Placing Chinese Morphemes in a Coordinate System: Reflecting on the Scientific Nature of Linguistic Theorization

Yafei Li

Nanjing University
University of Wisconsin-Madison

Chinese resultative words offer evidence that derivational morphology is not always a meaningful concept for languages, and that a coordinate system measuring the degrees of shift in semantic content, phonological content and boundness can more accurately handle the facts of word formation and reflect a language's unique historical heritage. This approach raises issues with the methodology of the Principles-and-Parameters model of language, which is critically examined in the second half of the paper.

Key words: coordinates, derivational, Galilean idealization, resultative

This paper consists of two parts. Part 1 articulates a framework in which Chinese morphemes and morphologically complex words can be more accurately described and classified. The framework is shown to help understand or clarify some synchronic and diachronic issues surrounding Chinese word formation. It also raises methodological questions about the mainstream practice of syntactic theorization, which is addressed in Part 2.

1. Morphemes in a coordinate system and what the system can do for us

It is common practice in the study of word-formation to group morphologically complex words into discrete classes, the most commonly used of which are inflectional, derivational and compounding (see Spencer 2001). A question from the theoretical perspective is whether such classifications are crosslinguistically valid or even useful. This section argues that at least in Chinese, derivational morphology is not a meaningful concept, that trying to squeeze various morphemes into this slot either is ad hoc or obscures certain salient properties of Chinese. Instead, a coordinate system in which each morpheme is measured and situated according to how far it shifts from its lexical/ historical origin in semantics, phonology and boundness not only avoids brute-force grouping, but also captures various synchronic and diachronic facts and correlations with more naturalness and accuracy.

1.1 *Si*: To die or not to die, that is only part of the question

Consider how *si* 'die' is used inside a word:

(1) a. si-zhan 'die-fight = fight mortally'
     b. si-ke 'die-knock = conflict incessantly'
     c. si-ying 'die-hard = unbendingly stubborn, extremely hard'
(2) a. da-si 'beat-die = beat to death'
     b. fan-si 'bother-die = be extremely bothersome'
     c. nan-si 'difficult-die = extremely difficult'

The two groups illustrate *si* 'die' to be either the first or second morpheme of a verbal compound. In the first example of each group (1a)–(2a), *si* retains its lexical meaning (at least in the default use of the compound), whereas in all other cases (1b and c)–(2b and c), *si* no longer means 'to die', but instead expresses an extreme degree of intensity appropriate for the event or state coded by the other morpheme.

In isolation, the semantic shift of *si* in certain contexts is not surprising—English also allows *X bored Y to death*. But the shift is related to two other facts in Chinese and thus raises corresponding questions. First, the compounds in (2) are the resultatives (r-words hereafter) extensively investigated in the literature (Cheng & Huang 1994; Gu 1992; Her 2007; Li 1990, 1993, 1995, 2005; Lin 1990; Starosta et al. 1998; Zou 1995; among many others). An r-word has the form of V-R, with V denoting the event/situation of cause and R that of the result.[1] Important to this work is that in the default case, both V and R carry their own tones and R bears either an identical stress as V or a stronger one. But *si* as R in (2) departs from this pattern by being almost always unstressed and toneless whether it exhibits the semantic shift or not. Furthermore, this phonological depletion cannot be attributed to *si* being inside a compound because in (1), where *si* is the first morpheme of a compound, it necessarily carries the full stress and the base tone. How is this unique trait of *si* while in the R position to be understood in the general context of Chinese grammar?

The second fact can be illustrated through the comparison below.

(3) a. ta   gen   wo   si-ke-le   san   nian.
         he   with   me   die-knock-ASP   three year
         'He was a determined enemy of mine for three years.'
     b. zhe   ji-ge   ren   fan-si   wo   le.
         this   few-CL   person   bother-die   me   LE
         'These few people bothered me to death.'
(4) a. zhe-pian   mianbao   si-ying,   genben   buneng   chi.
         this-CL   bread   die-hard   totally not.can   eat
         'This slice of bread is extremely hard. It's not edible.'

---

[1] For ease of discussion, both V and R are taken to be verbs. The fact that some such morphemes may likely be adjectival is of no significance to this work. See my discussion of Chinese adjectives in Chapter 2 of Huang et al. (2009). Also worth pointing out is that the English translations in (1) and (2) are only approximations to the Chinese examples and should not be depended on for analysis.

b. na-ci   kaoshi   nan-si-le.
    that-CL   exam   difficult-die-ASP
    'That exam was so hopelessly difficult.'

For a speaker of Chinese, the 'intensifying' *si* in (3) is in fact an exaggerative use of the morpheme. (3b), for instance, may be literally translated as 'those people are bothering me so much that they are killing me'. In other words, there is still a thematic relation between *si* 'die' and an argument of the sentence—someone will die, figuratively. But the thematic relation is completely lost in the examples in (4). The slice of bread and the exam are not thematically connected with *si* in any way, figurative or not. It may be said that *si* is truly used as an intensifier in these contexts.[2]

At the descriptive level, (3)–(4) are merely manifestations of gradational semantic shift, which is extensively discussed in terms of grammaticalization in the field of Chinese linguistics. For the theory of morphology, however, they raise the question of proper classification. While the *si*-containing words in (3) may still be taken as compounds with one of the morphemes inside interpreted figuratively, it is no longer as obvious which class the corresponding words in (4) fall into, especially when compared with other parts of Chinese morphology.

Consider the diminutive morpheme *-r/-er*[3] in many northern dialects of Chinese.

(5)   *hua-r* 'flower-R', *xiao-yu-r* 'little-fish-R', *hudie-r* 'butterfly-R', …

Etymologically, the morpheme comes from *er* 'child', which is still used in modern Chinese to mean 'son' or sometimes 'child'. Roughly put, the semantic change lies in removing from the original lexical item [+human, –adult] and keeping only [+small-sized]. By standard morphological classification, the diminutive *-r/-er* is a typical instance of derivational affixes. At the semantic level, exactly the same can be said of the thematically bleached *si*, with all defining features removed except the high-intensity element of the dying event. However, this use of *si* is less clearly derivational. First, its linear relation with the other morpheme correlates with perceptibly different semantics—*si*-V encodes a modifier relation and V-*si* a causal one.[4] This order-meaning correlation

---

[2] One may suspect that in (4b) there is a phonologically empty object because *nan-si* can indeed take an overt one: *na-ci kaoshi nan-si wo le* 'that-CL exam difficult-die me LE = That exam is so difficult that it kills me.' But such an analysis is refuted for two reasons. First, native speakers do not have the clear sense that (4b), when said in isolation, necessarily has an understood object. Secondly, there are examples where adding an overt object actually makes them less acceptable:

(i) zhe-tiao   lu   chang-si (??/*wo)   le.
     this-CL   road   long-die   me   LE
     'This road is so hopelessly long (to me).'

[3] Some subdialects treat the morpheme as a full syllable.

[4] When asked about *chang-si* 'long-die' and *si-chang* 'die-long', two native speakers without any linguistic background both said the latter is 'more objective'. I take this to mean that *si*-V expresses a simple—hence perceptually more matter-of-fact—modifier–modifiee relation, meaning 'extremely/unbearably V'. But V-*si* still somehow retains the basic causal relation, with *si* corresponding to the result, which is pragmatically tied to the mental state of the speaker.

is characteristic of compounds (see Jackendoff's 2002 discussion of *dog house* vs *house dog*). Second, dividing the words in (1) and (2) into the discrete classes of compounding and derivation may be convenient for the standard theory of morphology, but obscures a unique fact about Chinese, namely the different word-internal uses of morphemes like *si* are all intrinsically related etymologically and/or historically, displaying in effect a continuum from the most lexical to the least lexical. This brings us to §1.2.

1.2 The axes to measure shifting

We start with *r*-words.

With respect to the semantic shift of R, one can identify at least four stages in terms of how far a given R 'deviates' from the base verb. At one end are those that retain the base meanings. Almost all examples in Li (1990) belong to this type. S1 will be used to refer to R at this stage.

The next stage, S2, is represented by *zhu*. Used alone, it is often translated as 'reside in, live in'. But as R in an r-word, it always means to 'stay in one place'. The third word in (6), for instance, means to 'tie … securely'.

(6)   yao-zhu 'bite-ZHU', zhua-zhu 'grasp-ZHU', bang-zhu 'tie-ZHU', …

As it turns out, glossing the base meaning of *zhu* as 'reside in' is not accurate. Instead, it really means 'to stay in a residential place'.[5] It follows that *zhu* as R in (6) constitutes a shift from being specific to being generic. Another example is *diao*. When used alone, *diao* has the non-Agentive meaning 'to move downward away from the initial position' and therefore always gets translated as 'drop'. But as R, the 'downward' direction is removed from its semantics, resulting in r-words like *reng-diao* 'throw-DIAO = throw away' and *na-diao* 'take-DIAO = remove, take off'.

Skipping the third stage for now, at the farthest end of the semantic shift is S4, including the aspectual suffixes *-le*, *-zhe* and *-guo*. Take *-le* for example. In the Tang Dynasty, the resultative construction took the form of *chi fan liao* 'eat meal finish', which later changed to *chi liao fan* 'eat finish meal' (see Mei 1994). Eventually, *liao* in this position (i.e. as R) lost its lexical meaning and became the marker *-le* for perfective aspect. In fact, my own dialect still allows this perfective *-le* to be pronounced as *liao* in a playful context. *-zhe* 'progressive aspect' and *-guo* 'experiential aspect/past tense' both have corresponding lexical ancestors and similar routes of diachronic change.

The existence of S3, represented by the thematically depleted use of *si*, is best appreciated when compared with S4. In terms of semantic content, this *si* can be properly defined as containing only

---

[5] Evidence for this translation comes from the verb's ability to take a hotel room as the object: *ni zhu na-ge fangjian?* 'you ZHU which room?' is a natural question between two guests in a hotel. Note how bizarre it is to ask *Which room do you live in?* in that context.

[+high-intensity] and therefore is comparable to the single-featured functional suffixes in S4 (e.g. *-le* is presumably nothing but [+perfective]). On the other hand, while aspect is crosslinguistically manifested through structural means typical of functional categories (structural location in a clause, morphological form, etc.), languages are not known to give the content of intensity a uniform functional treatment. At least in Chinese, the distinction between the theta-less *si* and the perfective *-le* seems to lie plausibly in the former being lexical and the latter functional. Note that this characterization is not limited to R. The preverbal *si* in (4a) has the same semantic content. By the same token, the diminutive *-r/-er* is single-featured (i.e. [+small-sized]) and lexical.

As R's semantic variations have at least these four stages—fully lexical → genericized → lexically single-featured → functional—it is no longer straightforward to match the r-words with the standard morphological classifications. With R = S1 or S4, V-R can certainly be taken as a compound or the product of 'inflection' respectively. But simply squeezing the rest into any such discrete class would be arbitrary in one way or another—do the intensifying *si* and the diminutive *-r/-er* belong to the same class or not, and why? It would also muddle a non-trivial fact about Chinese: that there exists a diachronic trajectory from a lexical R to a functional suffix which is explicitly reflected in modern Chinese. The solution to all these concerns is an axis measuring the amount of semantic content which a morpheme still keeps from its base form, with 0 = no semantic shift and 1 the maximal shift.

(7)   The semantic (*s*) axis

        S1   S2   S3   S4
        ●─────────────●─────────────●─────────────●──────────→
        0   1
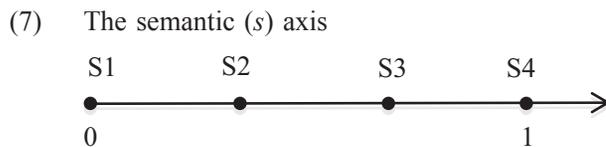
A clean classification of r-words is made even less practical with multi-stage shifts of R in two other dimensions. First, it is possible to define morpheme boundness with two tests: whether R with its particular value on the *s*-axis can be used as a standalone verb and whether the modality infixes *de* 'can' and *bu* 'cannot' may occur inside V-R.

(8)   *xiao* 'laugh, smile':
        a.   tamen   xiao-le.
               they   smile-ASP
               'They smiled.'
        b.   dou-(de/bu)-xiao
               cheer-(can/cannot)-smile
               '(can/cannot) cheer (someone) up and make (him) smile'
(9)   Genericized *zhu* (glossed as 'stay' merely to save space):
        a.   No standalone use with generic interpretation
        b.   yao-(de/bu)-zhu
               bite-(can/cannot)-stay
               '(can/cannot) bite and hold'

(10) *che* 'spread though':
    a.   No standalone use in modern Chinese
    b.   xiang-(*de/*bu)-che
          make.loud.sound-(*can/cannot)-spread.through
          '(can/cannot) spread loud sounds all over'

As (9) and (10) show, a morpheme incapable of being used by itself does not automatically reject infixation between V and R. By this standard, there are at least three values for R on the boundness (*b*) axis. The theta-less *si* patterns with *che* in this respect, with no standalone use and rejecting infixation (compare (4b) with **na-ci kaoshi nan-de/bu-si*. 'that-CL exam difficult-can/cannot-SI').

    When all word-internal morphemes are taken into consideration, at least one more measurement factor should be included: whether a given morpheme M has multiple linear options inside a word. The thematically depleted *si*, for instance, may serve as either the first or second morpheme of a word (see (1) and (2)), in contrast with the diminutive suffixal -*r*/-*er*. This factor participates in measuring boundness in Chinese because more linear freedom feels inversely correlated with being less bound. *Xiao* 'smile', *zhu* 'stay' and *che* 'spread through', exemplified in (8) through (10), all have the freedom (see *xiao-sha* 'laugh-dumb', *zhu-shou* 'hold-hand', *che-di* 'spead.through-bottom') whereas the diminutive -*r*/-*er* and the aspectual markers can only be suffixal.

    Since each of the three factors (standalone use (SU), infixation (INF), linear freedom (LF)) is binary-valued, the morphemes considered so far can be calculated quantitatively for their boundness. To be consistent with the *s*-axis on which departure from the base (or historical) form increases the value, the boundness values (*b*-values) are calculated by assigning 0 to the 'yes' answer and 1 to 'no'.

(11) Values on the *b*-axis

	SU	INF	LF	*b*-value
xiao	0	0	0	0/3=0
zhu	1	0	0	1/3=.33
che/*si*	1	1	0	2/3=.67
r-/-*le*	1	n/a	1	2/2=1

(12) The *b*-axis

Note that fully lexical morphemes by the semantic standard, such as *che* 'spread through', may be closer to being bound than free a direct result of the well-known transition from monosyllabic to polysyllabic words in Chinese over the millennia. Similarly, morphemes of the S3 group split on the

b-axis. This fact indicates that the *s*- and *b*-axes, though correlated for many morphemes, nonetheless describe separate paths of change and neither can be taken as a simple function of the other.

The phonological (*p*) axis measures the third dimension of shifting in word-internal morphemes. As mentioned previously, both V and R carry their own base tones by default, and R typically bears either an identical stress to V or a heavier one. But some instances of R deviate from this pattern. *Zhu* (=S2) is more commonly unstressed and toneless unless it is emphasized. Further down the axis, when *si* serves as R, it is almost always unstressed whether it keeps its base meaning of 'die', used exaggeratedly, or thematically depleted. Lastly, the aspectual suffixes are never stressed, nor is the *-r* form of the diminutive. Again, at least four stages of phonological 'bleaching' can be identified (with 0 being no bleaching at all and 1 being the total loss of stress and tone).

(13)　The *p*-axis

che, *xiao*	*zhu*/*kai*	*si*	Aspect affixes, *-r*
(S1)	(S2/S1)	(R1/S3)	(S4/S3)

$$\bullet \text{——————} \bullet \text{————} \bullet \text{————} \bullet \longrightarrow$$

0　　　　　　　　　　　　　　　　1
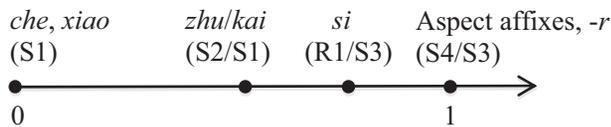
Admittedly, the positioning of the morphemes on the *p*-axis lacks a quantitative basis. But it is not entirely arbitrary. Since 0 stands for bearing either an equal stress or a heavier one, a morpheme like *zhu* with the options of an equal stress or no stress retains 'half' of the default pattern and thus is placed right between 0 and 1. Meanwhile, *si* is only very rarely stressed and intuitively should be very close to the necessarily unstressed aspect markers. Its position in (13) reflects this.

Also worth noting in (13) is *kai*, almost always translated as 'open' when used alone as a verb. But the translation fails to explain the semantics of the morpheme as R.

(14)　zou-kai 'walk-KAI = walk away'
　　　na-kai 'hold-KAI = take away'
　　　li-kai 'leave-KAI = depart'

As R in these *r*-words, *kai* generally adds the reading of *off* or *away* to the event of V. This interpretation seems so removed from the meaning of *open* that Starosta et al. (1998) argue that as R, *kai* is not a verb but a derivational suffix. I suggest that the meaning of the morpheme is to describe the event of two objects (e.g. a door and its frame) moving apart, which is also what *kai* means as R: let S be the subject of the verb and L be S's discourse-determined initial location, *S zou-kai* simply means S's walking resulting in S moving away from L. In other words, *kai* retains its base meaning even when used as R, and thus belongs to S1. On the *p*-axis, though, *kai* exhibits the same pattern as the genericized *zhu*, unstressed by default with the option of a full stress and the base tone. The point to be made is that morphemes in a group on one axis may belong to multiple groups on another axis, as clearly seen through (7), (12) and (13). This fact argues against a clean division of Chinese polymorphemic words into standard morphological classes.

1.3 Does Chinese have derivational words?

The two 'extremes' of morphological classification, compounding and inflection are easy to establish in Chinese. Words composed exclusively of free morphemes like *xiao* 'laugh, smile' and *nie* 'pinch' are clearly compounds (e.g. *zuo-deng* 'sit-await', *dou-xiao* 'cheer.up-smile'), and those containing aspectual suffixes (e.g. *xiao-le* 'laugh-ASP') are inflectional.[6] But it is unclear whether derivational morphology can be meaningfully justified in the language. We consider a few candidates for derivational affixes to illustrate the problem.

First consider *-hua*, which is perhaps the closest to the causative/inchoative affix in other languages (for more examples of *-hua*, see Lü 1984).

(15)   xiandai-hua 'modern-become → become modern, modernize'
        chengshi-hua 'city-become → become urbanized'
        you-hua 'excellent-ize → optimize'

Looked at alone, the examples in (15) can certainly be viewed as derivational words. But two more facts should not be overlooked. First, even in modern Chinese, *hua* can still be used to mean 'change, transform', with both the inchoative and causative uses. Second, when a verb and its internal argument form a compound in Chinese, the linear order is VO if V is transitive but OV if V is unaccusative.

(16)   V = transitive
        dong-yuan 'move-personnel → mobilize'
        de-zui 'receive-blame → offend'
        xing-mu 'wake.up-eye → attract attention'
(17)   V = unaccusative
        di-zhen 'earth-quake → earthquake'
        rou-ma 'flesh-numb → sicken'
        dan-qie 'gall.bladder-fear → be timid'

Given these two facts, the word order inside *xiandai-hua* 'modern-become' simply patterns with (17), with *hua* being essentially a raising verb and thus thematically an unaccusative with a single internal argument. What appears to be the causative use of *-hua* in *you-hua* 'optimize' may be viewed as resulting from further causativizing the entire word *you-hua* through a zero-morpheme (Pesetsky 1995) rather than using *-hua* as the causative morpheme in itself.

This analysis of *-hua* offers a possible explanation for the general tendency that X-*hua* is most naturally used as the causative transitive only when X is monomorphemic (see Lü 1984 for many examples that display this pattern). If *-hua* simply has the options of being inchoative and causative,

---

[6] The plural marker *-men* is often treated as an inflectional suffix (see Li 1999). But there is evidence (Li 2012) that it is most likely to be a clitic to a full DP.

this sensitivity to the syllable number of the root seems without reason. But if *-hua* is only inchoative and causativization happens to the entire X-*hua* word W, the strong preference for W to be disyllabic (which in turn forces X to be monosyllabic) may be attributed to a separate zero-form causative morpheme. It is more plausible that a separate morpheme may have its own requirements, including that it targets stems with only a certain number of syllables.

Returning to (17), to the extent to which these O-V words are unlikely to be derivational, it is no longer obvious that X-*hua* is a derivational suffix given the facts that *hua* is independently lexical, X is the single thematic argument of the inchoative *hua*, and X-*hua* has the same linear order as comparable compounds in (17).

Another tempting candidate for a derivational affix is *zhe* 'person, *-er*'.

(18)   ji-zhe 'report-person → reporter'
       xianxing-zhe 'first.move-person → pioneer'
       shengli-zhe 'victory-person → victor'

Indeed, *zhe* is never used as a standalone noun in colloquial Chinese, but only occurs as part of a word. However, a well-known fact proves that even such a bound morpheme should not be equated to *-er/-or* in English.

(19)   renhe   zi   ji-ri   qi   weifan   gai-xiang   xin   zhiding
       any   from   this-day   onward   violate   this-Cl   newly   formulate
       zhi   tiaoli zhe
       ZHI   rule   person
       'any person who violates this newly made rule from this day onward'

What precedes *zhe* is an entire relative clause. The choice of words generally has an archaic flavor, but the construction itself is highly productive in certain formal genres, suggesting that *zhe* is but a morphologically bound noun with a strongly formal/archaic connotation. And because there is no known reason to postulate different *zhe*'s for (18) and (19), it makes more sense to describe *zhe* as a clitic than a derivational suffix.

The last case to be looked at is the diminutive *-r*. It patterns with the functional suffixes along both the *b-* and *p-*axes, but shares with the theta-less *si* the same semantic property of having a single lexical feature (see the discussion leading to (7)). This semantic quality allows it to create a new word (or more precisely a lexeme) by affixing to an old one (*xiang* 'elephant' ≠ *xiangr* because *xiao xiangr* 'small elephant' is good but *da xiangr 'big elephant' is completely impossible), hence it seems to be a derivational morpheme. But the same can be said about *si* from S3. For sure, *si* may occur on either side of a lexeme (recall (1) and (2)) and displays less productivity than *-r*, but so does causativizing *en* in English (*en-large*, *en-able* vs. *bright-en*, *light-en*). If *si* is taken to be derivational, however, how about *che* in *che-ye* 'throughout-night = whole night', *che-gu* 'throughout-bone = bone-penetrating' versus *xiang-che* (=(10b)) and *tou-che* 'pass.through-throughout = thorough'? Whether by semantic content (or any of the other axes defined above) or

by productivity, there exists a spectrum of graduated change from *-r* to *si* to *che*. Arbitrariness is unavoidable no matter where one draws the dividing line for derivational morphemes.

## 1.4 A coordinate system for Chinese morphemes and words

I suggest that derivational words are not a universally meaningful morphological class, and that the more accurate framework to describe the various properties of Chinese words is the one in (20).

(20)   The morphological typology in Chinese:
    a.   The *s-axis* reflects a gradual shift from full lexical content (marked as 0) to unequivocal functional content (marked as 1).
    b.   The *p-axis* reflects a gradual shift from a full stress and a tone (marked as 0) to no stress and loss of tone (marked as 1).
    c.   The *b-axis* reflects a gradual shift from (+SU, +INF, +LF; marked as 0) to (–SU, –INF or no INF, –LF; marked as 1).
    d.   Each morphologically complex word is represented as a triplet of coordinates (s, p, b) in the space defined by the three axes above:
        i.   Let W be a word composed of morphemes $m_0 \ldots m_n$ and $0 < i \leq n$. The coordinates of W is the m_i with the largest $|s+p+b|$.
        ii.   A compound C consists of morphemes each of which has (0, 0, 0); the coordinates of C = the coordinates of any one of its components = (0, 0, 0).
        iii.   An inflectional word I consists of at least one morpheme of (1, 1, 1); hence the coordinates of I = (1, 1, 1).

(21)   Illustration of the compound *nie-jin* 'pinch-tight' and the inflectional word *nie-le* 'pinch-ASP'.
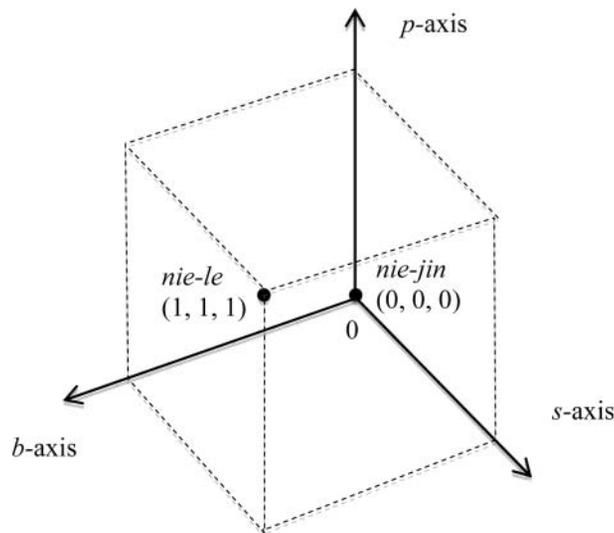
Compounds and inflectional words are represented at the two diagonal corners of this system. Other than these two extreme types defined in (20d-ii) and (20d-iii), all other 'in-between' words are distributed in the space marked out by the dotted cube. There is no derivational class that is descriptively or explanatorily insightful for Chinese.

What this system gets us is summarized below. First, it makes explicit the steps of diachronic change from lexical to functional morphemes in the case of R. A trajectory is clearly seen on the plane defined by the *s*- and *p*-axes (*jin* 'tight' = S1, *zhu* = S2, *si* = S3, *-le* = S4; see (7). The trajectory is equally reflected on the *b*-axis, which is not given here for lack of effective 3-dimential graphics).

(22)
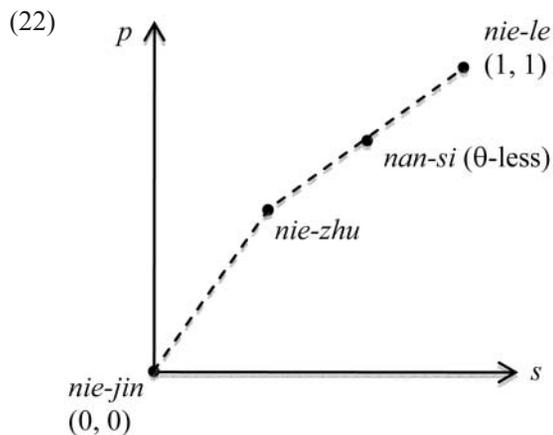
It is a unique fact about modern Chinese that a lexical origin underlies practically every functional morpheme as well as all those 'grammaticalized' words (such as *ba* and *bei*, the object and passive markers respectively) and as such still affects their syntactic and morphological behavior. Any grammatical theory of Chinese would be inadequate without the apparatus to capture this fact; it would also run the risk of omitting important information for explaining certain facts in the language. This brings us to the next point.

Second, recall from §1.1 that *si* as R is typically stressless, contrasting both with the default R and with *si* elsewhere inside a word (see (1) and (2) and the discussion thereafter). In itself, this is only an isolated idiosyncratic trait of *si*. But placed in (22), it becomes a meaningful link in the chain of changes that ultimately resulted in the aspect system of modern Chinese. In nature, an r-word explicitly encodes a complex event with V for the cause and R for the effect. From the perspective of event aspectuality, R constitutes the termination point of V (cf. Li 1993 and references therein). As it happens, the sole grammatical function of perfective aspect is to mark an event's termination as well. What happens in Chinese is that a lexical morpheme (=R) that denotes the intrinsic terminal point of an event is eventually turned into a functional suffix that expresses the speaker's focus on the end of an event (i.e. Smith's 1991 perfective view-point). (22) indicates that this lexical-to-functional shift has multiple stages, each of which is 'fossilized' in modern Chinese.

Minimally, one should now question whether a historical change indeed always corresponds to simply resetting a parameter, a view held by enough researchers in the Principles-and-Parameters (PnP) framework (see Kroch 2001 for a review).

Third, the coordinate system in (20) completely removes any arbitrariness from the discretely defined compounding-derivation-inflection partition. No coercion or subjective decision is needed to squeeze Chinese morphemes or words into these classes. Words composed exclusively of free morphemes can be conveniently called compounds and those formed with unequivocally functional affixes, inflectional words. But many exist in between to form continua by various linguistic standards. A few may indeed qualify as derivational by one standard or another, but being one of the many dots in the coordinate system, they have no special status in Chinese grammar and deserve to be given a class name as much or as little as any other in-between words that do not readily fall into any of the morphological classes defined primarily on the basis of European languages.

A desirable side-effect of this coordinate-approach is the uniform structural analysis of all r-words in Chinese, including verbs suffixed with aspect. Observing how little semantic content is preserved in the thematically depleted R-morphemes such as *si*, Starosta et al. proposed to treat them as derivational suffixes which serve as the morphological heads of words like *nan-si* 'difficult-SI'. This conforms with the typical morphological analysis in the literature, but creates an odd picture in Chinese: there is clear evidence that r-words consisting of free morphemes are head-initial (Li 1990, 1993) and aspectually suffixed verbs are still verbs (=head-initial), but those in the middle, though identical in form and in the basic semantics to stereotypical r-words, are taken to be head-final. The coordinate theory eliminates the problem from its root since there is no need to classify the variously bleached Rs into different classes. Whether (1, 1, 1) or (0, 0, 0) or anywhere in between, they all follow the same general *r*-word format. All *r*-words are head-initial, with compounding, inflection or whatever class name is used simply being there for convenience and of little theoretical significance.

Finally, the coordinate description of morpheme/word types highlights a fact: Languages vary to some extent in what factors are significant in their morphology. Due to historical reasons that are still active or visible now, the phonological properties of a morpheme may still deserve attention in Chinese (e.g. the stressless *si* as R). For English and many other European languages, however, stress is typically calculated by counting the number of syllables from one end of a word and generally independently of morpheme boundaries. Hence a theory of morphological types only needs to refer to semantic content and boundness in such languages. Furthermore, the notion of boundness also varies in content. In English, standalone use is the only factor that matters to describe morphological behavior, whereas in Chinese at least infixation must be taken into consideration in order to be descriptively adequate for certain types of words. Such crosslinguistic variations may cause complications and inconsistencies if one wants to define, say, the derivational class in a universal theory of morpheme typology. But the problem would never arise for the coordinate-based framework, in which a class name appropriate for one language may not even be meaningful for another.

2. Methodological thoughts

Section 1 presents a theoretical model that departs from the mainstream practice of the PnP approach in two respects: whether to recognize the existence of data-spectra by designing explicit theoretical apparatuses for them, and how to understand and achieve the goal of theorization to uncover 'deeper' laws.[7] This section offers thoughts on each of the two subject matters, with a view to justifying the legitimacy and necessity of the approach in §1 as part of any empirical inquiry, and to pointing out the conceptual and empirical flaws in the exclusive binary-valued deeper-is-better attitude toward linguistic research.

2.1 On squish-o-phobia

It is part of the theory in (20) to recognize the various intermediate states between the lexical base form and the functional affix. This approach to data is certainly not new. In a series of articles of which Ross (1972) is a representative, various spectra of data, called squishes, are investigated.

Ross's work is not received well by PnP for multiple reasons, one being technical in nature and involving empirical and theoretical details in Ross's presentations. For concreteness, I shall use the nouniness squish as an example. Ross's data continuum is given below.

(23) a. I wonder who he resented that I went steady with.
b. I wonder who he would resent for me to go steady with.
c. *I wonder who he resented how long I went steady with.
d. ?I wonder who he resented me going out with.
e. ??I wonder who he resented my going out with.
f. ?*I wonder who he resented my careless examining of.
g. ?*I wonder who he resented my careless examination of.
h. ?*I wonder who he resented the daughter of.

According to Ross, 'the nounier a complement is, the less accessible are the nodes it dominates to the nodes which command the complement' (Ross 1973:174). Rephrased in modern terminology, the nounier a complement is, the harder it is to extract out of it.

One objection to squishes takes the form of questioning the data itself. Newmeyer (1998), for example, denies the five levels of acceptability given by Ross and judges (23c,f,g) to be bad and the rest all good. It follows, for Newmeyer, that this squish does not exist in the first place. But this is not the way to refute a claimed phenomenon because, after all, Ross is a native speaker of the

---

[7] One may disagree with this remark by quoting Chomsky's (1986) *n*-Subjacency. Technical details aside, it is worth noting that the notion never received serious attention. Considering how enthusiastic the field generally is with Chomsky's ideas, the fact that this one was ignored shows how uneasy people are with spectra.

language as well. That one author does not find intermediate stages of acceptability does not auto-matically mean other speakers do not either. Below are the judgments from five graduate students, numbered from 1 to 5, with syntactic training but no stake in the debate.

(24)	example	1	2	3	4	5	Ross	Newmeyer
	a.	good	good	good	good	good	good	good
	b.	?	good	good	good	good	good	good
	c.	*	*	*	*	*	*	*
	d.	good	good	?	good	good	?	good
	e.	(?)	good	(?)	good	?	??	good
	f.	*	(?)	?	good	*	?*	*
	g.	*	(?)	?	good	*	?*	*
	h.	good	*	good	good	?	?*	good

All the informants share with Ross and Newmeyer that movement is good out of a non-*wh* com-plement clause but bad when a *wh*-island is crossed. The examples (d) through (h), however, are undeniably where native speaker assessment varies greatly. This divisiveness is especially worth attention given the consistency they demonstrate with examples (a–c). The ultimate explanation may or may not be from the core mechanisms of Universal Grammar (UG), and there may even be multiple factors at work here. But it is clear that some 'in-between' data does exist for what appears to be the majority of the native speakers of English. A denial from an individual speaker does not make the fact go away.

Another criticism of squishes is, as Newmeyer (1998:171) puts it, that 'the assignment of a rating between 0 and 1, struck many linguists as arbitrary and unmotivated. No reasonable set of criteria, for example, was ever proposed to determine if an abstract NP merited a rating of, say, .5 or .6 on the noun phrase scale'. For sure Ross did not provide such criteria, but this criticism of him is hardly fair, especially when it comes from the PnP tradition.

First of all, without systematically investigating the type of in-between data that Ross talked about, one would not know a priori whether there exist any linguistically significant rating criteria for squish measurement. When initial attempts to record such data are scorned or ignored, the field is pretty much left with the only logical route to ever reach a set of legitimate criteria, namely sitting still and waiting for overnight inspiration. Furthermore, it is common practice in the PnP literature to label marginal data with a full list of acceptability marks—(?), ?, ??, ?*, *, **, among others. In effect, the field of syntax has been assigning values to marginal data all along. Not converting them to 0.5 and 0.6 does not change the nature of the practice. What one ought to avoid is the situation described by a Chinese saying about the right to use fire: While the governor can set a house ablaze, an ordinary citizen is not even allowed to light a candle.

A small exercise with (24) is instructional. There are altogether four levels of acceptability: good, (?), ? and *, which can be converted numerically to 1, 2/3, 1/3, 0, respectively. The mean acceptability value (*m*-value) of each sentence in (23) is calculated by simply adding the converted

digital values from each speaker and dividing the sum by 5 (standard deviation is ignored in this illustrative simple exercise).

(25)   The *m*-values of examples (a–h), with the sample size = 5
   a.   1   e.   .67
    b.   .93   f.   .40
    c.   0   g.   .40
    d.   .93   h.   .73

Next compare (25) with Newmeyer's and Ross's judgments. Since examples (a) and (c) are clearly at the two extremes for everyone involved, they are placed at the two ends of the example axis. Furthermore, Ross gives five levels of acceptability (*, ?*, ??, ?, good), which we take to evenly divide the acceptability axis.

(26)   Graph 1
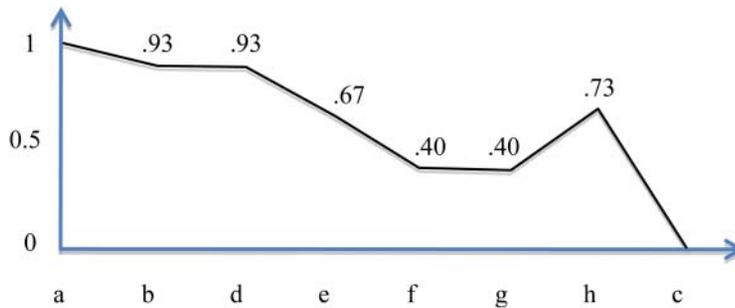
Graph 2
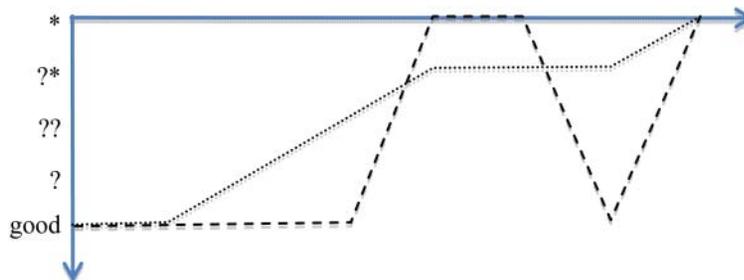
where Graph 1 = the five students, Graph 2 = Ross (⋯⋯⋯) and Newmeyer (- - -)

Except for example (h), Ross's judgments (⋯⋯⋯ in Graph 2) are very close to those of the five-people sample in Graph 1 in terms of the overall distribution of the degrees of acceptability. In contrast, Newmeyer's judgments, the - - - curve in Graph 2, exhibit much less correlation with Graph 1.

The graphs in (26) indicate that both at the individual level and in a collective sense, marginal acceptability is an undeniable reality when it comes to extractability. Whereas some speakers may only perceive a black-and-white picture (i.e. Newmeyer and Student 4), others—apparently more—clearly see different shades of gray. And in this respect, Ross's individual assessment of the data conforms better than Newmeyer's does to how the native speakers feel. Ultimately, it is the responsibility of linguistics to explain why examples (d) through (h) create such inconsistency and/or marginality among English speakers. For sure, the final explanation might not bear on the inner workings of UG, but logically it might. For instance, the very existence of such examples might be a clue to where exactly UG has 'functional gaps' so that other non-UG factors come into play, or it might end up proving that UG is not a completely binary-valued system as it is now formulated in PnP, that some core component of UG may be directly responsible for the nouniness squish. Individual researchers can certainly opt to not work on such a problem. For the discipline, though, both the fact itself and its theoretical implications—whatever they may be—deserve no less respect and attention than any other data. At the very least, efforts to record and analyze 'squishy' data should be encouraged and not dismissed.

This section obviously does not set out to defend all the squishes from Ross. What is important is for the field of syntax not to forget its empirical nature. It is a fact that intermediate data exist, both in the form of varying degrees of acceptability (see (23)) and through well-formed examples forming a spectrum by a certain linguistically relevant standard (see Chinese morphemes in §1). They are there to motivate further research, and deserve both a matter-of-fact recognition and a constructive attitude. Also worth remembering are similar examples from other empirical fields of inquiry. Take, for example, the famous example of Darwin's South American finches. The beaks of these birds vary in size and shape on different islands. This fact is not only a 'squish' but also one that helped discover the principles of biological evolution. How do we know, a priori, that some marginal or in-between data in syntax do not resemble the beak-sized squish and manifest a fundamental component of UG?

2.2 Is only abstract stuff good stuff?

To many scholars in the PnP model, the effort to describe minute factual details should be shunned for lacking the type of deep insight found only with abstracting away from superficial facts. Such a methodological choice is aptly expressed by Chomsky in a recent podcast interview by Jay Keyser:[8]

> … if you want to understand phenomena there's only two ways, one is brute force which would be some weak correlations, the rest just abstract away from them and study ideal cases. And that's called the Galilean style, attribute more reality to the ideal case even in

---

[8] http://www.mitpressjournals.org/userimages/ContentEditor/1256764313049/Keyser_Chomsky_ Podcast_ Transcript.pdf

a nonexistent system like a frictionless plane, than we do to the phenomena that are just too complicated because there's too many variables and too many interacting sort of things. And that's quite important for the study of language. In fact it's a live battle in the study of language. Should we just study the massive phenomena and use brute force methods to try to model them somehow. Or should we abstract away from the complex phenomenon, and try to find ideal cases that we'll be able to investigate and maybe find some principles.

Before this text, the 'brute force' approach is also likened to wanting to 'predict what's going on outside the window. . . . Take thousands of video tapes of whatever's happening outside the window, do some statistical analysis, and probably get a pretty fair prediction of what's going to happen next.' In brief, Chomsky advocates the Galilean route to search for principles via data idealization. Ultimately, this is also the philosophical justification for ignoring intermediate data.

There is no doubt in the mind of any theoretical linguist about the validity of the Galilean approach in syntactic research. In practice, however, there exist pitfalls in its implementation which the PnP model needs to watch out for.

The first one lies in what syntax views itself as an analogue of. In the previous quotation, Chomsky uses the nonexistent frictionless plane in the theorization of motion to justify data idealization in the study of language. If the velocity and distance of an object in motion are always affected by friction in reality but are not factored into what eventually led to the laws of Newtonian mechanics, in-between data can also be set aside, just like the real-life effect of friction, in formu-lating a theory of UG. The logic is straightforward; what is not, is whether syntax must completely resemble physics. Why, for instance, isn't biology at least a partial model instead?

Such a question is not entirely groundless. To begin with, the theory of UG is by hypothesis to investigate the nature and behavioral patterns of the biologically based 'language organ' (Chomsky 1965). At the conceptual level, then, one should not be surprised if some similarities are found among the behaviors of various biological systems. Take, for example, marginal acceptability in syntax. Whether it has any direct correspondence in Newtonian mechanics or not, it seems to find an analogue in interspecies crosses in the biological world, in the sense that due to their infertility, crosses are not fully 'well-formed' when compared with default species. Similarly, the aforemen-tioned beak shapes/sizes of South American finches correspond, at least at the descriptive level, to the Chinese morphemes ranging from being lexical, free or phonologically 'heavy' to functional, bound or phonologically bleached.

This discussion is not meant to turn analogies into evidence. Rather, it is simply not clear that a comparison between syntax and a single mature field such as physics can provide an exclusive model for syntactic research. It is also worth pondering this remark by Ernst Mayr, a principal architect of modern evolutionary biology and a philosopher of biological science (Mayr 1997:36): 'Perhaps the greatest failing of the philosophy of science … was that it took physics as the exemplar of science.' With respect to syntax, while various factors may force us to put aside some data at a given moment of research, arbitrarily deciding to idealize certain types of facts out of theorization,

especially by justifying it analogically through a single example of another field, runs the risk of throwing away clues to a hidden principle.

A methodological bifurcation such as Chomsky's has another danger: it may generate the impression—and actually has judging from the discouraging attitude of mainstream syntax toward squishes—that there is a zero-sum battle between two opposing approaches. In reality, the Galilean approach and the so-called 'brute force' approach complement each other for reasons that any non-partisan mind can see. Darwin's finches are an educational example where the detailed recording of the spectrum of beak sizes and shapes served as direct evidence for the theory of evolution. As a matter of fact, Darwin's *Origin of Species* is in essence the upshot of, and an abstraction from, a century's systematic data recording and classification, aka taxonomy, that was initiated and inspired by Linnaeus (see Mayr 1997:124). The modern version of taxonomy, called systematics, is defined by Simpson (1961) as 'the scientific study of the kinds and diversity of organisms, and of any and all relations among them'. Taxonomy is no longer the primary concern of modern biology, but it has provided the solid empirical foundation for the latter. Aware from the previous paragraph that no simple analogy may be adequate for determining the research model of syntax, it is nonetheless worth noting that when it comes to the degree of complexity and diversity of data, as well their discreet nature, modern syntax does seem more comparable to the early days of biology than to those of physics. To say the least, then, there exists no fundamental reason for which one can be sure that syntactic taxonomy aided by the empirically motivated abstractions (such as c-command and the locality constraints) will NOT help both to discover new facts and to eventually lead to a deeper understanding of UG.

The third consideration with overstressing the Galilean approach in the field of syntax is that in the absence of constant and conscious guard, doing so may let researchers close their eyes to apparently relevant facts in the name of abstraction and idealization. The short history of the PnP model has been full of such incidents.

The government-and-binding theory (GB) made the conscious effort to construct a modular system of language built on the unifying notion of government, a perfect example of seeking physics-like principles through abstraction (and presumably guided by Occam's Razor). Li (1987) noted that case theory, one of the core modules defined with government, both failed to account for certain facts (such as the distribution of the infinitive's subject) and suffered from theory-internal contradictions (how it was supposed to interact with θ-assignment). The proposed fix was to refor-mulate case-driven movement as the destination of movement (e.g. INFL) carrying a feature that must be matched and removed by a certain type of syntactic argument. The removal of this feature through matching allows both INFL and the argument to be interpreted properly, and government plays absolutely no role in this process. Obviously inconsistent with the highly abstract government-centric design of the GB model, this solution was vehemently opposed and dismissed by Chomsky as having 'methodological problems' (Noam Chomsky, personal communication). It should be instructional to the field of syntax, then, that a few years later, the essence of Li's solution was generalized, by Chomsky himself, into two core characteristics of the Minimalist Program: that movement is driven by feature matching and removal, and that government is no longer at the core

of syntactic computation.[9] What we learn from this incident is that without proper constraining, placing abstraction and idealization over the basic respect for facts and straightforward logic can lead research vastly awry, while doing the opposite may prove to be more insightful in the long run.

This may be an extreme case, with the whole field spending over a decade on a notion (i.e. government) proposed only to pursue some subjective implementation of Galilean abstraction. But it is by no means an isolated case. As the relation between data and theory is the central concern of this paper, I shall consider another example that has been having a wide impact on the field.

Baker's (1988) theory of incorporation has played a critical role in popularizing head-movement within the PnP framework of linguistic analyses. The subject–object asymmetry found in noun-incorporation (NI) is one of the two strongest arguments[10] for deriving a morphological phenomenon through syntactic head-movement.

(27)

```
              S
           /     \
        NP1       VP
         |       /    \
        N1      V      NP2
                        |
                        N2
```
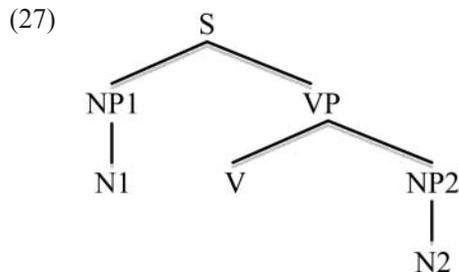
Moving N2 to V is good because V c-commands N2. Since V does not c-command N1, merging the two is impossible. Hence only the object noun can incorporate to the verb.

The analysis is simple, elegant and it abstracts away from the superficial syntax–morphology distinction. But from the very beginning, it faces an empirical problem: English is not an NI language because it is no good to say *These students room-cleaned yesterday*, but compounds of the form N-V-*ing* exhibits the same subject–object asymmetry: *room-cleaning* is fine but *student-cleaning* does not have *student* understood as Agent. Given the theoretical elegance of Baker's analysis, however, this obviously relevant fact has been being ignored by the entire field.

Baker's theory of NI was used in Hale & Keyser (1993) to account for English denominal verbs. One of their questions is why the following contrast exists, where (28b) is intended to mean the same as (28a) with *it* being the expletive.

(28)   a.   A cow calved.
       b.   *It cowed a calf.

---

[9] No credit is given by Chomsky to Li's pioneering work.
[10] The other is anaphor binding in verb-incorporation, which faces its own problems (see Li 2005).

Hale and Keyser's explanation is that denominal verbs are derived from an X'-style Lexical Relational Structure (LRS). Details aside, both examples in (28) have the LRS $[[_{NP1}$ cow$]$ $[V$ $[_{NP2}$ calf$]]]$, with V roughly meaning 'bring about'. The subject–object asymmetry with NI applies directly, letting only *calf* merge with V to yield *calve*.

An unexpected consequence of Hale and Keyser is that the very LRS analysis leads to a lexicalist tool to explain the subject–object asymmetry and therefore undermines Baker's syntactic approach to morphology. Because LRS allows recursion, it is necessary to prevent causativization of an unergative verb such as *sneeze*: there is no sentence like **The alfalfa sneezed the colt* that means the alfalfa made the colt sneeze. With irrelevant details omitted, Hale and Keyser's solution is for the 'true external argument', namely the Agent/Causer, to be provided by a functional head outside the VP structure. The same idea is independently articulated in Marantz (1984) and Kratzer (1996) (to handle V-O idioms) and Chomsky (1995) (by postulating v to accommodate Burzio's Generalization, among other things). But if the agent role is provided by v, then lexically merging a lexical V with N will never give N the Agent reading because by definition, there is no Agent role around at the word-formation level. The subject–object asymmetry follows automatically. This outcome applies not only to English compounds such as *room-cleaning* versus *student-cleaning (with the intended agent interpretation), but also to all cases of NI provided that NI is a lexical process.

In sum, taking NI as evidence for word-formation in syntax has been flawed since the beginning: (a) English compounding appears to be a straightforward problem; (b) there exists an easy lexicalist explanation for the subject–object asymmetry through v, which is motivated for multiple reasons (Chomsky 1995; Hale & Keyser 1993; Kratzer 1996; Marantz 1984). Adding to all this is (c) Li's (2005) detailed proof that Baker's (1996) new evidence for deriving NI syntactically all has a straightforward lexicalist explanation and in some cases a simpler one (not to mention the semitic data therein which argue against forming words in syntax). But despite all the empirical and theoretical considerations going against it, merging heads via syntactic movement has been widely used in the PnP literature without the slightest sign of caution. This is a live example of theoretical abstraction and idealization exempted from empirical burdens, all in the name of pursuing the Galilean approach.

2.3 The Fact First Principle (FFP)

The emphasis of a field changes over time, both in terms of the subject matter and in terms of the primary approach. Mayr (1997:126) laments, for instance, taxonomic research's 'neglect and low prestige in recent years'. The fact remains, though, that taxonomy was a major part during the young age of biology and provided enough empirical nutrients for the field to grow future branches and theoretical fruits. There is no clear reason to believe that linguistics has grown out of the need for a healthy attention to its own taxonomy. Armed with concepts such as c-command, barriers and minimality, linguistic systematics may promise to lay the foundation for a more robust implementation of the Galilean approach.

From another perspective, the methodological 'live battle in the study of language', as Chomsky put it, reflects how little control we have over the subject matter. There are simply too many factors to be considered, discovered or put aside in every step of our research and too many possibilities for whatever needs to be analyzed. When uncertain which choice to pick, people argue over the ways to reach decisions. This is also why the field of syntax seems to draw on Occam's Razor constantly. But Occam's Razor has had a mixed track record in linguistics—government being a prime example—because its successful application crucially relies on 'everything else being equal'. And this condition happens to be the hardest to meet in linguistic research given the complexity of the subject matter.

For this reason, I propose to apply Occam's Razor to linguistic theorization only when at least most things can be guaranteed to be equal (which is already very hard to guarantee) and to use the Fact First Principle (FFP), defined below, as the guiding light in all other scenarios.

(29)   The Fact First Principle
       Let F be a relevant set of directly observable linguistic facts and C be a theoretical claim regarding F. If C is not straightforwardly supported by F, compelling evidence must be provided that there is no C′ satisfying these conditions:
       a.   C′ is more directly supported by F, and
       b.   C′ is adequate for F.

In itself, the FFP does not rule out, say, Baker's syntactic account (SA) of NI. But because NI manifests itself as words and not phrases and because compounds in English, a typical non-NI language, displays the same subject–object asymmetry, the adoption of SA should only be after giving robust proof that no lexicalist account (in the same general framework, for practical purposes) can accomplish the same task. And in the case that the latter option is not yet technically available (v came several years after Baker 1988), the field ought to be always ready to re-evaluate SA with the newly developed or uncovered technical tools.

Because the two existing options (SA vs the lexicalist account) have the same data coverage, it is considered a safe context for Occam's Razor to remove redundancies. In particular, since lexical word-formation is enough to derive all NI-related facts (see §2.2 for references) as well as compounds in English, syntax's ability to derive the same facts becomes totally redundant. Interestingly, not only is the redundancy easy to remove, but the removal has theoretically desirable consequences. Let head-movement always target an empty position (=substitution) just as phrases do (as is made explicit through merge to the edge of derivation in the Minimalist Program). Such a theory of movement is maximally uniform and devoid of the conceptually messy and technically ad hoc adjunction to heads. The theory is also derivable from exclusive lexical word-formation—if all morphologically complex words are lexically constructed, there would be no occasion where morphemes need to be merged through head-movement. See Li (2005) for details. What is especially worth noting is that the option favored by the FFP is also the one that is the simplest and the maximally consistent with the rest of the movement theory.

If the FFP had been in place, the field would not have spent over a decade on the notion of government only to watch it disappear completely afterwards. As mentioned above, it had obvious flaws from the beginning and was 'justified' primarily in the name of Galilean abstraction and idealization. A little more attention to the facts it was designed to handle, as Li (1987) did, would have pointed the discipline at what has proven to be a more correct direction (no government and no clumsy m-command in UG, feature-checking as motivation for movement), not to mention the resources to be saved.

Back to morpheme classification in Chinese, it is a known fact that morphemes inside a word exhibit varying degrees of shifting from their base or historical forms and that these shifts affect their semantic, morphological and phonological behavior. The FFP favors any theoretical effort to represent this fact accurately over trying to idealize some pieces of it into the discrete classes of the popular theory of morphology. More generally, the FFP demands due attention to various kinds of factual squishes and does not exclude the possibility, a priori, that intermediate data may also reveal the inner workings of UG. It must be noted that the FFP does NOT object to data-idealization and theoretical abstraction. It merely wants a researcher to situate his/her Galilean effort constantly on the factual ground, so that the research on UG proceeds as the route to discover the truth and not as a competition for fancy story-telling. The ultimate basis for the FFP is the assumption that linguistics is on its way to becoming a science and not something like mathematics. See Li (2014) for more applications and implementations of the FFP.

References

Baker, Mark C. 1988. *Incorporation: A Theory of Grammatical Function Changing*. Chicago: University of Chicago Press.

Baker, Mark C. 1996. *The Polysynthesis Parameter*. Oxford & New York: Oxford University Press.

Cheng, Lisa L.-S., & C.-T. James Huang. 1994. On the argument structure of resultative compounds. *In Honor of William S-Y. Wang: Interdisciplinary Studies on Language and Language Change*, ed. by Matthew Y. Chen & Ovid J. L. Tzeng, 187–221. Taipei: Pyramid Press.

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

Chomsky, Noam. 1986. *Barriers*. Cambridge: MIT Press.

Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge: MIT Press.

Gu, Yang. 1992. *The Syntax and Resultative and Causative Compounds in Chinese*. Ithaca: Cornell University dissertation.

Hale, Kenneth, & Samuel Jay Keyser. 1993. On argument structure and the lexical expression of syntactic relations. *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, ed. by Kenneth Hale & Samuel Jay Keyser, 53–109. Cambridge: MIT Press.

Her, One-Soon. 2007. Argument-function mismatches in Mandarin resultatives: a lexical mapping account. *Lingua* 117.1:221–246.

Huang, C.-T. James, Y.-H. Audrey Li, & Yafei Li. 2009. *The Syntax of Chinese*. Cambridge & New York: Cambridge University Press.

Jackendoff, Ray. 2002. *Foundations of Language: Brain*, *Meaning*, *Grammar*, *Evolution*. Oxford & New York: Oxford University Press.

Kratzer, Angelika. 1996. Severing the external argument from its verb. *Phrase Structure and the Lexicon*, ed. by Johan Rooryck & Laurie Zaring, 109–137. Dordrecht & Boston: Kluwer.

Kroch, Anthony S. 2001. Syntactic change. *The Handbook of Contemporary Syntactic Theory*, ed. by Mark Baltin & Chris Collins, 699–729. Malden: Blackwell.

Li, Y.-H. Audrey. 1999. Plurality in a classifier language. *Journal of East Asian Linguistics* 8.1: 75–99.

Li, Yafei. 1987. Case in categorical terms. General exam paper. Cambridge: MIT.

Li, Yafei. 1990. On V-V compounds in Chinese. *Natural Language and Linguistic Theory* 8.2: 177–207.

Li, Yafei. 1993. Structural head and aspectuality. *Language* 69.3:480–504.

Li, Yafei. 1995. The thematic hierarchy and causativity. *Natural Language and Linguistic Theory* 13.2:255–282.

Li, Yafei. 2005. X^0: *A Theory of the Morphology–Syntax Interface*. Cambridge: MIT Press.

Li, Yafei. 2012. On the internal structure of the Chinese NP. Manuscript. Madison: University of Wisconsin-Madison. (Submitted to *Zhongguo Yuwen*)

Li, Yafei. 2014. On optimization in syntactic theorization. *Linguistics*, ed. by Zhijie Wang & Dong-dong Chen, Chapter 9. Beijing: China Renmin University Press, forthcoming.

Lin, Fu-wen. 1990. *The Verb-Complement* (*V-R*) *Compounds in Mandarin*. Hsinchu: National Tsing Hua University MA thesis.

Lü, Shuxiang. 1984. *Xiandai Hanyu Babai Ci* [*Eight Hundred Words in Modern Chinese*]. Beijing: The Commercial Press.

Marantz, Alec. 1984. *On the Nature of Grammatical Relations*. Cambridge: MIT Press.

Mayr, Ernst. 1997. *This is Biology: The Science of the Living World*. Cambridge: Belknap Press of Harvard University Press.

Mei, Tsu-Lin. 1994. Tangdai, Songdai gongtongyu de yufa he xiandai fangyan de yufa [The grammars of T'ang-Sung *koine* and the grammar of Modern Chinese dialect]. *Chinese Languages and Linguistics*, Vol. 2: *Historical Linguistics*, ed. by Paul Jen-kuei Li, Chu-Ren Huang & Chih-Chen Jane Tang, 61–97. Taipei: Institute of History and Philology, Academia Sinica.

Newmeyer, Frederick J. 1998. *Language Form and Language Function*. Cambridge: MIT Press.

Pesetsky, David M. 1995. *Zero Syntax: Experiencers and Cascades*. Cambridge: MIT Press.

Ross, John R. 1972. The category squish: Endstation Hauptwort. *Chicago Linguistic Society* (*CLS*) 8:316–338.

Ross, John R. 1973. Nouniness. *Three Dimensions of Linguistic Theory*, ed. by Osamu Fujimura, 137–258. Tokyo: TEC Company.

Simpson, George G. 1961. *Principles of Animal Taxonomy*. New York: Columbia University Press.

Smith, Carlota S. 1991. *The Parameters of Aspect*. Dordrecht & Boston: Kluwer.

Spencer, Andrew. 2001. Morphology. *The Handbook of Linguistics*, ed. by Mark Aronoff & Janie Rees-Miller, 213–237. Malden: Blackwell.

Starosta, Stanley, Koenraad Kuiper, Siew-ai Ng, & Zhi-qian Wu. 1998. On defining the Chinese compound word: headedness in Chinese compounding and Chinese VR compounds. *New Approaches to Chinese Word Formation: Morphology*, *Phonology and the Lexicon in Modern and Ancient Chinese*, ed. by Jerome L. Packard, 347–370. Berlin & New York: Mouton de Gruyter.

Zou, Ke. 1995. *The Syntax of the Chinese Ba-constructions and Verb Compounds: A Morphosyntactic Analysis*. Los Angeles: University of Southern California dissertation.

Department of Linguistics
University of Wisconsin-Madison
Madison, WI 53706
USA
yafeili@wisc.edu

漢語詞素與坐標系：
兼評語言學理論的科學性

李亞非

南京大學
威斯康辛大學麥迪遜分校

對漢語動補詞的研究表明，派生詞並非是一個對所有的語言都具意義的理論概念。在漢語中，通過座標系來測量和定位一個詞素的語義、語音及自由度漂移，可以更加精確地處理各類構詞現象和反映其歷史緣由。在此基礎上，本文第二部分對當代句法理論的研究方法提出一些質疑。

關鍵詞：動補詞，派生詞，座標系，理想化