⟨NP⟩ → (⟨$NP_1$⟩)-⟨$NP_2$⟩

The individual elements in *Qí Huán gōng* are, again, in large-to-small semantic order.

130. Hán Yù “Shīshuō” 韓愈《師說》.
131. *Shǐ jì*, “Chǔ shìjiā” 《史記》楚世家.
132. *Hàn shū*, “Gǔjīn rén biǎo” 《漢書》古今人表.

#### 4.2.3.2 Composing Four or More Elements

Chaining is not limited to three elements but the only common case is the successive conjunction of items in series. But complex chaining, and deeply nested subordination, are uncommon — Wenyan is, after all, processed by human minds, rather than automated parsers. It is the norm, however, for groups of four to be composed into two groups of two, in series. Traditionally, the aesthetic of the Chinese high-register language tends to favor dualistic prosody, in which atomic units of two combine to form larger paired units. Example 180 is such a case.

(180)

| | |
|---|---|
| *háoqì* 號泣 | ⟨VP⟩ → $V_1$ $V_2$ ‘to wail and weep’ |
| *bēi’āi* 悲哀 | ⟨VP⟩ → $V_1$ $V_2$ ‘to be sad and to sorrow’ |
| *háoqì bēi’āi* 號泣悲哀 | ⟨VP⟩ → ⟨$VP_1$⟩ ⟨$VP_2$⟩ → $V_1$ $V_2$ $V_3$ $V_4$ ‘to wail and weep, and to be sad and sorrowing’ |

號泣悲哀奔走道路[133]
They wail, weep, mourn, as they rush on the road (to the funeral).

Here is composition of four ⟨NP⟩ elements built up into a single ⟨NP⟩:

(181)

| | |
|---|---|
| *gān* 肝 | N ‘liver’ |
| *dǎn* 膽 | N ‘gallbladder’ |
| *chǔ* 楚 | N ‘Srhah (a place)’ |
| *yuè* 越 | N ‘Gwat (a place)’ |
| *gāndǎn chǔyuè* 肝膽楚越 | ⟨NP⟩ → ⟨$NP_1$⟩ ⟨$NP_2$⟩ → $N_1$ $N_2$ $N_3$ $N_4$ ‘liver and gallbladder, Srhah and Gwat’ |

自其異者視之、肝膽楚越也[134]
If you view them in terms of their differences, [they are as different as] the liver [is from] the gallbladder, [or as] the state of Srhah [is from] the state of Gwat.

Example 182 is a series of four ⟨VP⟩ expressions — three examples and then a summary line explaining the overall principle:

133. *Báihǔ tōng*, “Bēnghōng” 《白虎通》崩薨.
134. *Zhuāngzǐ*, “Déchōng fú”《莊子》德充符.

(182) 制禮作教 They formalized social relationships as an example.
立法設刑 They instituted laws and arranged punishments.
動緣民情 In the way they took action, they remained in keeping with the social conditions of their people.
而 *And in all this,*
則天象地[135] they took Heaven and Earth as their models.

It is possible to have a great many ⟨VP⟩ expressions in series. ⟨NP⟩ expressions in series are less common, but at times there are long catalogues of things — for instance, in *fù* 賦 celebrating (or satirizing) the opulence of a ruler's private parks.

### 4.3. What about Rules Seen in Other Models of Wenyan Grammar?

Wenyan is imbued with a certain mystery by many readers, and with mystery comes inordinate attention to particular syntactic patterns, which seem to call for special handling. The model in this paper makes it unnecessary to turn to special rules to explain the most familiar cases of special handling. One such case is the pivot noun, discussed above in Section 3.2.4; another is the pre-posing of pronouns, discussed at the end of Section 3.3.3; a third is putative verbal semantics, discussed at the end of Section 3.3.6. Below I discuss four other such cases: verb-object inversion, passive verbs, auxiliary verbs, and intrinsic adverbs.

#### 4.3.1. Verb-Object Inversion

Some hands advocate a pattern called "inversion," in which a verb and its object appear reversed: under this principle, N 之 V is read as though it meant "V N." For instance, *cǐ zhī wèi* 此之謂 (a common phrase in classical texts) is explained as *wèicǐ* 謂此 'to talk about this, to mean this [the foregoing]'.

Tactically (that is, as a mnemonic), reversing the substantial words this way is useful if it helps clarify semantics: that *zhī* 之 the object of the verb *wèi* 謂. Syntactically, though, *wèi* is functioning not as a verb here but as a noun: the structure of *cí zhī wèi* is actually ($N_1$ 之)-$N_2$ rather than (N 之)- V:

(183) *cǐ zhī wèi* 此之謂 ⟨NP⟩ → $N_1$ *zhī* $N_2$ 'the saying or meaning of this thing' (rather than *謂此 ⟨VP⟩ 'to mean this')

135. *Hàn shū*, "Xíngfǎ zhì" 《漢書》刑法志.

書曰、維齊非齊、此之謂也[136]
The *Documents* says, "Make balanced what is not balanced." That [quotation] is an expression of this [the aforementioned idea].

It is usually no problem to translate as though the grammar were ⟨VP⟩ → ⟨VP⟩ ⟨NP⟩ — in this case, "it means this [aforementioned idea]" — but syntax is ⟨NP⟩ → (⟨NP⟩)-⟨NP⟩.

Examples are plentiful. Below are two more, enough to make the point quickly:

| | | |
|---|---|---|
| (184) | *hé qīng* 何卿 | ⟨NP⟩ 'what high official?, which high official?' |
| | *wèn* 問 | V 'to inquire' → N 'inquiry' |
| | *héqīng zhī wèn* 何卿之問 | ⟨NP⟩ → (⟨$NP_1$⟩ *zhī*) ⟨$NP_2$⟩ 'the act of asking about which ministers?' (rather than *問何卿 ⟨VP⟩ → ⟨VP⟩ ⟨NP⟩); but translate as "to inquire about which high officials?" |

齊宣王問卿、孟子曰、王何卿之問也[137]
The King of Dzei, entitled "The Vast," inquired about high officials. Mencius said, "Your Majesty's is an inquiry about which high officials?"
☞ "About which high officials is Your Majesty inquiring?"

| | | |
|---|---|---|
| (185) | *kuāngqiè* 筐篋 | ⟨NP⟩ 'special box for storing valuable cloth or documents' |
| | *cáng* 藏 | V 'to store' |
| | *kuāngqiè zhī cáng* 筐篋之藏 | ⟨NP⟩ → ($NP_1$ *zhī*) $N_2$ 'the storing [of things] in a special box' (rather than *藏 筐篋 ⟨VP⟩ → ⟨VP⟩ ⟨NP⟩); but translate as "to store [things] in a special box" |

約者有筐篋之藏、然而行不敢有輿馬[138]
Misers have the habit of storing valuables in special boxes, and so when they go out they don't dare to go [long distances] by carriage or horse.

136. *Xúnzǐ*, "Wángzhì" 《荀子》王制.
137. *Mèngzǐ*, "Wàn Zhāng, xià" 《孟子》萬章下.
138. *Xúnzǐ*, "Róngrǔ" 《荀子》榮辱. Graph 藏 could also be *zàng* 'treasury', a noun.

What is needed to make sense of expressions like these is a flexible mindset toward parts of speech and how parts of speech affect meaning; no special grammar rule is needed.

### 4.3.2. Passives

It is widely taught that any verb is passive if it follows certain other verbs, especially:

(186) *jiàn* 見 'to see, meet with'

(187) *wéi* 為 'to make; to act as; to become; to consider'

(188) *kě* 可 'to be possible'

(189) *zú* 足 'to be sufficient'

and a few others. These cases already appear in the first full native grammar of Wenyan, that of Mǎ Jiànzhōng 馬建忠 (1845–1900).[139] As with inversion, this is really an observation about semantics — and it is frequently effective in guiding translation — but it is not an observation about syntax. Wenyan syntax does not actually show grammatical voice distinctions, and it is hard to imagine they would turn up in just a few special environments. The reality is that Wenyan has only what might be called a middle voice, in which passive and active are not distinguished explicitly. Rather than admitting a special passive voice in Wenyan, it is simpler to draw on patterns already well established.

#### 4.3.2.1. *Jiàn* 見 with Following Verb as Object

For *jiàn* 見, what appears to be a following verb is an object semantically, which is to say a noun syntactically:

(190) *jiàntú* 見屠 ⟨VP⟩ → V N "to meet with slaughtering"
☞ 'to be slaughtered'

> 齊趣下三國、不且見屠[140]
> Dzei must submit at once to the three powerful states, or else it will get slaughtered.

---

139. Mǎ distinguishes the behaviors of these various particles; see Mǎ 1954:204–5 section "惟以「為」字先於外動者"; 1954:207–11 sections "四，以「見」「被」等字加於外動之前者" and "五，「可」「足」兩字後動字，概有受動之意."

140. *Shǐjì*, "Qí Diàohuì wáng shìjiā"《史記》齊悼惠王世家. Reading *fǒu qiě* 不且 "if not, then as an alternative" ☞ 'otherwise'; reading *cù* 趣%促 'at once'.

In other words, this is a straightforward example of ⟨VP⟩ → ⟨VP⟩ ⟨NP⟩ (Section 3.3.2).

#### 4.3.2.2. *Wéi* 為 with [⟨NP⟩] ⟨VP⟩ as Object

The seeming passive construction with *wéi* 為 takes a semantic verb phrase or a whole semantic subject-predicate expression as object (Section 3.2.4–3.2.5). The essential mnemonic translation for it is "the situation becomes one of [⟨NP⟩] ⟨VP⟩," drawing on the meaning of *wéi* 為 as "to become."

There may or may not be an explicit semantic "agent" of the verb following *wéi*:

(191) *qín* 禽 — V 'to capture'
*wéi qín* 為禽 — ⟨VP⟩ → ⟨VP⟩ ⟨NP⟩ → V N "to become [a case of] capture" ☞ 'to be captured'

[夫差]使人請食於越、越王弗與、乃攻之、夫差為禽[141]
[P'tshaj] sent people to ask Gwat for food, but the King of Gwat would not give them any and attacked, instead, and P'tshaj got captured.

(192) *wéi qín qín* 為秦禽 — ⟨VP⟩ → $V_1$ $NP_{1=2}$ $V_2$ "to become [a case of] Dzin capturing" ☞ 'to be captured by Dzin'

韓氏之兵非削弱也、民非蒙愚也、兵為秦禽、智為楚笑[142]
Gar's troops were not weak and its people were not stupid, but the troops were captured by Dzin and its wisest heads were made laughingstocks of by Srhah.

We may be tempted think of the noun as a modifier of the second verb:

(192′) *wéi qín qín* 為秦禽 — *⟨VP⟩ → ⟨VP⟩ ⟨NP⟩ → $V_1$ (N)-$V_2$ "to become [a case of] capturing through the agency of Dzin" ☞ 'to be captured by Dzin'

The argument against this otherwise attractive analysis has to do with production and is therefore out of scope here: when the noun is represented by a pronoun, that pronoun is *zhī* 之. Now, *zhī* as a pronoun is normally only the object of a verb; it does not normally function as a modifier. An example:

---

141. *Lǚshì chūnqiū*, "Xiàoxìng lǎn, chánggōng" 《呂氏春秋》孝行覽·長攻.
142. *Zhànguó cè*, "Hán 1, Qín Hán zhàn yú zhuózé" 《戰國策》韓一·秦韓戰於濁澤.

(193) *wéi zhī qín* 為之禽 ⟨VP⟩ → $V_1$ $NP_{1=2}$ $V_{22}$ "to cause it to become him (etc) capturing" ☞ 'to be captured by him (etc.)'

敵人執數、動則就陰、以虛應實、必為之禽[143]
Suppose the enemy holds the advantage in numbers. If you move, stay in shadow. If you respond to his substance with your own lack of substance, you will surely be captured by him.

Overall, the *wéi* 為 structure displays all the same patterns of behavior as the so-called pivot structure (Section 3.2.4). Semantically it is hard to see *wéi* as a pre-pivot verb, for reasons that, once more, have to do with production and so are outside the scope of this analysis.[144] That, however, is a lexical limitation of *wéi* itself as a verb; *lìng* and *shǐ*, too, do not have exactly the same meaning as solitary verbs that they do as when they introduce a pivot structure. The *wéi* structure has become a special pattern, and perhaps we should think of *wéi* in it as a grammaticalized verb. That may be why it developed a special disambiguation pattern. Consider the following example and its later explanation:

(194) *kùn* 困 V 'to beset, hem in, hold back'
*wéi jiǔ kùn* 為酒困 ⟨VP⟩ → ⟨VP⟩ ⟨NP⟩ → $V_1$ $N_{1=2}$ $V_2$ "for the situation to become 'liquor holds [one] back' " ☞ 'to be badly affected by liquor'

喪事不敢不勉、不為酒困[145]
In conducting rites for the dead, never dare fail to exert yourself fully. And do not get addled by strong drink.

(195) *wéi jiǔ suǒ kùn* 為酒所困 ⟨VP⟩ → ⟨VP⟩ ⟨NP⟩ → V (⟨$NP_1$⟩)-⟨$NP_2$⟩ "to become someone that liquor holds back" ☞ 'to be badly affected by liquor'

143. *Huáinán zǐ*, "Bīnglüè xùn" 《淮南子》兵略訓.
144. Because unlike with *lìng* 令 and *shǐ* 使, we cannot truncate after the noun and retain part of the sense of the sentence.
145. *Lúnyǔ*, "Zǐ hǎn" 《論語》子罕.

> 《論語》不為酒困、《註》言不爲酒所困而及亂也[146]
> *Analects*: Do not get addled by strong drink. *Comment*: It means to reach the point of losing control, as a result of getting addled by strong drink.

Use of the particle *suǒ* 所 to disambiguate *wéi* [⟨NP⟩] ⟨VP⟩ is not syntactically passive any more than is the original structure, since *suǒ* V is a noun phrase, meaning the object of V (see 64 and 138). After *wéi*, the construction [⟨NP⟩] *suǒ* V is a noun phrase: ⟨NP⟩ → (⟨$NP_1$⟩)-⟨$NP_2$⟩.

#### 4.3.2.3. *Kě* 可 and *Zú* 足 with Following Verb as Object

In the cases of *kě* 可 and *zú* 足, a following verb phrase is semantically their object:

(196)

| | |
|---|---|
| *dào* 道 | V ‘to say’ |
| *kě* 可⟨VP⟩ | ⟨VP⟩ → ⟨VP⟩ ⟨NP⟩ ‘to make or consider ⟨VP⟩ possible’ |
| *zú* 足⟨VP⟩ | ⟨VP⟩ → ⟨VP⟩ ⟨NP⟩ ‘as make or consider ⟨VP⟩ sufficient’ |
| *kědào* 可道 | ⟨VP⟩ → ⟨VP⟩ ⟨NP⟩ → $V_1$ $V_2$ “to consider saying possible” ☞ ‘to be possible to say’ (rather than *‘to be possible to be said’) |

> 君子則不然、言思可道、行思可樂、德義可尊、作事可法、容止可觀、進退可度、以臨其民[147]
> The gentleman, on the other hand, is not like that. In what he has to say, he considers whether it is possible to say it; in his actions, he considers whether they are possible to take pleasure in. His moral power and his sense of obligation it is possible to put forward as respectable; it is possible to imitate them his initiatives and his service. His bearing one may gaze upon; one may measure him against the way he steps forward or withdraws himself. And in all these things he superintends his people.

(197)

| | |
|---|---|
| *chēng* 稱 | V ‘to praise’ |
| *zúchēng* 足稱 | ⟨VP⟩ → ⟨VP⟩ ⟨NP⟩ → $V_1$ $V_2$ “to consider praising sufficient” ☞ ‘to be worth praising’ (rather than *‘to be sufficient to be praised’) |

146. *Kāngxī zìdiǎn*, “Wéibù 4” 《康熙字典》口部四·困古文朱.
147. *Xiàojīng*, “Shèngzhì” 《孝經》聖治. Reading *duó* 度.

> 管仲蒙恥辱以存亡不足稱也[148]
> The fact that Middleson Kwanh suffered humiliation in order to preserve himself from death is not praise-worthy.

Note that there seem to be no cases of 見 or 為, supposed passive markers, appearing in connection with these supposedly passivizing verbs 可 and 足.

To repeat the main point, the ⟨VP⟩ following *jiàn* 見, *wéi* 為, *kě* 可 , *zú* 足, and other such cases can still be translated with a passive construction in Mandarin or English, but there is no reason to make that a feature of syntax, rather than semantics.

### 4.3.3. Modal Auxiliaries

A small number of words function in a way that makes them look like modal auxiliary verbs in translation. The commonest of them — *yù* 欲 'to desire to; to desire', *néng* 能 'to be able to; to enable', *dé* 得 'to be able, manage to; to obtain, succeed at', *yuan* 願 'to wish for, will' — are well attested as ordinary verbs with transitive semantics. In this semantic "auxiliary" sense, they simply take as objects the verb phrases that follow them.

A few cases are not otherwise known with transitive semantics, such as *kěn* 肯 'to be willing to' and *gǎn* 敢 'to dare to; to make bold to'. But they are well attested with intransitive semantics, and it seems a short leap to model them behaving the same way as *yù*, *néng*, and the others above.

An exceptional case is *qí* 其 'to be likely to, to tend to; would', which may have been a true modal auxiliary at some point, but which seems to have ceased to be productive early on and is normally seen in fossilized idioms in the Hàn and later. Because it is a unique case, I think it is best treated as a particle.

### 4.3.4. The Question of Intrinsic Adverbs

As noted in Definition 22 and Sections 3.3.4–5, adverbial semantics is expressed by subordinating a noun or verb to a verb phrase. But are there any words that are intrinsically adverbial? If so, they appear to be too rare to constitute a regular semantic category.

Below I consider three kinds of adverbs: intensive adverbs (for meanings akin to "very" and "most"), some common examples of general adverbs, and "totalizers" (akin to Mandarin *dōu* 都 'all').

The main intensive adverbs are *shèn* 甚, *zhì* 至, and *jìn* 盡; of these, *zhì* 至 and *jìn* 盡 are already well attested as verbs: 'to reach' and 'to come to an end'. Is *shèn* 甚 in-

148. *Yántiě lùn*, "Lùnrú" 《鹽鐵論》論儒. Reading *cúnwáng* 存亡 as V N "to preserve being lost" ☞ 'to salvage preservation from his lost state'.

trinsically adverbial? No; it is well attested in verbal (and also nominal) usage, for instance:

(198) *shèn* 甚 V ‘to be severe’

雖有過、其不甚矣[149]
Even if there were a transgression, it would not be severe.

(199) *shèn* 甚 V → N ‘a case of being “very much so”’

稽首、服之甚也[150]
Kowtowing is an intense case of submission.

Among common adverbs, *bì* 必 ‘necessarily’ and *gū* 姑 ‘tentatively, transiently’ are representative. Are they intrinsically adverbial? No, both are attested in verbal usage:

(200) *bì* 必 V ‘to be necessary; to treat as necessary’

聖人以必不必、故無兵、眾人以不必必之、故多兵[151]
The sage takes the necessary to be unnecessary, and so uses no troops. Ordinary people take what is unnecessary and “necessitize” it, so they have lots of troops.

(201) *gū* 姑 V ‘to pause at’

姑純懿之所廬[152]
I rest transiently in the place where pure virtue resides.

Wenyan has a number of functioning like Mandarin *dōu* 都, a “totalizer” particle meaning “entirely, altogether, all the foregoing.” Like *dōu*, they usually appear between topic and comment. Most are verbs whose meanings can be related to the sense “all”:

(202) 總/統/齊/全/備 “to encompass as a whole” ☞ ‘all’

(203) 專/一/並/共 “to act together, to combine” ☞ ‘all’

149. *Lǐ jì*, “Biǎojì” 《禮記》表記.
150. *Lǐ jì*, “Jiāo tèshēng” 《禮記》郊特牲.
151. *Zhuāngzǐ*, “Liè yùkòu” 《莊子》列御寇.
152. Zhāng Héng, “Sīxuán fù” 張衡《思玄賦》.

(204) 盡/畢 “to use up, complete” ☞ ‘all’

(205) 通/徧 “to reach everywhere” ☞ ‘all’

Other *dōu*-like words — *xián* 咸, *jiē* 皆, *xī* 悉, *jù* 俱 — lack obvious verbal sense. Are any of them intrinsically adverbial? No; all of these have straightforward verbal usage:

(206) *xián* 咸 V “to act together” ☞ ‘to remain in harmony’

故和聲入於耳、而藏於心、心億則樂、窕則不咸、總則不容[153]
So when harmonious sounds enter the ear, they are stored away in the heart. If the heart is at ease, the listener is happy. If the sounds are quiet, they do not reach everywhere; and if they are overwhelming, they cannot be taken in.

(207) *jiē* 皆 V ‘to reach everywhere’[154]
*kǒng* 孔 V ‘to be large in scale or scope’

為酒為醴、烝畀祖妣、以洽百禮、降福孔皆[155]
Make ale, make sweet wine; make offering to ancestors male and female; meet thereby the hundred ritual obligations; may they send down good fortune, and may it reach everywhere.

(208) *xī* 悉 V ‘to count up, tally up, get an overview of’

今王自行、悉國中武力以伐齊、而子胥諫不用[156]
When the King set out in person and combined all the military strength within his state in order to attack Dzei, Sir Sra’s remonstrations were not taken into account.

153. *Zuǒ zhuàn*, “Zhāo, year 21” 《左傳》昭公二十一年.

154. Kǒng Yǐngdá 孔穎達, “Máo shì zhuàn shù 毛氏傳疏”: 皆、徧也、與偕通. Graphs 偕 and 皆 are historically homophones (古諧切 {kei-2b}, pointing to Mandarin *jiē*), but recent Mainland dictionaries read the graph 偕 as *xié* (*Xiàndài Hànyǔ cídiǎn*; *Xīnhuá zìdiǎn* 新華字典 1971). *Xié* 諧 {ghei-2b} may be exerting influence. Examples of this morpheme written 偕 in received texts are clear as V “to act together.” For instance: 故事與時並、名與功偕 [And so this work (of conducting ritual) blends in with different times; its good effects match its names] (*Lǐ jì*, “Yuè jì” 《禮記》樂記).

155. *Shījīng*, “Zhōu sòng, Chén gōng zhī shí, Fēngnián” 《詩》周頌·臣工之什·豐年.

156. *Shuōyuàn*, “Zhèngbiàn” 《說苑》正諫.

(209) *jù* 俱 V ‘to come together into a group, interact; to accompany’[157]

偏則風、俱則雷[158]
With [*yīn* and *yáng*] unbalanced, the wind blows; when they interact, it thunders.

All these examples are verbal. Intrinsic adverbs, if they do exist, must be rare.

### 4.4. Multiple Readings

Chinese may not always have been isolating, but the writing system looks as though it was designed, and later standardized, to write an isolating language. That is true even at an early period. Even if we eventually uncover evidence that morphology was incorporated systematically into the script at one point, as seems elusively possible, it is significant that no such practice survived into later use. I have suggested that Chinese may have undergone a *restructuring* event early in its history, shearing away inflectional and derivational morphology to leave an isolating language — presumably because it was adopted by large numbers of new speakers.[159]

But there is a paradox connected with reconstructed morphology. The primary native evidence of it is multiple readings associated with a single graph, which includes readings that we understand to reflect semantic differences, together with *cognates* or *doublets*.[160]

---

[157] The graph 俱 is now standardly read *jù*, but it is useful to retain the traditional reading ***jū***, to reinforce the fact that semantically it and *jù* 具 are quite different words. 具 {guoH-3c} → *jù* ‘to possess and fully display’; also: ‘tool’; 俱 {kuo-3c} → ***jū***. Is there an etymological relationship between *jù* 具 and ***jū*** 俱? In the Baxter-Sagart system, *jù* 具 ***[g](r)o-s*** may be understood as ***m-k(r)o-s***, from ***jū*** 俱 ***k(r)o***, by affixation of ***-$s_3$*** for outwardly directed action and perhaps also ***$m_{1a}$-*** for volitional-causative sense. Baxter & Sagart 2014:54–55, 59, 2015.

[158] *Dà Dài Lǐ jì*, “Zēngzǐ tiān yuán” 《大戴禮記》曾子天圓.

[159] Branner 1999:159–66.

[160] A term common in modern Chinese philology, *tóngyuán zì* 同源字 “graphs of common origin” (Wáng Lì 1982:3–12), appears to render English *cognate*, but in practice it normally applies to what we would call *doublets*: words descending from a single ancestor at different stages of history but coexisting in a single language. The name doublet implies especially “the coexistence of [an] old form with the innovated form”; over time they often become become either differentiated as to function or meaning or else reconciled through blending and amalgamation (Hock 1991:169, 189–91). In contrast, English *cognate* includes words in different languages that descend from a single word in a common ancestral language; they are “genetically equivalent” (Fox 1995:62). Apparent doublets are so numerous in Chinese that Karlgren used them for inter-

The majority of the doublets that are applied to reconstructing early Chinese morphology involve the interchange of verb and noun, a relationship traditionally called *sìshēng biéyì* 四生別義 "distinguishing meaning by tone category." The paradox is that apparently all Chinese verbs and nouns are able to undergo this interchange, not only in received texts (which may have been redacted) but also in excavated bronze inscriptions (which could not have been). Yet only a fraction of them are associated with doublet readings in the received scholia and are not reflected in the organization of the received script.

If morphology was productive and widely known at one time, then the tradition must have been lost early enough that most of it did not find its way into scholia. And yet there must have been a continuous history of people being able to read. Without knowledge of morphology, they must have been reading Wenyan as an isolating language, just as we do, and they must have been manipulating parts of speech more or less as described here.

*Sìshēng biéyì* is commonly called "derivation by tone change" in English, but it is less tendentious to term a set of alternate readings like these a *relationship* or *alternation* and avoid referring to *change* or *derivation*, because the existence of these readings does not imply temporal progression from one to the other. we use primarily those doublets that are related to one another in a systematic way; derivation is temporal, while interchange is not. The proposed reconstructions began coexisting at some point and it may be that they have always coexisted, as part of an organic derivational system.

Like many other contemporary readers, I have mused aloud about the possibility that the ancient term *yǎyán* 雅言 'elegant speech' meant a high diglossic register that, coexisting with lower registers, reflected its own distinct regional origin.[161] But perhaps "elegant speech" was in fact a tradition of reading more in keeping with the structure of the writing system, and perhaps even morphologically more explicit than what has come down to us in fragmented and reliquary form in the scholia.

### 4.5 Counterbalancing and "Part-of-Speech" Yoga

I end with one of the particular pleasures of reading Wenyan: something a colleague and I once described as "syntactic yoga": the free manipulation of parts of speech that is possible in Chinese.[162]

The lines in example 182, above, display *counterbalancing* (*duìchèn* 對稱): the author has composed four-character verb phrases in series, but in their internal syntax as we

---

nal reconstruction, adopting the phrase *word family* to describe them (Karlgren 1933). Chinese renders that expression *tóngyuán zìzú* 同源字族.

[161.] Branner 2000:89–94.

[162.] Branner & Meng 2010a.

are describing it they are not in perfect parallel. An uncomplicated reading of the first couplet happens to be in perfect parallel:

| | | |
|---|---|---|
| (210) | 制禮作教 | ⟨VP⟩ → ⟨$VP_1$⟩ ⟨$VP_2$⟩ → $V_1$ $N_1$ $V_2$ $N_2$<br>They formalized social relationships and made an example of them. |
| | 立法設刑 | ⟨VP⟩ → ⟨$VP_1$⟩ ⟨$VP_2$⟩ → $V_1$ $N_1$ $V_2$ $N_2$<br>They instituted laws and arranged punishments. |

Each pair of words is a verb-object compound, and the whole couplet has the same syntax in the first line as in the second, all the way down to the level of individual words.

The second couplet is not in perfect parallel:

| | | |
|---|---|---|
| (210′) | 動緣民情 | ⟨VP⟩ → ⟨VP⟩ ⟨NP⟩ → ($V_1$)-$V_2$ ($N_1$)-$N_2$<br>In the way they took action, they remained in keeping with the social conditions of their people. |
| | 則天象地[163] | ⟨VP⟩ → ⟨$VP_1$⟩ ⟨$VP_2$⟩ → $V_1$ $N_1$ $V_2$ $N_2$<br>… They took Heaven and Earth as their models. |

Line one of 210′ is verb-object, while line two is a pair of verb-object expressions. Verb phrases in lengthy series are common in high-register writing, but internally — at the level of compounds or individual words — they are not necessarily counterbalanced.

But the distinction between Wenyan verbs and nouns is fluid — a word may be made to change its type at the reader's whim, and sometimes it is possible to recast syntactically unbalanced lines into "good" form — that is, reading them as more closely counterbalanced.

Here is another example:

| | | | |
|---|---|---|---|
| (211) | 常民 | | *Ordinary folk* [*have*] |
| | | 文杯畫案 | decorated cups, painted tables — |
| | | 机席緝⟦屨⟧ | armrests, seat-mats, close-stitched insoles — |
| | 婢妾 | | *Womanservants and concubines are* |
| | | 衣紈履絲 | garbed in fine silk, beclogged in plain silk — |
| | 匹庶 | | *For the common people,* |

163. *Hàn shū*, "Xíngfǎ zhì" 《漢書》刑法志.

糔飯肉食[164] barnyard millet is their food, flesh what they eat.

The quartets of characters counterbalance at higher levels, but in the most natural reading, they do not counterbalance at the most granular level:

(211′) 文杯畫案 ⟨NP⟩ → ⟨$NP_1$⟩ ⟨$NP_2$⟩ → ($V_1$)-$N_1$ ($V_2$)-$N_2$
decorated cups, painted tables —
机席緝屨 ⟨NP⟩ → ⟨$NP_1$⟩ ⟨$NP_2$⟩ → $N_1$ $N_2$ ($V_1$)-$N_3$
armrests, seat-mats, close-stitched insoles —

It is straightforward to read the first line as ($V_1$)-$N_1$ ($V_2$)-$N_2$. In the second, *jīxí* 机席 is most natural as "armrests and seat-mats," which is $N_1$ $N_2$. On the other hand, if we feel pressure to see *jī* 机 as a verb, we could read *jīxí* as (V)-N: seat-mats to which armrests have been added, "armrested seat-mats."

We can read the succeeding couplet in at least two different ways:

(211″) 衣紈履絲 ⟨VP⟩ → ⟨$VP_1$⟩ ⟨$VP_2$⟩ → $V_1$ $N_1$ $V_2$ $N_2$
garbed in fine silk, beclogged in plain silk —
糔飯肉食 ⟨VP⟩ → ⟨$VP_1$⟩ ⟨$VP_2$⟩ → ($N_1$)-$V_1$ ($N_2$)-$V_2$
they feed on barnyard millet, fleshfully they eat.

(There is our friend *ròushí* 肉食 again, "to eat in a fleshly manner.") Alternately:

(211‴) 衣紈履絲 ⟨NP⟩ → ⟨$NP_1$⟩ ⟨$NP_2$⟩ → $N_1$ $N_2$ $N_3$ $N_4$
their clothes are fine silk; plain silk their sandals
—
糔飯肉食 ⟨NP⟩ → ⟨$NP_1$⟩ ⟨$NP_2$⟩ → $N_1$ $N_2$ $N_3$ $N_4$
barnyard millet is their food, flesh what they eat.

The first reading (211″) seems preferable to me, and perhaps it is my own bias to enjoy coercing nouns into verbal senses.[165] But the second reading (211‴) is counterbalanced, down to its individual words. And that offers its own aesthetic satisfaction.

---

164. *Yántiě lùn*, "Guójí" 《鹽鐵論》國疾.

165. In fact, we can even force the line about barnyard millet into parallel with the line about fine silk: "糔飯肉食 [they use barnyard millet as their food, flesh as their eating]," yielding ⟨VP⟩ → ⟨$VP_1$⟩ ⟨$VP_2$⟩ → $V_1$ $N_1$ $V_2$ $N_2$.

Just as we are not obliged to insist on one syntactic output of a context free grammar, so the parts of speech and syntactic relationships implied by Chinese counterbalancing are fundamentally fuzzy.

That fuzziness contributes much to our impression of dense, Humboldtian "pure thought." In fact, I think that "pure thought" is not the whole story of what makes reading Wenyan such a powerful experience — the really decisive satisfaction, and the really interesting feature of Chinese, is the ability it grants us to manipulate parts of speech this way in the actual act of reading.

And that is possible because in Wenyan, whatever its origins, parts of speech are fundamentally indeterminate.

## ACKNOWLEDGMENTS

I developed most of this model in the classroom, over the two decades 1995–2014, while teaching students of Wenyan at the Universities of Washington, Minnesota, and Maryland, at National Taiwan Cheng-chi University, and at Columbia University. I am most grateful to Yuan-Yuan Meng 孟苑苑 (of Lùgǎng 鹿港 and Columbia) for conversations about Chinese syntax in general over a number of years. I also wish to thank to Paul R. Goldin, L. Amber O'Hearn, Constance Cook, Geoffrey Redmond, Chao Fang-yi 趙芳藝, Barbara Meisterernst, Roy I. Blankman, and Rose Ames for their instructive comments and criticism. Jean-Michel Fortis kindly directed me to excellent resources on the pre-Chomsky history of syntax trees. I encountered Peter Boodberg's approach in the 1980s, under Gari K. Ledyard, William G. Boltz, and Jerry Norman (1936–2012), all former Boodberg students.

## CONVENTIONS

- *Context-free grammar*. Conventions for the grammar are introduced at the beginning of Section 3 and conventions for part-of-speech notation at the end of 3.1.
- *Punctuation*.
  - 、. In Wenyan examples the only punctuation I use is the traditional *dùnhào* 頓號 [pause mark]: 、. In modern orthography, the *dùnhào* has a new and altered meaning: it separates items in a series, between which there is normally no audible pause in reading. Some native speakers are surprised to see this unfamiliar pause usage, but it is historically valid and true to the actual meaning of the name *dùnhào*. I avoid modern punctuation of Wenyan texts because it adds semantic content — a violation of the separation of concerns.

    - ％. I use a percent sign to suggest iconically the swapping that is involved in loangraph (*tōngjiǎ* 通假) usage. So 雍％壅 means that 雍 appears representing the word normally written 壅.
- *Translation issues*:
    - *Official titles*. I render ancient titles in such a way as to avoid unapt Western feudal terminology, responding to the ideas of Li Feng.[166] Where rulers are named by reign titles or posthumous titles, I generally translate them; such titles were chosen for meaning and so they are best rendered in translation.
    - *Literal versus idiomatic renderings*. When the literal rendering of some Chinese is significantly different from an idiomatic rendering, I include both, with a pointing finger (☞) between them and the literal in quotation marks. For example:

        *bǐshuō* 筆說　“to argue with pen” ☞ to make written arguments

        不日成之　“not have a day’s duration, and accomplish it” ☞ to get it done in under a day

    - *Romanization of proper nouns in Old Chinese phonology*. To transcribe proper nouns and other terms in translations, I render recent Old Chinese reconstructions. Primarily I use Baxter-Sagart, which charmingly incorporates many relationships from character-structure into the reconstructed forms. At times I also use ideas and phonological assignments of Schuessler.[167] A prime goal in using Old Chinese is to avoid the anachronistic effect of naming things from much earlier eras with contemporary Mandarin. I replace all special phonetic symbols: *ə* with *e*, *-ʔ* with *-h*, *ŋ* with *ng*, as well as making other reductions. Normally I ignore the A/B syllable-type distinction; where it is occasionally useful for disambiguation (or for preventing reconstructed forms from confusion with Pīnyīn), I replace the pharyngealization token ˁ with the apostrophe originally used by Norman (1994). Norman considered his proposal “an open system” and did not intend to insist on a narrow phonetic interpretation — that is part of the reason he called it “Early Chinese” rather than “Old Chinese.”[168] I also sometimes use an apostrophe to attach affixes to words. The goal of all this is not to render Early or Old Chinese exactly but to give the translation vividness of sound without anachronism. This is hardly a new idea. You see it

166. Li Feng 2003.

167. Baxter & Sagart 2014, 2015; Schuessler 2009.

168. Personal communication, 1991.

in full flower in Edward Schafer's tour-books of Táng literature,[169] and here is David Knechtges's 1982 rendering of line from a Hàn *fù* — he uses Pīnyin and the Old Chinese of Li Fang Kui in one breath:

> 憚夔龍兮怖蛟螭[170]
> Frighten the *kui* dragon,
> Scare the *kog* and *tya*.

- *Phonetic tokens for characters*:
  - *Pīnyīn*: To notate sound in glosses (as opposed to translations) I use Mandarin, rather than any reconstruction or metasystem representing pre-modern language, since my purpose is to provide the contemporary reader with something familiar and pronounceable. I do not capitalize proper nouns in this context, again because doing so adds semantic content to the representation of sound and that violates separation of concerns. (I do, however, capitalize proper nouns in translation.) I generally use the 1955 Mandarin standard, but prefer the 1932 standard for the tones of words of *rùshēng* 入聲 origin.
  - *Romanization of proper nouns in pre-modern phonology*. Medieval phonology, which is mentioned occasionally here, uses the Branner "neutral transcription," whose goal is to transcribe standard phonological categories without reconstructing them.[171]

## REFERENCES

BACKUS, J. W. 1959. The Syntax and Semantics of the Proposed International Algebraic Language of the Zurich ACM-GAMM Conference. *Proceedings of the International Conference on Information Processing*, UNESCO, 1959, pp. 125–132. Typewritten preprint.

BAXTER, WILLIAM H., and LAURENT SAGART. 2014. *Old Chinese: A New Reconstruction*. Oxford: Oxford University Press.

BAXTER, WILLIAM H., and LAURENT SAGART. 2015. The Baxter-Sagart Reconstruction of Old Chinese. Downloadable XLSX table of 4967 morphemes and reconstructions. Said to be Version 1.1, dated 20 September, 2014, but the filename ("Baxter-

169. Schafer 1963 and especially 1967.

170. Zhāng Héng, "Nándū fù" 張衡《南都賦》. Knechtges 1982:329.

171. Branner 2006.

SagartOC2015-10-13.xlsx") suggests 13 October, 2015. Website: http://ocbaxtersagart.lsait.lsa.umich.edu. Accessed January, 2017.

BOODBERG, PETER A. 1943. *Exercises in Chinese Parallelism*. Berkeley and Los Angeles: University of California Press. University of California Syllabus Series, No.284.

BRANNER, DAVID PRAGER. 1999. *Problems in Comparative Chinese Dialectology — the Classification of Miin and Hakka. Trends in Linguistics* series, no. 123. Berlin: Mouton de Gruyter.

BRANNER, DAVID PRAGER. 2000. 'Red Cliffs' in Taiwanese Hàmbûn. *CHINOPERL Papers*, 2002: 67–100.

BRANNER, DAVID PRAGER. 2006. Comparative Transcriptions of Rime Table Phonology. In David Prager Branner, *The Chinese Rime Tables: Linguistic Philosophy and Historical-Comparative Phonology*. Appendix II. Amsterdam: John Benjamins Publishing Company.

BRANNER, DAVID PRAGER. 2015. The Programming Language Called Classical Chinese. Presentation at Strange Loop 2015, St. Louis, MO, September 26. Slides on line at https://github.com/strangeloop/StrangeLoop2015/; video at https://youtu.be/vBhg2p8aAQ0. Accessed January, 2017.

BRANNER, DAVID PRAGER. 2018. How Simple is the Grammar of the Bronze Inscriptions? National Meeting of the American Oriental Society, 17 March, 2018, Pittsburgh.

BRANNER, DAVID PRAGER, and YUAN-YUAN MENG. 2010a. Syntactic Yoga in Chinese-English Lexicography. Proceedings of the 2009年《康熙字典》暨詞典學國際學術研討會 International Seminar on Kangxi Dictionary and Lexicology, 2009: *Zhōnghuá zìdiǎn yánjiù* 中华字典研究 2: 627–38. Delivered as a plenary speech as "Hàn-Yīng cídiǎnxué zhōng de 'yǔfǎ yújiā' 汉英词典学中的'语法瑜伽' ['Syntactic Yoga' in Chinese-English Lexicography]," Huángchéng Xiàngfǔ, Yángchéng County, Shānxī 山西省陽城縣皇城相府, on 16 July, 2009.

BRANNER, DAVID PRAGER, and YUAN-YUAN MENG. 2010b. Review of John DeFrancis, *ABC Chinese-English Comprehensive Dictionary*. *Journal of the American Oriental Society* 130.1: 83–87.

BRANNER, DAVID PRAGER, and YUAN-YUAN MENG. 2010c. Review of Lǐ Xíngjiàn 李行健, *Xiàndài Hànyǔ guīfàn cídiǎn* 现代汉语规范词典 [A standard dictionary of modern Chinese]. *Journal of the American Oriental Society* 130.2: 484–86.

BRANNER, DAVID PRAGER, and YUAN-YUAN MENG. 2013. A Curious Lexicographic Relic of the Cultural Revolution. *Journal of the Royal Asiatic Society*, Ser. 3, 23/4: 551–82.

BRITTAIN, RICHARD COULTER. 1973. *A Critical History of Systems of Sentence Diagramming in English*. Doctoral dissertation, The University of Texas at Austin.

CHAO, YUEN REN. 1968a. *Language and Symbolic Systems*. Cambridge: Cambridge University Press.

CHAO, YUEN REN. 1968b. *A Grammar of Spoken Chinese*. Berkeley, University of California Press.

CHOMSKY, NOAM. 1956. Three Models for the Description of Language. *IRE Transactions on Information Theory* 2.3:113–24.

CHOMSKY, NOAM. 1959. On Certain Formal Properties of Grammars. *Information and Control* 2:137–67.

CHURCH, ALONZO. 1940. A Formulation of the Simple Theory of Types. *The Journal of Symbolic Logic* 5.2:56–68.

CHURCH, ALONZO, and W. V. QUINE. 1952. Some Theorems on Definability and Decidability. *The Journal of Symbolic Logic* 17.3:179–87.

CIKOSKI, JOHN A. 1978. Three essays on Classical Chinese grammar (3): an analysis of some idioms commonly called 'passive' in Classical Chinese. *Computational Analysis of Asian and African Languages* 9.133–208.

*DÀ CÍDIǍN* 大辭典. 1985. Taipei: Sānmín Shūjú 三民書局.

DIJKSTRA, EDSGER W. 1969. EWD 273: The Programming Task Considered as an Intellectual Challenge. MS on line at https://www.cs.utexas.edu/users/EWD/ewd02xx/EWD273.PDF. P. 13.

DIJKSTRA, EDSGER W. 1974. EWD 447: On the role of scientific thought. In *Selected Writings on Computing: A Personal Perspective*, (Springer-Verlag, 1982). MS on line at https://www.cs.utexas.edu/users/EWD/ewd04xx/EWD447.pdf. Accessed January, 2017.

DOUGLAS, CARSTAIRS. *Chinese-English Dictionary of the Vernacular or Spoken Language of Amoy*. London: Presbyterian Church of England, 1873. Revised edition of 1899. Reprinted at Princeton: Chinese Linguistics Project, 1970, and numerous times since then in Taiwan.

FOX, ANTHONY. 1995. *Linguistic Reconstruction: An introduction to theory and method*. Oxford: Oxford University Press.

*GUǍNGYÙN*. 1008. *Xīnjiào zhèngqiè Sòng běn Guǎngyùn* 新校正切宋本廣韻. Ed. Lín Yǐn 林尹. Taipei: Límíng Wénhuà Shìyè 黎明文化事業, 1976.

*HÀNYǓ DÀ ZÌDIǍN* 漢語大字典. 1986–90. Chéngdū 成都: Húběi, Sìchuān císhū chūbǎnshè 湖北，四川辭書出版社.

HARBSMEIER, CHRISTOPH. 1998. *Language and Logic*. Joseph Needham, *Science and Civilisation in China*. Vol. 7, Part 1. Cambridge: Cambridge University Press.

HOCK, HANS HEINRICH. 1991. *Principles of Historical Linguistics*. Berlin: Mouton de Gruyter.

VON HUMBOLDT, WILHELM. 1836. *Über die Verschiedenheit des menschlichen Sprachbaues und ihren Einfluß auf die geistige Entwickelung des Menschengeschlechts*. Berlin, Gedruckt in der Druckerei der Koniglichen Akademie der Wissenschaften. Edition: *Wilhelm von Humboldts gesammelte Schriften*, ed. Albert Leitzmann. Berlin: B. Behr's Verlag, 1907; repr. Berlin: Walter de Gruyter & Co., 1968.

KARLGREN, BERNHARD. 1933. Word Families in Chinese. *Bulletin of the Museum of Far Eastern Antiquities*. 5:9–120.

KENNEDY, GEORGE A. 1938. *A Beginner's English-Chinese Vocabulary*. [New Haven] Department of Oriental Studies, Yale University.

KENNEDY, GEORGE A. 1939. *Chinese Reading for Beginners*. [New Haven] Department of Oriental Studies, Yale University.

KENNEDY, GEORGE A. 1964. *Word Classes in Classical Chinese*. In Tien-yi Li, ed., *Selected Works of George A. Kennedy*. New Haven: Far Eastern Publications, Yale University. Reprinted mainly from *Wennti* 9 (April, 1956), and partly from Kennedy's unfinished manuscript.

KERNIGHAN, BRIAN W., and ROB PIKE. 1984. *The UNIX Programming Environment*. Upper Saddle River, NJ: Prentice Hall. See esp. "Input-output redirection" and "Pipes," pp. 29–33.

KNECHTGES, DAVID R. 1982. *Wen xuan or Selections of Refined Literature*. Author listed as Xiao Tong; translated, with annotations and introduction by David R. Knechtges. Vol. 1: Rhapsodies on Metropolises and Capitals. Princeton: Princeton University Press.

KNUTH, DONALD E. 1964. Backus Normal Form vs. Backus Naur Form. Letters to the Editor. *Communications of the ACM*, 7.12:735–6.

KNUTH, DONALD E. 1968. Sec. 2.3, Trees; Sec. 2.3.4.6, History and Bibliography. In *The Art of Computer Programming*. Vol. 1, *Fundamental Algorithms*. Reading, Massachusetts: Addison Wesley, pp. 308–15; 405–6.

KNUTH, DONALD E. 1997. Sec. 2.3, Trees; Sec. 2.3.4.6, History and Bibliography. In *The Art of Computer Programming*. Vol. 1, *Fundamental Algorithms*. Third ed. Upper Saddle River, NJ: Addison Wesley, pp. 308–18; 406–7.

KROLL, PAUL W. 2015. *A Student's Dictionary of Classical and Medieval Chinese*. Leiden, Boston: Brill.

LI, FENG. 2003. "Feudalism" and Western Zhou China: A *Criticism. Harvard Journal of Asiatic Studies*, 63.1:115–144

MǍ, JIÀNZHŌNG 馬建忠. 1954. *Mǎshì wéntōng jiàozhù* 馬氏文通校注. Ed. Zhāng Xíchēn 章錫琛. Beijing: Zhōnghuá Shūjú.

MAZZIOTTA, NICOLAS. 2016. Drawing Syntax before Syntactic Trees. Stephen Watkins Clark's Sentence Diagrams (1847). *Historiographia Linguistica*, 43.3:301–342.

MAZZIOTTA, NICOLAS, AND SYLVAIN KAHANE. 2017. To What Extent is Immediate Constituency Analysis Dependency-Based? A Survey of Foundational Texts. *Proceedings of the Fourth International Conference on Dependency Linguistics*, Pisa, Italy, September 18-20, 2017, pp. 116-126.

MEISTERERNST, BARBARA. 2016. The Expression of Epistemic Modality in Classical and Han Period Chinese. In Barbara Meisterernst, ed., *New Aspects of Classical Chinese Grammar*. Wiesbaden: Harrassowitz Verlag, 105–29.

NAUR, PETER (ed.). 1960. Report on the Algorithmic Language ALGOL 60. *Communications of the ACM*, 3.5:299–314.

NORMAN, JERRY. 1994. Pharyngealization in Early Chinese. *Journal of the American Oriental Society*, 114.3:397–408.

PERCIVAL, W. KEITH. 1976(1967). On the Historical Source of Immediate-Constituent Analysis. *Notes from the Linguistic Underground*. Ed. James D. McCawley. New York: Academic Press, 1967. *Syntax and Semantics*, 7:229-242. Presented originally at the Linguistic Circle of Madison at the University of Wisconsin, 1967. Electronic version on line at https://people.ku.edu/~percival/ICanalysis.html. Accessed November, 2018.

PERLIS, A. J., and K. SAMELSON. 1960. Appendix Two: Preliminary Report of ACM-GAMM Committee on an International Algebraic Language (1960). Reports on a meeting of 1958. In *Annual Review in Automatic Programming*, Vol. 1. Series: International Tracts in Computer Science and Technology and Their Application, Vol. 3. Oxford: Pergamon.

SCHAFER, EDWARD H. 1963. *The Golden Peaches of Samarkand; A Study of T'ang Exotics*. Berkeley, University of California Press.

SCHAFER, EDWARD H. 1967. *The Vermilion Bird; T'ang Images of the South*. Berkeley, University of California Press.

SCHUESSLER, AXEL. 2007. *ABC Etymological Dictionary of Old Chinese*. Honolulu: University of Hawai'i Press.

SCHUESSLER, AXEL. 2009. *Minimal Old Chinese and Later Han Chinese: A companion to Grammata serica recensa*. Honolulu: University of Hawai'i Press.

SPRINKLE, RUSSELL, *et al.* 1976. *Amoy-English dictionary*. Taichung: Maryknoll Fathers.

WÁNG, LÌ 王力. 1982. *Tóngyuán zìdiǎn* 同源字典. Beijing: Shāngwù Yìnshūguǎn.

WÁNG, LÌ 王力 ET AL. 2000. 王力古漢語字典. Beijing: Zhōnghuá Shūjú.

*XIÀNDÀI HÀNYǓ CÍDIǍN* 現代漢語詞典 1978. Beijing: Shāngwù Yìnshūguǎn.

*XĪNHUÁ ZÌDIǍN* 新華字典 1971. Beijing: Shāngwù Yìnshūguǎn.

ZÁDRAPA, LUKAŠ. 2016. *Lǔshì chūnqiū*: Object Words Denoting Processes. In Barbara Meisterernst, ed., *New Aspects of Classical Chinese Grammar*. Wiesbaden: Harrassowitz Verlag, 149–74.

Proceedings of the 29th North American Conference on Chinese Linguistics (NACCL-29). 2017. Volume 1. Edited by Lan Zhang. University of Memphis, Memphis, TN. Pages 85-98.

# THE CHINA URBAN LANGUAGE SURVEY PROJECT 2003 - 2016

范德博
Marinus VAN DEN BERG
*Leiden University, IIAS, Emeritus*

In this presentation, three major hypotheses developed in the context of the China Urban Language Survey Project will be discussed. Changes in the urban language environment are mainly related to large scale migration from the country side and other places toward the newly developed and developing industrial centers in China's southeastern provinces. The project started in 2003 in Nanjing University's Sociolinguistic Laboratory under the guidance of professor Xu Daming, attracted researchers from various universities both in China, Hong Kong, Taiwan and Europe, and obtained funding from various sources including a major grant from the Netherlands' Organization for Scientific Research (NWO). Progress reports were presented in yearly conferences, and major findings appeared in a 2010 book publication (van den Berg & Xu 2010). We will start this presentation with the notion "long-term accommodation", based on work by Van den Berg in Taiwan in 1977-78 (van den Berg 1986), extend that notion to developments in mainland China, and present three major hypotheses developed in the context of the China Language Survey Project, a triglossia/diglossia hypothesis based on work in Hainan province (Tsou et al. 2010), the fundamentals of Speech Community Theory developed in studies of the language situation in the Inner-Mongolian city of Baotou (Xu 2004; 2010), and the concept of network density, developed in a study of the language situation in Beijing (Song & Zhu 2016).

## 1. Long-term accommodation

In a questionnaire survey among National Taiwan University students in 1977, it was possible to demonstrate across generation adjustment to the national language configuration at each of the time frames involved. For grandparents of the students, who were born during the Qing Dynasty or during the beginning years of the Japanese colonial period, Minnan was the dominant language, and more so for grandmothers (71%; 74%) than for grandfathers (52%; 48%). The first of these figures giving maternal grandparents language backgrounds, and the second those of students' paternal grandparents. These data confirm the limited options for female education during the last of the empirical dynasties, whereas Minnan-Japanese bilingualism data show the first signs of that gender gap's closure. For maternal grandparents, the bilingualism data were 22% and 27% for

grandmothers and grandfathers respectively, whereas Minnan-Japanese data for paternal grandparents showed a wider gap, 21% for the grandmothers, comparable to the 22% found for the maternal grandmothers, and 35% for the grandfathers, extensively higher than the figure found for the maternal grandfathers (27%).

The interesting observation for the generation of parents born during the Japanese period is that a majority of them became trilingual in Minnan, Japanese, and Mandarin Chinese, called Guoyu 國語 under the Republican government. A gender gap still existed, but both sexes had in majority become trilingual. The data for the students' mothers was 46% and for the fathers 57%. Parents who were older or were otherwise disadvantaged in learning a new language variety did not pick-up Guoyu 國語, and remained bilingual Minnan-Japanese speakers, and this occurred more typically for mothers (18%) than for fathers (12%). Minnan mono-lingualism showed a similar gender difference, 20% for mothers and 9% for fathers. These latter differences most likely have a correlation with economic activity (farm labor) and income.

The language situation of the students themselves gives a totally different picture. Japanese has disappeared from the language repertoire, and the best language claimed by almost all students (93%) is Mandarin Chinese (Guoyu). With an average age of 20, these students were born around 1957, and had participated in a Mandarin Chinese dominant education system, moving from elementary school, to high school, and on to university, in this case, one of the top universities in Taiwan, implying that the results are not those of the average student, but of a selection of Taiwan's top students, who also studied in Taiwan's political, economic, and educational center, Taipei.

Keeping this in mind, these data allow the interpretation that members of a national community adjust to the national level language configuration of that community. This adjustment, using insights from interpersonal accommodation theory (Giles & Powesland 1986), we called *long-term accommodation*, suggesting that over-time community members, while keeping accommodating to everyday language requirements, are forced to adjust to the norms set at the national level. The result is not unexpected. It was observed for the development of Latin in the Roman Empire and related to factors such as a central government supporting economic development, making it worthwhile to acquire the language of government and education, the presence of social mobility, which helps to create multilingual areas, thereby creating the need for a lingua franca.

Using this insight and turning now to mainland China, it is possible to predict that over- time the national standard language, Putonghua, will spread. The conditions for that spread, using the Taiwan and Roman data, are a well-organized education system, a language market supporting the use of Putonghua, economic development, and social mobility which will make Putonghua the aspired to lingua franca. As we will see, education got reorganized after 1980, the language market was strongly influenced by mass-migration, and the spread of Putonghua awaited the emergence of an economically developing China.

## 2. Developments in Modern China

Without paying attention to the pre-1980 language situation, we must conclude that the long-term accommodation hypothesis is not confirmed by developments in the 1980s. The language market in the first ten years after the start of the 1980 Open Door policy was in favor of Cantonese, the language variety of the capitalist entrepreneurs with a Hong Kong background, who invested in factories in Shenzhen and the Pearl River Delta (Zhan 1993; Guo 2004). The first wave of workers came from the surrounding country-side and were all speakers of some Cantonese dialect, and as a result oriented themselves on Cantonese rather than on Putonghua for communication. Shenzhen, the first Special Economic Zone, and the surrounding areas soon also attracted many people from other parts of China, including the Mandarin dialect area. Migrants in a new city like Shenzhen mixed home dialect languages (Hakka, Siyi (Taishanese), Swatow (Shantou), etc.) with Putonghua and Cantonese, making Putonghua the dominant language for the technical professions and for business transactions in new districts (Van den Berg 2009; Tang 2016), whereas in Cantonese speaking Guangzhou, northern and better educated migrants formed their own Putonghua based networks (Van den Berg 2010). One question that in this setting comes to the fore is how will these developments work-out for China as a whole? Various answers are possible, Putonghua will destroy the regional dialects, the dialects will stop the spread of Putonghua, or a new bilingual balance will develop. We will not discuss each of these possibilities further at this moment, but in the following we will discuss three proposals, the triglossia/diglossia hypothesis, Speech Community Theory, and network density. Each gives a somewhat different view on what is most likely the future development of the Chinese urban language market. The first of these is the strongest proposal and is the one that addresses the national language situation. It predicts a general tendency, so let's see what that line of research has to say.

## 3. From triglossia to diglossia

In an elementary school survey in Sanya, Hainan province, researchers observed that students, in addition to some use of the regional language variety Hainanese, mainly used the home dialect (six different language varieties), when talking to their grandparents, whereas in communication with parents, some Mandarin Chinese was introduced. When talking to each other the level of Mandarin Chinese got more extensive, reaching even higher levels when an everyday task such as shopping was involved. In public transportation, finally, Mandarin Chinese obtained its maximum use. In the latter case, we can imagine the lingua franca effect of the use of Mandarin Chinese, particularly when public transportation personnel have different backgrounds. Hainanese is still being used as lingua franca for around 20 percent of the cases, but in the remaining settings Putonghua is dominant, taking over in effect the lingua franca role of Hainanese (Tsou et al. 2010).

The researchers concluded that given rapid modernization taking place in Sanya, there is a language shift taking place from the home language to, what they call, the

Supreme Language, Putonghua. They see this as a shift from a triglossia situation (home dialect-Hainanese-Putonghua) toward a diglossia setting (home language-Putonghua). And they even went one step further claiming that this development is taking place throughout China, the motivation being that a triglossic setting demands a higher psychological burden than a diglossic setting, and as a result it would be only natural to see a shift in the direction of diglossia (home dialect-Supreme language), whereas in the long run the home dialects too will have to disappear, mono-lingualism assumedly providing even less of a cognitive burden.

Given this analysis, there are quite a few questions that can be asked, and we will do so in a moment, when we will address the language situation in Shanghai. One element, however, that from the perspective of urban language studies is clearly missing in the discussion is the impact of social stratification, and in order to get that more clearly into perspective it is necessary to focus on Speech Community Theory, which states that in an economically developing urban environment social stratification and language differences must occur.

**4. Speech Community Theory**

Bloomfield, as early as 1933, devoted a full chapter to the discussion of the speech community (Bloomfield 1933: 42-56). He defined a speech community as: "a group of people who interact by means of speech" (Bloomfield 1933: 42). Simple as this definition might seem, Bloomfield was clearly aware of many of its implications, for instance, he did not specify what kind of speech, and in this definition, this can be any kind of speech, the fundamental and essential ability of humans to interact with each other. He further specified the speech community "as the most important kind of social group", different from "other phases of social cohesion, such as economic, political or cultural groups". More importantly, quite early in the chapter, Bloomfield mentioned "the assimilation into a speech community of whole groups of foreigners, such as immigrants, conquered people, or captives," an issue that is at the heart of the present presentation. Unfortunately, he has not much to say about the process itself, but he is clearly aware of the implications. Limiting ourselves to complex, large sized, speech communities, a distinction Bloomfield makes too, he observes that everywhere differences in speech develop, the main sources are distance, which is *geographical*, and *social* differentiation. As to the latter, he points to the speech differences which develop within the standard language under the influence of differences in family tradition, schooling, occupation, and income, which results in subgroups we recognize as *social classes*. Of course, in this general introduction, Bloomfield did not discuss details of the formation process of a speech community. That issue was taken up more recently in a series of real-time studies I like to introduce now, and which we know under the title of Speech Community Theory (Xu 2004; 2010; 2016).

The observations by Bloomfield got a new impetus when the New York speech community was analyzed in an empirical way (Labov 1966/2006). That approach

confirmed in essence most of the distinctions made by Bloomfield, but now certain of these distinctions were given real content such as a precise distribution of the phonemes [-r] and [aw] across social space. Often too new labels, such as style shifting, were introduced. This empirical approach became the new standard, and the empirical study of a speech community was fundamentally different from earlier more theoretically oriented approaches (Patrick 2002). It is this line of research that was taken up in the study of a developing Chinese speech community, the new Kundulun district in the city of Baotou in Inner Mongolia (Xu 2010).

Speech community theory starts from the observation that the social system of speakers (think of the social groups mentioned by Bloomfield) is the basis for understanding the way of speaking in a community. Where linguists tend to focus on speech sounds and grammatical rules, they lose contact with the social organization of the speaker group. In order to clarify this, Xu Daming and his students over a period of twenty years studied the changing language situation in the already mentioned new industrial district of Kundulun, part of the city of Baotou in Inner- Mongolia. This district was established in 1956 as the residential area for workers of the large state-owned Baotou Steel Corporation, housing around three-hundred thousand people. Migrants came from all parts of China, but the majority originated from surrounding provinces, and spoke varieties of Mandarin dialects. Dialect contact resulted in variation in forms of nasalization. A comparison of the results of two studies, one in 1987 and a follow-up study in 2006, showed that the relationship between internal phonetic variation and selected social variables gets more complex over time, and this led to the formulation of a theory of speech community formation. The theory stresses the development of shared forms of communication as the result of day-to-day communication between same-group community members, compare Bloomfield's *mutual adjustment*. The relation between 'occupation' and other variables ('social network', 'place of origin' as the more persistent ones) with phonetic variation made clear that, since occupation, education, and income are correlated, this kind of variation creates a socially stratified society. And in such a society, speech styles are stratified too. The importance of these studies is that they trace, in great detail, the development of a new speech community in terms of intrinsic constraints in relation to a set of social variables. The implication for the triglossia>diglossia hypothesis is that is needs to be evaluated in the context of a socially stratified society, and that is what we will do in the next section, where we will study a large, complex, socially stratified, speech community and address the issue of migration as raised by Bloomfield and see to what extent this leads to assimilation, as he noted, or will take different forms. The analysis will also allow us to test to what extent the triglossia hypothesis put forward in the previous section can be maintained.

## 5. The Shanghai language situation

Neglecting the formative period of the Shanghai dialect in the 19th and 20th centuries (see for this issue Van den Berg 2016), the developing Shanghai language

situation after the civil war, can be divided into two periods, a relatively stable period after 1949 and a rapidly changing period with massive migration after 1990. Let us start by looking at the demographic data in order to see when and to what extent migration occurred. The population development data since 1952, shown in figure 1, shows a curve that at first during the beginning years of communist control mildly increased, then, during the cultural revolution, fell back under the influence of the *xia xiang*, 'back to the country side' policy. After 1990, the curve rises rapidly, increasing from 12 million at that time to 23 million in 2010, an increase of almost hundred percent when compared with 1980, the year the 'Open Door' policy started. In the first period after 1949, Shanghai was economically strangled by the new power holders, as a punishment for its days of glory and Western involvement in the previous hundred or so years. That first period was the time of social control and the 'iron rice bowl,' and social stratification the result of family tradition, occupation, schooling and income did not shape Shanghai society, unless we accept a class difference between communist cadres and the population at large. In any way, that would be a completely different form of stratification from what we saw developing after 1990, when Shanghai was given the opportunity to develop again.

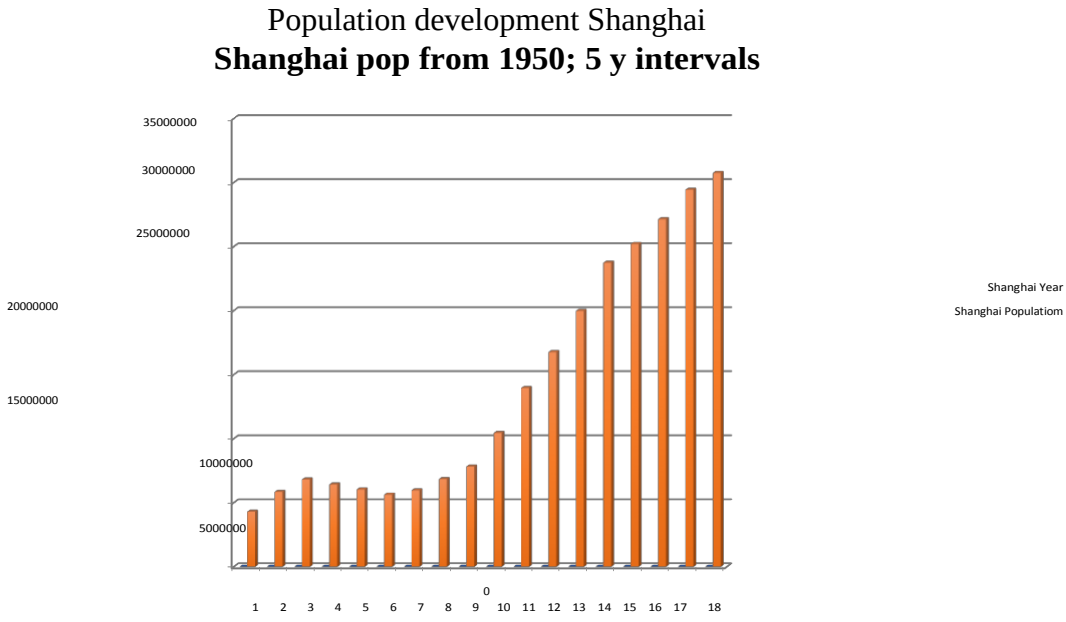


Figure 1. Shanghai's population development after 1950

After 1990, Shanghai started to develop again after it finally was allowed to do so, and sprinted forward, becoming China's economic center again in just a few years. In this

second period, international cooperation intensified, the stock market reopened, new occupations were added, and higher incomes obtained. This happened in particular after 2000 when university education had modernized, and the younger generation could compete in many fields. The result was a stratified society, with the under forty generation developing most rapidly. Language use data in four economically stratified department stores in the Xujiahui area support that view (see Van den Berg 2016 for the details).

In the second period, one of the driving forces in Shanghai's rise to prominence was migration. The 2010 census showed that of Shanghai's 23 million inhabitants, 9 million were migrants with long-term resident status, and those with origins from the Mandarin dialect areas Anhui (29%), Jiangsu (17%), Henan (9%), and Sichuan (7%) were the largest groups. The 2010 census also showed that four-fifth of the migrants had a rural background, suggesting that at that time manual labor was still needed most, but that higher educated groups were increasing in number. Data of language use in public places shows that the manual labor group communicated in their home dialects or used an accented form of Mandarin Chinese. It is the second group that in its various daily contacts, both inside the job, and during shopping or leisure, uses Putonghua, the language of higher education (van den Berg 2016). This form of communication is possible only when, at the Shanghainese side, bilingualism has developed as well. Assuming that is the case, we can map the various social groups and their level of bilingualism as four social classes (Figure 2).

Language Stratification in Shanghai
Statistics copied from New York 1963 (Labov, 1963; table 7.5)

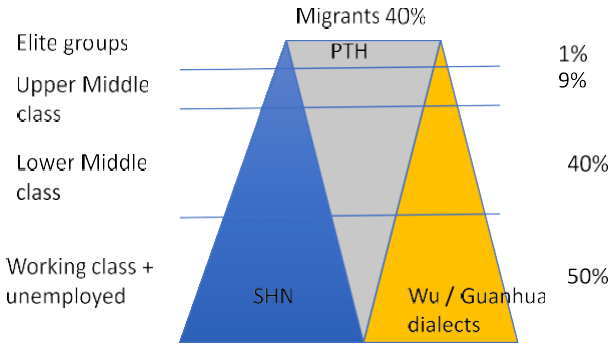


Figure 2. Social class and language stratification in Shanghai

At the top we have the one percent elite group, representing very rich entrepreneurs, in majority speaking Putonghua, but also people speaking their home dialect, Shanghainese, a Wu dialect, or a Mandarin (Guanhua) dialect. The majority of the population (50%), however, belongs to the lower and upper working classes, and it is they who speak in majority Shanghainese, or one of their many home dialects, be it from the Wu group or the Mandarin group. Among them, and most likely more extensively among the upper working class, various people have a high school background and are accustomed to speaking Putonghua. The next largest group is the lower middle class (40%), education among them is higher and so is their income, their level of bilingualism and the use of Putonghua. In this group, as the figure suggests, better educated Shanghainese speakers, Wu dialect and Mandarin dialect speakers, tend to match the group of Putonghua speakers. Our fourth and last group (9%) represents members of the upper-middle class, these are all higher educated individuals, the majority of which is accustomed to speaking Putonghua. Upper-middle class membership, however, is not restricted to Putonghua speakers, we also find people with a Shanghainese background, who prefer to speak Shanghainese. Most Wu dialect speakers will switch to Shanghainese when interacting with locals, switching to Putonghua when appropriate. In the same way, Mandarin dialect speakers, in an interaction with a local person, most likely will use Putonghua, helping to increase that percentage that way. When talking to a *guxiang*, 'home dialect person' the shared dialect will be used (Fig. 2).

Having established the status of social diversification in modern Shanghai society, we now need to determine the extent to which the triglossia>diglossia hypothesis holds in this context? Shanghainese and some of the Wu and Mandarin dialects are well being

maintained among members of the working class, either by themselves or with support from the home county or township. The chances that Shanghainese, as the regional language, and the one comparable to Hainanese, will be given up, however, is small. Shanghainese is a high-status language variety and will be maintained for the foreseeable future by millions of people. Wu dialects spoken in Shanghai are under pressure from this more prestigious Shanghainese, whereas Mandarin dialects will feel the pressure from the Supreme variety, Putonghua. Both dialect groups, Wu dialects and Mandarin dialects, however, have strong backing from their home dialect areas and those areas will not be directly influenced by what happens in Shanghai. These various home dialects might be used less in Shanghai but will remain supported at the home county area. They remain under influence of national level education policy though. The triglossia>diglossia hypothesis therefore is not supported by the Shanghai data, and a reformulation will be necessary.

What is needed, in addition to the urban phenomenon of social stratification, -- the Hainan survey does no address that issue adequately either -- is a distinction which describes differences in *vitality* of the language varieties in use in Shanghai. A language variety's vitality is determined by three factors, *status factors*, *demographic factors*, and *institutional support factors* (Bourhis et al., 1981). Shanghai's status is related to Shanghai's position as China's main economic center. In Hainan that would be a comparison with Haikou, the Hainan provincial capital, but that information was not provided in the Tsou et al. study. Demographic data further set Shanghainese apart from Hainanese, which attracted far smaller numbers of speakers. The third factor, institutional support, favors Putonghua as the Supreme variety, with support in both education and the media. Shanghainese, however, still has a certain amount of support in those domains, whereas that support might be there for Hainanese, but was not documented in the quoted paper and we assume that support is limited.

Given this observation, what we see in Shanghai is the demise of home language varieties, the ones that are passed on from generation to generation, since it is the modern younger generation who are modernizing rapidly and come to see those varieties as superfluous. In Shanghai it is not the regional language that is pushed out under influence of Putonghua, Shanghainese has high status and is maintained, what did change was the level of bilingualism among Shanghainese speakers, which strongly increased. What does tend to disappear, in contrast, is the original home dialects in Shanghai (Chu 2001). These dialects have little vitality, in terms of status, number of speakers, and institutional support, and are under pressure from both the regional language Shanghainese and the Supreme language Putonghua. It is in particular the younger generation that in this rapidly changing modern society find it difficult to support these home language varieties, despite their emotional link to the home county, especially when ancestor worship still is located there.

Having set the stage for bilingualism in a multi-million and multi-lingual city as Shanghai, it is now time to look at bilingualism in another multi-million city, Beijing.

How will networks function in that city, and what can we learn from it for the triglossia>diglossia hypothesis and for social stratification research?

**6. Network density**

Bloomfield observed that in a complex speech community we can observe differences in density of communication. Some speakers communicate more to certain contacts than to others. Theoretically, these differences could be mapped, even patterns across-time could be discovered, but, as Bloomfield observed, in reality this is impossible to do (1933: 42-56). Taking up this challenge, Song and Zhu (2016) designed a technique for comparing the density of communication among young (20-44 years) native Beijing dialect speakers, with the intent to being able to evaluate the strength of maintenance of the Beijing dialect versus the incoming force of Modern Standard Chinese, Putonghua. Using the network concept, they called a person's density of communication with other contacts, his *network density*. Their procedure was to select a sample of Beijing dialect speakers (n=269) from each of the ten central city districts, five inner-city districts and five inner-suburb districts, and asked respondents to select five persons with whom they communicated most frequently on a daily basis. They also asked the respondents to order these contacts according to the amount of time they were communicating. This procedure resulted in a classification that compared the number of Beijing dialect speaking contacts in a respondent's network of five, with the number of Putonghua speaking contacts in that same network. Using the abbreviation of BAF for Beijing Accent Friend, and PAF for Putonghua Accent Friend, this classification matched a BAF of 5 (all five contacts speaking Beijing dialect) with a PAF score of 0 (zero), and a BAF score of 4 with a PAF score of 1, etc.

Respondents data were collected either in face-to-face encounters or through telephone interviews. The procedure chosen was time-consuming but rewarding. A standard set of questionnaire questions was used to encourage the respondents to comment on their age and background, how long they were living in that particular district, detail their attitude toward Beijing dialect and Putonghua, report the language variety they were most commonly speaking, and give examples of their language use and forms of code-switching. Direct observation further allowed the researchers to evaluate the level of maintenance of Beijing dialect. The demographic details showed that the sample contained two groups of people, an Old Beijing group whose parents already lived in the area, and a New Beijing group, whose parents arrived in the city only more recently.

Respondents attitude toward Beijing dialect and Putonghua was in agreement with earlier studies, comparing dialect and standard language (Fishman 1972). Beijing dialect was generally described as kind, pleasant to hear, giving a sense of belonging, creating a feeling of interest and of history, giving a sense of identity, a language one is accustomed to, and one that is efficient. The listing is in the order of frequency of reporting. In contrast the respondents feel that Putonghua is convenient for between group

communication, is considered normal and formal, whereas Beijing dialect sounds more like country-side speech, whereas Putonghua is easy to handle and easy to communicate in, and gives a sense of superiority. Clearly these answers more strongly reflect, respectively, Old Beijing backgrounds and New Beijing backgrounds, as we will see.

Another interesting result was that respondents were not particularly clear as to the kind of language variety they were speaking. Having a set of criteria to evaluate a respondent's speech, the researchers were able to categorize the answers given. One of the striking results of this approach was the conclusion that a large number of respondents claiming to speak Putonghua, actually were speaking Beijing dialect (n=44). After correcting for this discrepancy, the analysis showed that attitude, maintenance, selection and code-switching were all under the influence of the top three Beijing dialect speakers. A positive attitude toward Beijing dialect was also found when the spouse was a local, Beijing dialect speaking, person. That positive attitude was, further, more strongly supported in the five inner-city districts (as compared to the inner-suburb districts).

The results for the Putonghua speakers was similar. Attitude toward Putonghua, selection of that variety instead of Beijing dialect, and code-switching were all three strongly supported when the first three listed speakers were Putonghua speakers. The second strong support came from having a New Beijing background. That background resulted in having a positive attitude toward Putonghua and having a preference for using it. A local spouse, further, not only helped to create a positive attitude but was also reported as instrumental in using Putonghua. In the same way, living in an inner-suburb district resulted not so much in a positive attitude toward Putonghua, but rather was supportive for selecting Putonghua as the preferred language variety.

Is the triglossia hypothesis supported by these data? It seems difficult to make that claim on the basis of these data. Rather, these data add various dimensions to the triglossia/diglossia debate and allow the construction of a dialect maintenance model, that combines geographical, personal and network variables: a person's background (Old Beijing versus New Beijing), the language preferences of the three persons he most frequently talks to, his place of residence (Inner district; Outer district), and the language preference and background of the spouse. This way a model is created that helps to understand the changing fortunes of dialect and standard language in a person's personal life and can be applied to all dialect-standard language settings. Clearly, the potential difference between the established and newcomers is demonstrated in this study as well (Elias & Scotson 1965/1994). The model does not predict language behavior in public settings, such as public transport and shopping environments. Giving the overwhelming number of migrants in Beijing, the situation, most likely, will be similar to that reported for Sanya, increased use of Putonghua in public places. There of course one has no choice as to the selection of the kind of person one is interacting with, and the presence of a shared lingua franca will facilitate the choice of that language variety.

## 7. Conclusion

We started the presentation with introduction of the concept long-term accommodation, based on Taiwanese data from the mid 1970s. When applying these to the Chinese mainland, we expected to find large scale adjustment toward the use of Putonghua, but for the first ten years after the 'Open Door' policy that could not be confirmed. Rather during those first ten years, Cantonese, the language of jobs and opportunity for a higher income, spread. A study of language acquisition of elementary school children in Hainan in more recent time, however, demonstrated that indeed the national standard language, Putonghua, the language of education, was spreading throughout the various groups making-up the Sanya community. The hypothesis based on that research stipulated that there would be nationwide a tendency to remove regional languages in favor of the national language Putonghua. A first test for the triglossia>diglossia hypothesis was provided by a real-time study of speech community formation. That study showed how, over-time, mutual adjustment between speakers takes place, and how, given the nature of a person's occupation, place-of-origin, and social network, a speech community diversifies and gets socially stratified. This study for the first time demonstrated in detail how this mechanism works and how in the future we can understand language behavior in large-scale urban settings. Using these findings, we need to observe that the triglossia>diglossia hypothesis does not take into consideration the development of social stratification in Sanya, and that might very well force the authors to adjust the hypothesis in these terms. We expect dialect maintenance to be stronger among the lower educated working classes, whereas dialct loss will be stronger among elite groups.Finally, when testing this hypothesis in Shanghai and in Beijing, that hypothesis could not be supported. Additional forces needed to be added. Social stratification was already mentioned. The second addition to the discussion is the application of the full force of language vitality theory, which distinguishes between language status, demographic strength, and institutional support. The latter addition made clear that Shanghainese, as regional lingua franca, cannot be compared to a more local language variety such as Hainanese, which on each of the three dimensions mentioned has lesser status. The Beijing network density study, we introduced, allowed, in addition, the formulation of three more variables that need to be taken into consideration when a language variety is evaluated. One needs to consider the location and role of the city center as center of dialect maintenance, a speaker's personal background (time and location of residence), network density (number of dialect speakers in the network) and the tradition of the family, which includes the language background of parents and spouse.

## REFERENCES

BLOOMFIELD, L. 1933. *Language.* New York: Holt.

BOURHIS, RICHARD Y., GILES, HOWARD AND ROSENTHAL, DOREEN. 1981. Notes on construction of a 'subjective vitality questionnaire' for ethnolinguistic groups. *Journal of Multilingual and Multicultural Development* 2, 145-155.

BROSNAHAN, L.F. 1963. Some historical cases of language imposition. In, *Language in Africa*, John F. Spencer (ed.), pp. 7-24. Cambridge: Cambridge University Press.

CHU, XIAOQUAN. 2001. Linguistic diversity in Shanghai. *Journal of Asian Pacific Communication,*11 (1): 17-24.

ELIAS, N. AND SCOTSON, J.L. 1965/1994. *The established and the outsiders.* London: Frank Cass & Co./ Norbert Elias Stichting

FISHMAN, J. 1972. *Sociolinguistics*. Rowley, MA: Newbury House.

GUO, XI. 2004. *Zhongguo Shehui Yuyanxue (Chinese Sociolinguistics).* Hangzhou: Zhejiang University Press.

LABOV, W. 1966/2006. *The Social Stratification of English in the New York City.* New York: Cambridge University Press.

PATRICK, PETER L. 2002. The Speech Community." In J.K. Chambers, Peter Trudgill & Natalie Schilling-Estes, eds. *The Handbook of Language Variation & Change*. 573-597. Oxford: Blackwell.

SONG, JING AND ZHU, YINGMEI. 2016. Attitude, maintenance, and use of Beijing dialect and Putonghua among younger generation Beijing born residents. *Journal of Asian Pacific Communication,* 26/1: 32-55.

TANG, ZHIXIANG. 2016. Language choice and language attitude among different social segments in Shenzhen. *Journal of Asian Pacific Communication* 26/1: 143-160.

VAN DEN BERG, M.E. 1986. *Language Planning and language Use in Taiwan.* Taipei: Crane.

VAN DEN BERG, M.E. 2005. Vitality, Identity, and Language Spread: The Case of Shanghainese. *Journal of Chinese sociolinguistics* 5/2, 225-235.

VAN DEN BERG, M.E. 2009. Jingji Xiandaihua he Yimin: Yueyu, Putonghua Jingzheng Guanxi de Fenxi. (Economic Modernization and Migration: An Analysis of Cantonese-Putonghua Competition.) *Nanjing University Applied Linguistics Seminar*, May 18.

VAN DEN BERG, M.E. 2010. Socio-Economic Stratification in the Guangzhou Speech Community: Language Behaviour in Shopping Areas of Yuexiu and Tianhe Districts. In, M.E. van den Berg and Daming Xu (Eds.), (pp. 236-268).

VAN DEN BERG, M.E. 2016. Modernization and the restructuring of the Shanghai speech community. *Journal of Asian Pacific Communication,* 26/1: 112-142.

VAN DEN BERG, M.E. & XU, DAMING (EDS). 2010. *Industrialization and the Restructuring of Speech Communities in China and Europe*. Newcastle upon Tyne: Cambridge Scholars Publishing.

XU, DAMING. 2004.Yanyu Shequ Lilun (Speech Community Theory). *Journal of Chinese Sociolinguistics* 4.

XU, DAMING. 2006. Yuyan de bianyixing yu yanyushequ de yizhixing. (Variability of

language and uniformity of Speech Communities.) *Yuyan Jiaoxue yu Yanjiu* 5.

XU, DAMING. 2010. The Development of the Baotou Speech Community: A Quantitative Study of Nasal Variation in Mandarin Chinese. In, M.E. van den Berg and Daming Xu (Eds.), (pp. 120-140).

XU, DAMING. 2016. Speech community theory and the language/dialect debate. *Journal of Asian pacific Communication* 16/1: 8-31.

ZHAN, BOHUI. 1993. "Putonghua 'nanxia' yu Yue fangyan 'beishang' " 普通话 南下' 与粤方言 ' 北上' ["(Southbound" Putonghua and "Northbound" Cantonese)]. *Xueshu Yanjiu 学术研究* 4, 67-72.