

**Proceedings of the 13<sup>th</sup> International Symposium  
on Chinese Language and Linguistics  
(IsCLL-13)**

第十三屆中國境內語言暨語言學國際研討會論文集

*Edited by*

Yung-O Biq  
National Taiwan Normal University

Lindsey Chen  
National Taiwan Normal University

# Contents

## I . Keynote Speakers' Papers

<b>ON MACROVARIATIONS AND MICROVARIATIONS IN PARAMETRIC THEORY .....</b>	<b>1</b>
C.-T. JAMES HUANG	
<b>ON WORD ORDER, BINDING RELATIONS, AND PLURALITY WITHIN CHINESE NPs .....</b>	<b>19</b>
ŽELJKO BOŠKOVIĆ AND I-TA CHRIS HSIEH	
<b>AFFIX SUPPORT AND THE EPP .....</b>	<b>48</b>
NORVIN RICHARDS	
<b>IS UG LIKE A CHUNK OF SWISS CHEESE? .....</b>	<b>92</b>
YAFEI LI	
<b>ESTABLISHING TWO MORE STORES IN WORKING MEMORY TO BECOME A SKILLED CHINESE READER.....</b>	<b>105</b>
OIDVID J. L. TZENG, DENISE H. WU, ESTHER Y.-C. LIN, AND WEN JUI KUO	

## II . Presenters' Papers

<b>CLASSIFIERS AND MEASURE WORDS: EVIDENCE FOR A UNIFIED UNIVERSAL STRUCTURE .....</b>	<b>110</b>
ONE-SOON HER	
<b>NUMERAL CLASSIFIERS AND BARE NOMINALS .....</b>	<b>138</b>
BYEONG-UK YI	
<b>CLASSIFIERS IN ERSU .....</b>	<b>154</b>
SIHONG ZHANG	
<b>UNIVERSAL GRAMMAR AND REGIONAL GRAMMAR: CONSTRUCTING A TWO-TIER SYNTAX.....</b>	<b>184</b>
HSIN-I HSIEH	
<b>DISTRIBUTED MORPHOLOGICAL ANALYSIS OF MANDARIN REDUPLICATION .....</b>	<b>202</b>
WEI HU	

<b>THE HOLE OF <i>WHOLE</i> AND THE SILENT AMOUNT</b> .....	<b>230</b>
WEI-WEN ROGER LIAO	
<b>UNCERTAINTY REDUCTION AS A PREDICTOR OF READING DIFFICULTY IN CHINESE RELATIVE CLAUSES</b> .....	<b>245</b>
ZHONG CHEN, LENA JÄGER, AND JOHN HALE	
<b>MANDARIN DITRANSITIVES AND THE APPLICATIVE THEORY</b> .....	<b>262</b>
PEI-JUNG KUO	
<b>ON THE VP PERIPHERY AND DISPOSAL-AFFECTIVE ALTERNATIONS: A CASE STUDY OF <i>JIONG</i> CONSTRUCTION IN HAKKA</b> .....	<b>278</b>
JUI-YI ZOEY CHUNG	
<b>LANGUAGE UNIVERSALS AND THE TYPOLOGY OF MANDARIN MID VOWEL ASSIMILATION</b> .....	<b>304</b>
YEN-HWEI LIN	
<b>EFFECTS OF PERCEPTUAL ASSIMILATION ON L2 TONAL PROCESSING</b> .....	<b>326</b>
JUNG-YUEH TU	
<b>明清時期萬曆乾隆年間閩南語指示詞的演變: 從指示+樣式到程度加強副詞</b> .....	<b>343</b>
CHINFA LIEN	
<b>IS THE TAM SYSTEM A UNIVERSAL SYSTEM? THE INTERPLAY OF ASPECT AND MODALITY IN HAN PERIOD CHINESE</b> .....	<b>357</b>
BARBARA MEISTERERNST	
<b>客語兒童早期習得否定句的研究</b> .....	<b>387</b>
RUEIHLIRNG FAHN AND RUEI-ZHU LIN	
<b>CONTROL IN EARLY PRODUCTION OF MANDARIN-SPEAKING CHILDREN</b> .....	<b>413</b>
XIAOLU YANG AND CHENG YANG	
<b>FOCUS AND SCOPE IN CHILD LANGUAGE</b> .....	<b>434</b>
PENG ZHOU AND STEPHEN CRAIN	

# On Macrovariations and Microvariations in Parametric Theory

C.-T. James Huang  
*Harvard University*

The Principles-and-Parameters Theory (P&P), which took shape in the early 1980s, marked an important step forward in the history of generative grammatical studies. It offered a plausible framework in which to capture both the similarities and differences among languages within a rigorous formal theory. It enabled the discovery and explanation of grammatical patterns of variation across languages. Most of all, it offered an explanatory model for the empirical analyses toward meeting the challenge of “Plato’s Problem” posed by children’s effortless—yet completely successful—acquisition of their grammars under a condition of “poverty of the stimulus”.

Specifically, the P&P model led linguists to expand their scope of inquiry and enabled them to look at an unprecedented number of less studied languages, not only in syntax and phonology but also in semantics, and raising new questions about the nature of language which could not even have been formulated earlier. Another consequence was that it became perfectly reasonable, as has been happening every day, that one could discover properties of one language (say English) by studying aspects of a distinct, genetically unrelated language (say, Chinese or Gungbe), and vice versa.

Most of the original proposals of parameters in the early days of P&P were in the form of *macro*-parameters. They have the characteristic property of capturing the fact that parametric variations occur in clusters. As the theory developed, it became clear that such a model is inadequate for the description of numerous micro-scale parametric variations across languages. In addition, certain correlations that were expected in proposed well-known macro-parameters turned out not to hold as more languages were brought into consideration. In the mean time, considerations of theoretical parsimony led to widespread adoption of the lexical-parameterization hypothesis (known now as the Borer-Chomsky conjecture), ruling out much of the theoretical vocabulary used in earlier macro-parametric proposals. The developments led some to doubts about the existence of macro-parameters, and even the feasibility of the P&P program. With the advent of the Minimalist Program and its goal to go “beyond explanatory adequacy” to address the problem of language emergence within a brief evolution, further questions arise about the status of parameters.

In my talk, drawing upon recent discussions, I will support the position that both macro-parameters and micro-parameters exist, and there is really no real good alternative to a parametric model of language acquisition. Consistent with the “three-factors” conception of language design, parametric variations may be seen as different modes of externalization from a system with a simple set of unspecified functional or lexical items. Using Chinese examples as a paradigm case, I show that both macro-parameters and micro-parameters exist, and the tension between descriptive and explanatory adequacy is resolved by the view (Roberts and Holmberg 2010) that macro-parameters are aggregates of micro-parameters acting in concert, with correlating values as driven by economy considerations and conservative learning strategies.

## 1. The Principles-and-Parameters theory

- Plato's Problem and knowledge of language
- The logical problem of language acquisition: poverty of stimulus
- The biological model of grammar: Grammar = nature + nurture
- The P&P Model
  - a. UG → Primary Linguistic Data → PG [I-grammar]
  - b. UG = Principles + parameters
  - c. Language acquisition = fixing the parameter values
- The P&P model:
  - a. Solves Plato's Problem, the logical problem of language acquisition
  - b. Explains language uniformity, diversity, change
  - c. Consequence of linguistic methodology: you definitely can discover properties of the English language by investigating Chinese (or any other language), without knowing a word of English at all.

## 2. Principles and parameters in GB

- Some notable examples
  - Head parameter: [head initial] or [head final] (Stowell 1981, Huang 1982)
  - Null subject parameter (Rizzi 1982, and many others): [yes] or [no]
  - Null topic parameter (Huang 1984): [yes] or [no]
  - Wh-movement parameter (Huang 1982, etc.): [overt] or [covert]
  - Non-configurationality parameter (Hale 1983): PP [yes] or [no]
  - The Polysynthesis parameter (Baker 1996): [yes] or [no]
  - Nominal mapping parameter (Chierchia 1998a,b): [±arg], [±pred]
  - Relativized X-bar parameter (Fukui 1986): FP [yes] or [no]
  - The DP/NP parameter (Bošković 2010, etc.): language with D [yes] or [no]
- **Macro-parameters:** Most of the parameters proposed have the character of being *macro-parameters*, in that their effects are observed across the board (e.g. the head parameter) and that they usually capture the fact that observed variations often *cluster* together. Examples:

- Head parameter: V-final, N-final, P-final, A-final, etc.
  - Null subject parameter: null subject, free inversion, apparent long subject-extraction [Rizzi 1982]; differences between French and Italian long clitic-climbing and infinitival V-movement [Kayne 1989, 2012], etc.
  - Nominal mapping parameters: bare argument, classifier, no plural marking, no determiners, etc.
  - The DP/NP parameter (Bošković 2008, 2010, etc.): left branch extraction, adjunct extraction from NP, scrambling, adnominal double genitives, superlatives with more than half reading, IHRC and locality, etc.
  - The Polysynthesis Parameter (Baker) [=Morphological Visibility condition]
    - A phrase X is visible for theta-role assignment from a head Y only if it is coindexed with a morpheme in the wording containing Y via:
      - An agreement relation (pronominal affixes) [a spec-head relation]
      - A movement relationship [incorporation] [antecedent-trace relation]
    - Yes on both: Mohawk, Nahuatl, Mayali, ...
    - Yes on agreement, no on incorporation: Warlpiri, Navajo, etc.
    - No on both: English, French, Chichewa, Japanese, Hindi, German, etc.
- The explanatory value of macro-parameters
- Problems encountered by macro-parameters
    - Empirical: certain macro-parametric predictions are often not borne out
    - Theoretical: (i) extra burden on linguistic theory, (ii) arbitrariness in where variation may or may not occur.
  - **Micro-parameters:**
    - Lexical parameterization hypothesis (Borer 1984 and Chomsky 1995): highly restrictive theory, by reducing all parametric variations to features of (functional) lexical items.
    - The existence of *micro-parameters* that may or may not have wide-ranging effects. Examples
      - Good enough, \*enough good, sufficiently good
      - John is likely to go, \*John is probable to go, etc.
      - ‘Pre-classifier *one-drop*’ (+ or -) in Mandarin, Cantonese, Taiwanese, etc.

### 3. Principles and parameters in Minimalism

- **Two goals** of the (exploratory program known as) Minimalism: ‘beyond explanatory adequacy’—beyond explaining the ‘poverty of stimulus’ problem:
  - a. Theoretical: in pursuit of the beauty of scientific theory
    - Galileo’s belief that there is a very simple explanation for Nature
    - Cf. Chen-Ning Yang (1982): three levels of beauty in science
  - b. Empirical: to explain the ‘brevity of evolution’ problem

- **The MP's** proposed answer to these goals: Reduce the size and options of UG
  - a. Three factors of language design (Chomsky 2006):
    - Genetic endowment: UG
    - Experience: PLD
    - Other independent systems that are recruited in language use, such as other cognitive abilities (logical reasoning, memory), computational efficiency, minimality, and general laws of nature
  - b. Many 'principles of grammar' are ascribed to the 3<sup>rd</sup> factor
  - c. All parameters are stated as micro-parameters

- **Problems for P&P in MP**

There are numerous empirical problems that have arisen under the very restrictive proposals in the programmatic MP, but that have been left untouched by the MP practitioners. Many of the results of the P&P in GB-Barriers-style have not been carried forward empirically, in part because some of the mechanisms and notions introduced earlier have been made unavailable. One important area is the subfield of locality theory and government. Another is the status of Parametric Theory:

- Since traditional macro-parameters cannot be stated within Minimalist vocabulary, all parametric variations are micro-parametric variations and stated as variations in the nature of formal features of individual lexical or functional categories. This means that traditional macro-parameters are completely excluded as such.
- Do macro-parameters exist, and if so, how are they accommodated in UG?
- Given the large number of micro-parameters as based on individual lexical items, doesn't Plato's Problem arise again?

- **Some recent views**

- Baker (2008)
- Kayne (2005, 2012)
- Holmberg (2010)
- Holmberg and Roberts (2011)
- Also Longobardi (2005) and Baker (2003)

- **Baker (2008):** argues for the need for macroparameters in addition to microparameters.

- Shows that certain macroparameters indeed go a long way towards reducing the range of actual occurring variations

- “The strict microparametric view predicts that there will be many more languages that look like roughly equal mixtures of two properties than pure languages, whereas the macroparametric-plus-microparametric approach predicts that there will be more languages that look like pure or almost pure instances of the extreme types, and fewer that are roughly equal mixtures.” (p. 361). The picture seems to fit the second vision, as two examples show.
  - Head directionality: (from World Atlas of Language Structures)
    - V-O P-NP = 417 (consistent languages)
    - O-V NP-P = 427 (consistent languages)
    - V-O NP-O = 38 (inconsistent languages)
    - O-V P-NP = 10 (inconsistent languages)
  - Many macro-parameters could probably never have been discovered simply by comparing dialects of Indo European languages
- **Kayne (2005-2012)**
    - The reality and existence of microparameters
      - Microparametric differences are everyday life phenomena
      - The value of looking at dialectal differences to detect many useful generalizations that would not have been possible from a macroparametric investigation
    - The methodology of microparametric research
      - Restrictiveness of parametric theory (to lexical parameterization)
      - Something close to a ‘controlled experiment’
    - Many or most (or all?) macroparameters could be broken up into small, microparameters: “Kayne 2012: It might also be that all ‘large’ language differences, e.g. polysynthetic vs. non- (cf. Baker (1996)) or analytic vs. non- (cf. Huang (2010)), are understandable as particular arrays built up of small differences of the sort that might distinguish one language from another very similar one, in other words that *all parameters are micro-parameters*. [emphasis added]”
    - What about the problem of language acquisition—so many parameters and how many languages?
      - Should not be a huge acquisition problem, if parameters are kept to a simple form. Learners are not asked to pick a language from 8 billion.
      - $2^{32} = 4,294,967,296$ ,  $2^{33} = 8,589,934,592$
      - $2^{100} = 1,267,651,000,228,229,402,496,703,205,376$   
一百二十六萬七千六百五十一兆兆 1,267,651 trillion trillion



- **Holmberg (2010):** it is equally possible to consider parameters as *underspecifications* in UG, and entirely in line with minimalist considerations.
  - “A parameter is what we get when a principle of UG is underdetermined with respect to some property. It is a principle *minus* something, namely a specification of a feature value, or a movement, or a linear order, etc.”
  - The change in theoretical perspective [brought about by the Minimalist Program] does not, in itself, invalidate the aims, or the methods, or the results achieved in the GB-era, nor is it inconsistent with P&P theory, once parameters are seen as points of underspecification, where, crucially, the alternative specifications are always strictly limited in number.

#### 4. Claims and proposals

- A:** Both macro-parameters and micro-parameters are needed in linguistic theory. Evidence from Chinese, and other languages.
- B:** Macro-parameters are simply aggregates of micro-parameters acting in concert on the basis of a conservative learning strategy (a la Roberts and Holmberg 2010).
- C.** The (micro-) parameters are themselves hierarchically organized (Roberts and Holmberg 2010, cf. also Longobardi 2005, Baker 2003), so fixing one parameter higher on the hierarchy automatically eliminates certain values for other parameters.

#### 5. Point A: Both macro-parameters and micro-parameters are needed in linguistic theory.

##### 5.1. Synchronic variations: macro-parametric properties of Modern Chinese (as opposed to e.g., English):

- (1) Modern Chinese:
  - a. Light verb constructions: 打鱼 ‘do fish’、打电话 ‘do phone’、打喷嚏 ‘do sneeze’. English: simply *fishi*, *phone*, *sneeze* (denominalization)
  - b. Pseudo-incorporation (Massam 2002) a.k.a. “phrasal compound”: 捉鱼 ‘catch fish’, 吃饭 ‘eat rice’, 剥皮 ‘peel skin’. English: *to fish*, *to feed*, *to skin*.
  - c. Accomplishments are complex: Compound and phrasal accomplishments: 打破、弄破、踢破. English: simplex *break*.

- d. Need overt classifiers for count nouns: 三本书. English: 3 books.
  - e. Need overt localizers: 走到桌子旁边. English: went to the table.
  - f. Canonical “Kaynean word order”: Subject-Adjunct-Verb-Complement. (= “V2 counting from the right”)
  - g. Wh-in-situ (instead of wh-movement)
  - h. No forms equivalent to *nobody*, *each other*, bi-nominal *each*, etc.
  - i. Adverbial vs. adjectival *fast*, *all*
  - j. No determiner (but numeral *one* or determiner *this*, *that*)
  - k. No coercion (e.g. *begin the book*) (Lin 2005)
  - l. No (canonical) gapping
  - m. No ‘ga-no conversion’
  - n. Syntax-semantics mismatches (三天書), etc.
  - o. Analytic passivization
- Note the clustering of these properties in the same language to the exclusion of them in, say, English. The properties (some of which having been attributed to some macroparameters) suggest a **mega-, macro-parameter of analyticity vs. synthesis**. This clustering is *macroparametric* in nature.

## 5.2. Macro-parametric properties of Old Chinese (vs. Modern Chinese)

- (2) Old (or Archaic) Chinese (500 BC to 200AD)
  - a. No light verb (but denominalization): 渔 *yu* ‘to fish’
  - b. No pseudo-incorporation: *fan* ‘have rice’
  - c. No compound: simplex accomplishments. 破 *po* ‘break’
  - d. No overt classifiers for count nouns (no need for ‘light noun’)
  - e. No need for overt localizers (no need for ‘light noun’)

- g. NP-movement (passives)
  - h. Wh-movement (吾谁欺 *wu shei qi* I whom deceive?)
  - i. *suo*-movement (我所欲 *wo suo yu* that which I want)
  - j. focus-movement (唯命是从 *wei ming shi cong* ‘only order have I followed’)
  - k. post-verbal adjuncts (易之以羊 *yi zhi yi yang* ‘exchange it with a sheep’)
  - l. Has canonical gapping
  - m. *qu* = ‘leave, depart from’ in OC (vs. = ‘go to’ in MnC)
  - n. has ‘ga-no conversion’ (之所 V、主之谓)
  - o. and more.
- Note again the clustering of properties that occur in OC to the exclusion of MnC. This clustering is *macro-parametric* in nature.

### 5.3. Macro-parametric properties of other languages, genetically unrelated languages.

- Similar clustering of properties in other languages. Gungbei, etc. (Badan, Cheng, Aboh, etc. Also Thai and other South East Asian languages)
- Point A: Macro-parametric variations exist, and the clustering of properties must be captured.

### 5.4. Micro-parameters clearly must also exist to account for smaller differences. Some micro-parameters exhibit certain degrees of value-clustering:

Synchronic micro-variations among Chinese dialects

- (3) Classifier stranding: (Cheng&Sybesma 2005)
- yi ben shu* ‘one classifier book’ → ~~yi~~ *ben shu* ‘classifier book’
- Mandarin:       \* Subject, ok Object
  - Cantonese:     ok Subject, ok Object
  - TSM:            \* Subject, \* Object

- (4) Development of postverbal suffixes
  - Mandarin: some aspectual suffixes (*zhe, le, guo*)
  - Cantonese: More postverbal suffixes
  - TSM: Fewer
- (5) Verb object order preferences (Liu 2002, Tang 2006)
  - Mandarin: mildly OV or VO
  - Cantonese: strongly VO
  - TSM: strongly OV
- (6) Position of *qu* ‘go’ (Lamarre 2008, etc.).

For the expression ‘Zhangsan went to Beijing’, Mandarin allows both the analytic and the synthetic strategy:

- Analytic: Zhangsan dao Beijing qu le.  
Zhangsan to Beijing go Perf ‘Zhangsan to Beijing went.’
- Synthesis: Zhangsan qu-le Beijing.  
Zhangsan go-Perf Beijing ‘Zhangsan went-to Beijing.’

Cantonese allows only the synthetic strategy involving head-movement, whereas Pre-Modern Chinese allows only the analytic strategy, with no head-movement.

- Mandarin: 张三到北京去了、张三去了北京。
- Cantonese: \*张三到北京去了、张三去了北京。
- MQ 朴通事: 张三到北京去了、\*张三去了北京。
- TSM: 张三对北京去矣、张三去北京矣。

→ Point A: Microparameters also exist, with varying degrees of clustering. See more examples of microparameters below.

**6. Point B: Macro-parameters as aggregates of micro-parameters with correlating values (Holmberg and Roberts 2010). Cf. also Longobardi (2005).**

- (7) Macro-parameters as aggregates of micro-parameters with correlating value
  - a. Aggregates:** each parameter is a micro-parameter that falls under the lexical parameterization hypothesis, limiting variation to formal features of lexical items.
  - b. Correlating values:** the micro-parameters acting in concert for reasons of markedness or conservatism.

c. **Markedness:** conservatism, “generalization of the input” as indicated below:

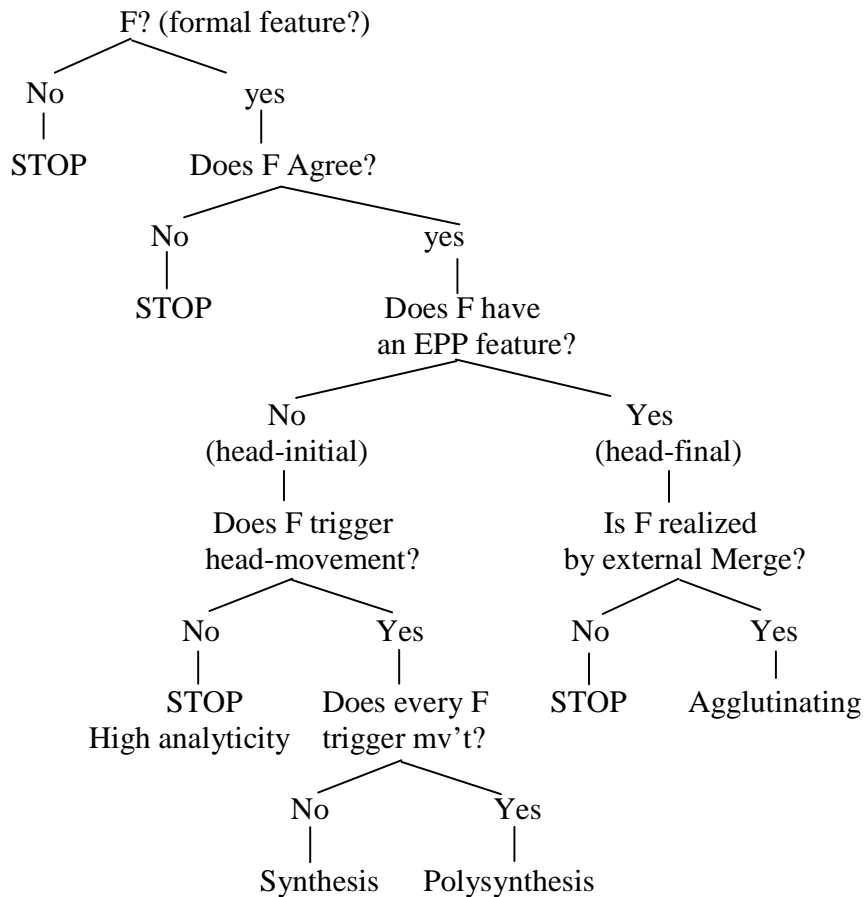
- If H *v* is [+strong] then H *n, a, p* are also [+strong]; [-strong] otherwise
- If H *v* is [+EPP] H *n, a, p* are also [+EPP]; [-EPP] otherwise.

→ Point B:

- There is a conservative learning strategy that generalizes the known to the unknown → “generalization of the input”
- The result of extensive generalizations gives rise to ‘clustering effects’ → aggregates of micro-parameters acting in concert → macro-parameters

7. **Point C: The hierarchical organization of (micro)-parameters** (following Roberts & Roussou 2003, Holmberg and Roberts 2011), organizes the following set of options relating to a given formal feature F on the basis of grammaticalization as a diachronic operation affecting functional categories.

(8)



- The elements of parameterization are hierarchically organized, and this reduces the potential number of options that a child has, thereby easing the abductive learning procedure. Hence, Plato's problem is solved.
- Macro-parameters are not primitives: they are the aggregates of microparameters that can be localized to lexical items and their features, where differences exist. UG is reducible to a bare minimum, hence further approaching 'beyond explanatory adequacy'.
- To be illustrated for Modern Chinese, Old Chinese, and Modern Chinese dialects, etc.

## 8. Implementation: Back to Chinese

- Summary

- There are macroparameters which capture important, sometimes sweeping, clusters of typological properties. There are also microparameters with less or no clustering effects.
- The macroparameters give rise to general patterns, while the microparameters give rise to the mixed, exceptional cases, plus details of variation or change.
- Macroparameters are aggregates of microparameters acting in concert for reasons of conservative learning strategy.
- Parameters conceived as cases of underspecification do not add to the burdens of UG and are consistent with minimalist theorizing.
- Microparameters are hierarchically organized so the number of occurring options to choose from is greatly reduced, as is the burden of the child.
- Binary-value parameters are not difficult and a small number of them are enough to give rise to large numbers of linguistic variations (languages, dialects, idiolects).

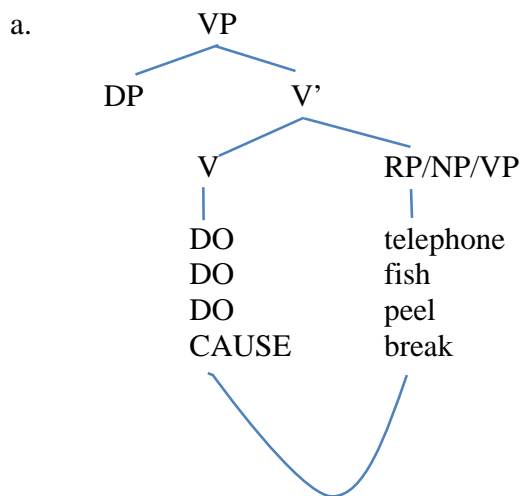
- Chinese macro-properties:

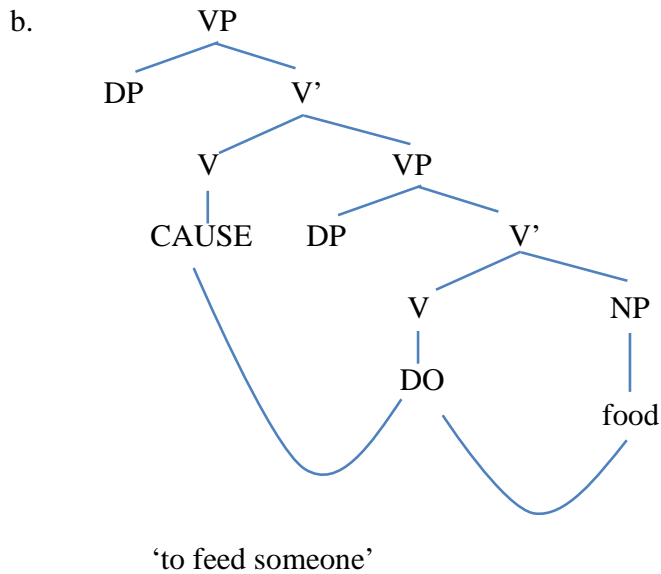
- Viewed synchronically: Modern Chinese exhibits high analyticity in a macro-parametric way, in systematic contrast to other languages, at all levels: lexical, functional, and the level of argument structure.
- Viewed diachronically: Old Chinese underwent macro-parametric change from a substantially synthetic language to a highly analytic language, at all levels, with analyticity peaking at the end of the 6-dynasties (late Middle Chinese), followed by small-scale new changes that result in the major dialects of Modern Chinese, with some varying degrees of small-scale synthesis. (A linguistic cycle: synthetic → analytic → synthetic → ...)

- How to characterize the macroparametric properties?
  - How about an ‘analytic-synthetic’ parameter, with the features [ $\pm$ analytic], [ $\pm$ synthetic], so that Chinese is [+analytic, -synthetic], Old Chinese (and say English) are [-analytic, +synthetic], and some other languages are [+analytic, +synthetic]?
  - The answer should be no, because:
    - Such a view is purely descriptive and does not reveal the real nature of linguistic variation. For one thing, there are exceptions that must be accounted for, and such exceptions must resort to micro-parametric descriptions. A binary-value parameter cannot reveal the nature of ‘continuum’ that characterizes cross-linguistics variations and diachronic changes.
    - Such a view makes use of concepts unavailable in the theoretical vocabulary of current linguistic theory, against the spirit of minimalist theorizing.

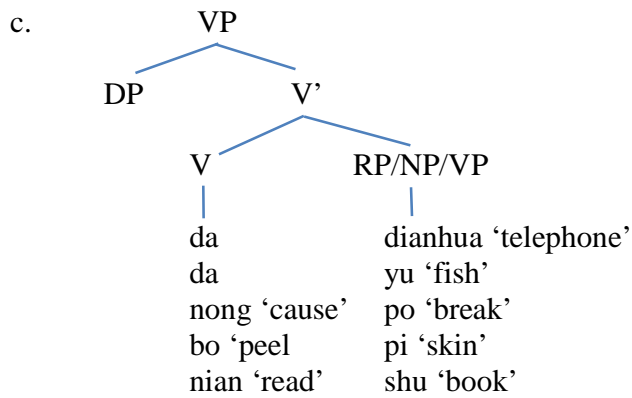
### 9. Implementation: lexical and phrasal domains

(9) Derivational morphology / syntax:

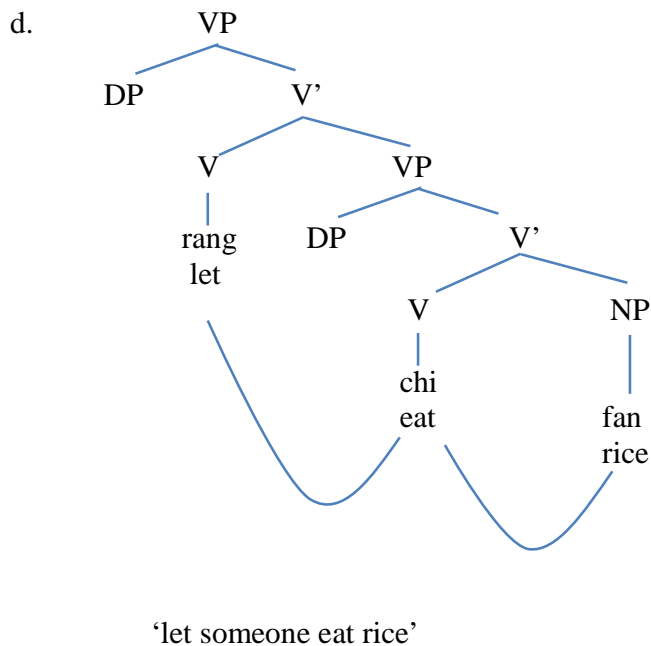




English *v*'s may be in the form of a phonetically null light verb DO or CAUSE, which are assumed to have properties: both have F features which need to Agree, do not contain EPP, and do trigger head-movement. By (7), this leads to synthesis, and English abounds in *simplex* denominal verbs like *telephone*, *fish*, *peel* and *simplex* causatives like *break* or *feed*.







By contrast, Chinese does not have the CAUSE and DO in lexical structure. It needs to resort to lexical (light or heavy) verbs which do not trigger head-movement, leading to STOP in (7) and high analyticity. Instead of simplex denominalized action verbs or simplex causatives, Chinese resorts to more complex expressions, and abounds in have light verb constructions, pseudo-incorporation, resultative compounds or phrases, and periphrastic causatives.

- **Clustering as aggregation:**

- (10) By (7c): If H *v* is [+strong] then H *n*, *a*, *p* are also [+strong]; [-strong] otherwise. Chinese has lexical classifiers, nominal localizers, adjectival degree marker, and prepositions, while English has *all* such categories in null or affixal form.
- (11) a. Chinese count nouns are formed by an overt ‘light noun’ (classifier):
- [CL *ben* [NP *shu* ]] = count noun
  - By (8), the light noun does not trigger head movement → “*ben shu*” is the Chinese ‘count noun’, i.e., an analytic ‘count noun phrase’.
- b. English count nouns are formed incorporating the noun root into an empty CL-head (cf. Borer 2005):
- [CL CL [NP *book* ]] → [CL *book* [NP *t* ]] = count noun
  - By (8), the CL has formal feature that Agrees, has no EPP, triggers head-movement → the count noun = synthetic.

- (12) a. Chinese forms locational NPs with overt localizers  
 □ [zhuozi [nali]] (nail 'place' = no head-movement → analytic)
- b. English forms such NPs by incorporating silent PLACE (cf. Kayne 2010)  
 □ [table [PLACE]] → table → synthetic
- (13) a. Chinese adjectives have lexical *hen* which marks absolutive degree: *hen hao*.
- b. English adjectives incorporate into null HEN → synthetic
- (14) a. Chinese complex PPs take 'discontinuous' form:  
 □ zai zuozi pangbian 'by the table side' → analytic
- b. English complex PPs are formed by incorporation:  
 □ be- the table –side → beside the table → synthetic

### 10. Implementation: clausal, inflectional domain

- (15) Mandarin Chinese has aspectual suffixes that:
- Are functional (grammaticalized) and formal
  - Enter into Agree with appropriate verb stems
  - [-strong], so do not trigger (overt) head-movement

Zhangsan [<sub>ASP</sub> PERF] [<sub>VP</sub> zuotian qu-le Kaohsiung] (PERF Agrees with *le*)  
 Zhangsan yesterday go-le Kaohsiung

- (16) English T and Asp heads are similar in this respect to Mandarin
- These clausal heads are functional
  - They enter into Agree with the inflected verbs
  - They do not attract the inflected verbs to T

John [<sub>T</sub> TNS] [<sub>VP</sub> often kisses Mary in the kitchen] (TNS Agrees with *kisses*)

→ Thus, English is only synthetic, but not polysynthetic, according to (8), since only some, but not all, Fs trigger head movement.

→ Chinese is more consistently analytic than English is synthetic. Markedness (7c) applies more broadly in Chinese to clausal heads as well.

### 11. Old Chinese

- Lexical categories: In the lexical domain, Old Chinese is similar to English

(17) Like English, Old Chinese possessed null DO and null CAUSE as higher lexical heads (both reconstructed as \*s- by Mei Tsulin 1973, and references). This gives rise to the properties of (2a-c):

a. No light verb, but denominalization: 渔 *yu* 'to fish': [VP \*s [NP *yu*]] → *yu*<sub>verb</sub>

b. No pseudo-incorporation: *fan* 'have rice': [VP \*s [NP *fan*]] → *fan*<sub>verb</sub>

c. No compound: synthetic accomplishments. *po* 'break'  
[VP \*s [VP-inchoative *po*]] → *po*<sub>verb-causative</sub>

And by markedness (7c), also:

d. No overt classifiers for count nouns (no need for 'light noun')

e. No need for overt localizers (no need for 'light noun')

- Clausal functional heads

(18) Old Chinese TP differs from Modern Chinese in the nature of a clausal functional head (or more than one) in the TP region, right below the subject. Let's call it FP (possibly focus phrase):

- F has the properties of an unvalued feature that requires it to agree with an appropriate element and an EPP feature requiring XP movement. This gives rise to the following XP movements in OC:

a. NP-movement (passives)

b. Wh-movement (吾谁欺)

c. *suo*-movement (我所欲)

d. focus-movement (唯命是从)

e. post-verbal adjuncts (易之以羊)

- Possibly the F also triggered head movement, giving rise to canonical gapping (Wu 200?, He 2010):

f. Has canonical gapping

## 12. Modern Dialects

- Microvariations with small degrees of clustering among Mandarin, Cantonese and TSM.
- Relatively speaking:
  - Cantonese has undergone more grammaticalizations, and is the least analytic (or most synthetic) of the three dialects
  - Mandarin (and Shanghai) have developed some suffixes but remain analytic in that these suffixes do not trigger head movement. They may undergo Affix Hopping.
  - Taiwanese SM remains the most analytic, in having developed the least number of suffixes. Most suffixes in other dialects remain full verbs in TSM.
- Such micro-parametric differences are sure to increase when more dialects are examined, either contemporary dialects or dialects at any given historical stage.
- Hence although there is the appearance of macro-parametric changes from OC to Modern Chinese, the truth must be that actual changes took place on a micro-parametric level.

## 13. On some micro-parametric properties of TSM vs. Mandarin

- (19)
- a. Distribution of aspectual markers:
    - Mandarin: aspect
    - TSM: resultative verb
  - b. Aspectual light verb:
    - Mandarin: [null ASP ... V-*le*]
    - TSM: [lexical *you* ... V]
  - c. Denominal or action VP
    - Mandarin [null DO + VP]
    - TSM: [lexical *zuo* + VP]
  - d. Position of object:
    - Mandarin: postverbal preferred
    - TSM : preverbal preferred

- e. Verb complement construction with objects
    - Mandarin: ok [ [V-R compound] + object]
    - TSM: requires object in preverbal position
  - f. Causative constructions
    - Mandarin: allows V-to-CAUSE raising (at syntactic level)
    - TSM: does not
  - g. Outer objects with V-movement
    - Mandarin: allows both constructions
    - TSM: does not
  - h. *hoo hoo hoo*
    - TSM: a full range of meaning of *hoo*
    - Mandarin: impoverished paradigm
  - i. *ka* vs. *ba*
    - TSM: *ka* has a wider range of semantic band-width
    - Mandarin: *ba* has a narrower band-width
  - j. Classifier stranding
    - Mandarin: allows one-drop in certain positions
    - TSM: does not allow one-drop at all
- Clustering as aggregation: results of (7c) = generalization of the input.
  - There are already 10 micro-parametric differences here between Mandarin and TSM. Logically there could be  $2^{10} = 1024$  independent TSM dialects that differ from each other by at least parameter. But it is unlikely that these parametric values are equally distributed. Rather, the likely norm is that they cluster together with respect to certain values. → medium-size macro-parameter.
  - Possibly all these parameters simply express special cases of TSM as *consistently more analytic than* Mandarin.
  - Finally, not all speakers agree on the observations made above, thus reflecting dialectal/idiolectal differences from me. This is not surprising, as microvariations typically arise among individual speakers.

#### 14. Summary and conclusion