

A+N constructions in Mandarin and the 'compound vs. phrase' debate¹

Martin Schäfer

Abstract

Mandarin Chinese has three A+N constructions with distinct formal properties. One construction is clearly phrasal, one clearly constitutes a compound. The status of the third construction is controversial, being analysed either as a compound or as a phrase. Frequently drawing on data from Germanic A+N constructions for comparison, I show in this article that this issue is undecidable on the basis of the Mandarin data. On the other hand, I argue that the third construction cannot be collapsed with either of the other two constructions, regardless of whether it is analysed as a compound or a phrase.

I Introduction

Mandarin Chinese has three A+N constructions with distinct formal properties, shown in the examples in (1).

- (1) (a) *hēi-ban* [Type I]
black-board
'blackboard'
- (b) *hēi de gǒu* [Type II]
black SUB dog
'black dog'
- (c) *dà pánzi* [Type III]
big plate
'big plate'

In the literature, there is no controversy about the first two of these constructions: (a) is analysed as a compound, and (1b), where the subordinating particle *de* 'SUB' appears between the A and the N, as a phrasal construction (see for example standard grammars such as Li & Thompson (1981)). The controversy revolves around the type

III construction. On the surface, it resembles the type I construction, but types I and III differ with regard to their formal properties (cf. Feng 2001; Paul 2005). Initial evidence for a structural difference between the two A+N constructions without *de* SUB comes from data on adjective ordering. Crosslinguistically, the ordering of multiple prenominal modifiers obeys the order SIZE > COLOUR (see Sproat & Shih 1991), cf. the German (and English) data in (2).

- (2) (a) *der große weiße Mercedes* [default]
 the big white Mercedes
 (b) *#der weiße große Mercedes* [prosodically marked]
 the white big Mercedes

These ordering constraints have to be obeyed for the A+N type III constructions, cf. (3), where the colour adjective *bái* ‘white’ and the size adjective *dà* ‘big’ are combined with the noun *pánzi* ‘plate’.

- (3) (a) *dà pánzi*
 big Plate
 ‘big plate’
 (b) *dà bái pánzi*
 big white plate
 ‘a big white plate’
 (c) **bái dà pánzi*
 white big plate
 Cf. Feng (2001)

In contrast, the ordering constraints are violated in A+N type I constructions, cf. (4), where the same adjectives co-occur with the noun *guàr* ‘gown’.

- (4) (a) *dà-guàr*
 big gown
 ‘unlined long gown’
 (b) **dà bái guàr*
 big white gown
 (c) *bái dà guàr*
 white big gown
 ‘a white unlined long gown’
 Cf. Feng (2001)

Paul (2005) argues, based on this and other data, that the type III constructions are phrasal constructions. Duanmu (1998) argues that all *de*-less A+N constructions are words, making type III constructions compounds. A third position is held by Feng (2001), who, similar to Duanmu (1998), also regards both constructions, type I and type III, as words, but distinguishes between morphological words (‘*cífǎcí*’) and syntactical words (‘*jùfǎcí*’).²

In this paper, I will review the arguments put forward by the respective authors, showing that the Mandarin-internal view does not decide this issue. Nevertheless,

I argue that the three-way-distinction for A+N constructions must be upheld, regardless of the final analysis given to the A+N type III constructions. Throughout, I will draw on data from German for comparison and contrast.

The article is organized as follows: In the remainder of the introduction, I will review the standard criteria used to distinguish A+N compounds and A+N phrases in languages such as German and English and show that these criteria cannot be applied to Mandarin. In the second section, the semantics of the A+N type III construction are discussed. The third section discusses the status of productivity data with regard to the compound versus phrase distinction. In the fourth section, I will look at syntax-based arguments. Both sections are split up into smaller sections, where the relevant data is presented, followed by the interpretation of this data in the pro- and anti-compound literature. The fifth section discusses the consequences and implications arising from the data and its different interpretations. The article ends with a short conclusion.

1.1 Distinguishing between A+N phrases and A+N compounds

From the vantage point of a language like German, the distinction between compound and phrasal A+N structures is trivial, since two different criteria, morphological and prosodic, converge to distinguish the two.

Thus, the phrasal *schwarze Wurzel* ‘black root’ contrasts with the compound *Schwarzwurzel* ‘salsify’ because inflection for case and number is expressed on both the adjective and the noun of the phrasal A+N construction. In A+N compounds, on the other hand, inflection is only marked on the head of the compound. This difference between the two types of constructions is exemplified in (5), which shows the two constructions in the nominative and genitive singular. (5a) gives the data for the phrasal A+N construction, (5b) for the A+N compounds. In the A+N compound, *schwarz* ‘black’ appears in its short form without any inflectional markers. In the phrasal A+N construction, the suffix *-e* expresses nominal case and the suffix *-en* expresses genitive case.

- | | | | | | | |
|-----|-----|-------|------------|-----------------------|---------------|--|
| (5) | (a) | [NOM] | <i>die</i> | <i>schwarz-e</i> | <i>Wurzel</i> | |
| | | | the | black-NOM.SG | root | |
| | | [GEN] | <i>der</i> | <i>schwarz-en</i> | <i>Wurzel</i> | |
| | | | the | black-GEN.SG | root | |
| | (b) | [NOM] | <i>die</i> | <i>Schwarz-wurzel</i> | | [GEN] <i>der</i> <i>Schwarz-wurzel</i> |
| | | | the | black-root | ‘salsify’ | the black-root ‘salsify’ |

The prosodic difference is witnessed by the divergent stress patterns, cf. *schwarze Würzel* versus *Schwarzwurzel*. Note that the different status of the two constructions is reflected in the orthography: the A+N compounds are written as one word, the A+N phrases are written as two words.

In English, although closely related to German, only the prosodic criterion remains, since A+N compounds and A+N phrases are not morphologically distinct. Thus, the A and the N in phrasal A+Ns are both stressed, cf. (6a), whereas in the A+N compound, only the A is stressed, cf. (6b).

- (6) (a) *bláck bírd*
 (b) *bláckbird*

According to Huddleston & Pullum (2002), from whom this data is taken, exceptions to the two prosodic patterns in (6) are very rare, but do exist: they mention *blackcurrant* and *hotdóg*.³

Mandarin Chinese is similar to English in that it does not inflect for case. In addition, even the prosodic criterion is absent, since there are no phonological effects that can distinguish compounds from phrases.⁴

This leaves semantics, productivity data, and syntax to establish the status of the type III construction.

2 Semantics

The three A+N constructions differ with regard to their semantics and their productivity. The construction with *de* is the most general of the three constructions. It is semantically always transparent, and there are few restrictions with regard to the type of adjective that can occur in this construction.⁵ The A+N type I constructions tend to be semantically non-transparent and are not productive. The A+N type III construction is somewhere in between: it is semantically transparent, but its productivity is severely limited.

A+N (and also N+N) constructions without *de* are traditionally described as functioning as names:

In general, adjectives that modify a noun without the particle *de* tend to be more closely knit with the noun. The consequence is that the adjective-plus-noun phrase tends to acquire the feature of being a name for a category of entities. (Li & Thompson 1981: 123).

That is, A+N constructions are semantically similar to common nouns. Paul (2005) discusses two important consequences of this semantic characteristic of the A+N type III construction. First of all, A+N type III can be used in some contexts where type II construction are not allowed, a finding Paul attributes to Fu (1987). One example for such a context is given in (7), classified as an identification context by Paul.

- (7) *Zhè shì heī (*de) tóujīn*
 This be black SUB scarf
 ‘This is a black scarf.’
 Cf. (33) in Paul (2005)

That is, *de* SUB is disallowed here when (7) is used very much like a definition or a classification, in other contexts, especially those involving contrast, the sequence *heī de tóujīn* ‘black scarf’ is allowed. Second, ‘[...] the “A/N+N” sequence has to result in a natural, plausible classification’ Paul (2005: 772). *Heī tóujīn* ‘black scarf’ in (7) is possible as an A+N type III construction, because it is natural and plausible to subclassify scarfs by their colour.

The most notable semantic difference between the type I and the type III construction is that type I constructions can be semantically non-transparent, cf. the data in (8) and (9).

- (8) (a) *huáng-yóu*
yellow-oil
'butter'
(b) *dà-yān*
great-smoke
'opium'

The examples in (8) show that the meaning is not transparent as it is not built up in a compositional way. (9) shows one consequence of this, namely non-contradictory interpretations if, of two incompatible adjectives, one appears as part of a type I compound, and the other as part of the type II phrase.

- (9) *bái de hēi-bǎn*
white SUB black-board
'white blackboard'
Cf. (12) in Duanmu (1998)

Paul (2005: 785, note 14) argues that type I constructions do not necessarily have non-transparent meanings but can also have 'more or less' compositional meanings, cf. (10).

- (10) (a) *xiǎo-shì*
little-matter
'minor matter, petty thing'
(b) *hóng-bāo*
red-envelope
'red envelope containing money offered as a gift, a gift of money, a bribe'
Cf. Paul (2005, note 14)

While (10a) is still transparent today, the situation with regard to *hóngbāo* is slightly more complex. The term was coined due to the practice of handing over the money-gift in red wrapping, but already in this usage its meaning is very specific. Although even today *hóngbāos* often come in the form of a red envelope, they do not have to come in this form, and in learning the meaning of the term, the etymology is irrelevant.⁶

Nevertheless, since there are cases like *xiǎo-shì* 'minor matter', the semantic transparency of an A+N construction is best viewed as only a one-way criterion: If an A+N construction is non-transparent, then it is not phrasal.

The fact that the A+N type III construction is semantically transparent and has a naming function does not point towards either a phrasal or compound analysis. In fact, these kinds of functions do not even have to be restricted to either phrases or words within one language, as a look at so-called lexicalized A+N phrases in German shows, cf. (11).

- (11) *wilde* *Ehe*
wild-e *Ehe*
 wild-NOM.SG marriage
 ‘cohabitation’
 = not officially sanctioned marriage-like way of living together

By the criteria discussed in section 1.1, *i.e.* morphology and prosody, *wilde Ehe* ‘cohabitation’ is clearly a phrase: the adjective carries inflection, and the accent falls on the second word. But its semantics is non-transparent and its meaning therefore needs to be listed in the lexicon.

3 Productivity

Type III constructions are productive, but their productivity is limited in a number of ways. In the following, we will distinguish two cases: a) adjectives that can never be used to form an A+N construction and b) adjectives that can only form A+N constructions in combination with a restricted number of nouns.

First, certain types of adjectives, the complex form adjectives, are not allowed in A+N constructions type III.⁷ Complex form adjectives can be either (a) reduplicated adjectives or (b) modifier-head compound adjectives. Neither of these are allowed in *de*-less modification structures, cf. (12) and (13).

- (12) (a) *gānjìng* (*de*) *yīfu*
 clean SUB clothes
 ‘clean clothes’
 (b) *gāngānjìngjìng* *(*de*) *yīfu*
 clean SUB clothes
 ‘(thoroughly) clean clothes’
 = (1) in Paul (2006)
- (13) (a) *bái* (*de*) *zhǐ*
 white SUB paper
 ‘white paper’
 (b) *bái bái / xuěbái / xuěbáixuěbái* *(*de*) *zhǐ*
 white/snow-white/sn.-wh-sn.-wh. SUB paper
 ‘(snow-) white paper’
 = (2) in Paul (2006)

As (12a) and (13a) show, the combination of *yīfu* ‘clothes’ and *zhǐ* ‘paper’ with adjectives of cleanliness and colour, respectively, is possible. It is the reduplication or the existence of an additional modifier of the A that leads to ungrammaticality: *gāngānjìngjìng* is the reduplicated form of *gānjìng* ‘clean’ and *bái bái / xuěbái / xuěbáixuěbái* are the reduplicated variant of *bái* ‘white’, a modifier-head compound adjective with head *bái* ‘white’, and the reduplication of that modifier-head compound, respectively.

Second, apart from the morphological restrictions mentioned in the previous section, there are many examples of restrictions on particular A+N combinations, cf. the following examples.

- (14) (a) *huáji diànyǐng*
 funny movie
 ‘funny movie’
 (b) **huáji rén*
 funny person
 (c) *huáji de rén*
 funny SUB person
 ‘funny person’
 Cf. (45d/46d/47d) in Duanmu (1998)
- (15) (a) *huáng zhìfú*
 yellow uniform
 ‘yellow uniform’
 (b) **huáng qìchuán*
 yellow steamboat
 (c) *huáng de qìchuán*
 yellow SUB steamboat
 ‘yellow steamboat’
 Cf. (45e/46e/47e) in Duanmu (1998)

As (15a) and (14a) show, the adjectives *huáji* ‘funny’ and *huáng* ‘yellow’ can be used in type III constructions. As (14b) and (15b) show, this is not always possible. Finally, the acceptability of (14c) and (15c) shows that this is not due to any semantic clashes between the adjective and the noun.

In the literature, the productivity data is accounted for in two ways: either by relating the productivity gaps to the idiosyncratic nature of word formation, or by explaining the gaps as a result to be expected from the naming function of the A+N type III construction. I will discuss both accounts in turn.

The first account builds on the common assumption that phrasal rules are fully productive, whereas ‘lexical rules should be conceived of as freely allowing un-systematic exceptions’ (Wasow 1977: 330). Duanmu (1998: 162) argues that the fact that *de*-less A+N constructions are ‘unproductive for most adjectives’ supports the view that these constructions are not phrasal. The criterion is again not absolute, since morphological processes can also be fully productive. The generalisation can only be that if a construction is not productive then it is not phrasal.

However, even this claim is too strong, as it does not cover the difference between lexicalized phrasal constructions and non-lexicalized phrasal constructions. The German lexicalized A+N phrases introduced in section 2 above are a case in point. The reason that they are labelled as lexicalized lies in their restricted productivity. Nevertheless, they are clearly phrasal, as is evident from (a) the fact that both the adjective and the noun inflect for number and case, and (b) their prosody, i.e. the noun

receives the main accent. Generally, the same point can be made for a large number of idioms, which are often phrasal, but nevertheless are not productive.

Paul (2005) gives a semantic explanation for the productivity gaps. According to her, the gaps in productivity are due to the conceptually different status of *de*-less versus *de*-modification structures. As mentioned above, A+N type III constructions are used like common nouns and, just as common nouns, denote properties that allow natural and plausible categorisation. Consequently, when *de*-less modification would yield unnatural and implausible categorizations, this construction type is not available. Paul (2005) uses this reasoning to account for the difference in acceptability of the examples in (16) and (17), and the argument naturally extends to examples (14) and (15) above.

- (16) (a) *zāng yīfu*
dirty clothing
'dirty clothing'
(b) **zāng táng*
dirty candy
= (48) in Paul (2005)
- (17) (a) *bái zhǐ/tóufa*
'white paper/hair'
(b) **bái shǒu*
white hand
= (49) in Paul (2005)

Paul (2005) points to Bolinger (1967), who also gives a semantic explanation for the gaps in English attributive modification, cf. (18).

- (18) (a) *absent friends* vs. **present friends*
(b) *deposited money* vs. **withdrawn money*

The problem for the semantic account of the data is that it is not always obvious why a certain A+N combination should be acceptable as a useful category for a given language, while others are not.⁸

The semantic explanation can also be used to explain partly the morphological constraint whereby no complex form adjectives can be used in this construction. Thus, Paul (2005) mentions Zhu's (1980) observation that the reduplication of adjectives usually express a subjective evaluation of the property expressed by the adjective from the point of view of the speaker. Therefore, Paul argues, it is not surprising that it cannot be used in constructions that establish new subcategories.

On the whole, the productivity data does not yield clear results.

4 Syntax-based tests

The syntactic arguments discussed by Duanmu (1998), Feng (2001), and Paul (2005) all refer to the lexical integrity hypothesis as stated in Huang (1984), cf. (19).

- (19) The Lexical Integrity Hypothesis (= 'LIH')
 No phrase-level rule may affect a proper subpart of a word
 = (12) in Huang (1984)⁹

Advocates of the compound analysis like Duanmu (1998) and advocates of the phrasal analysis, such as Paul (2005), both use test methods based on the LIH in their argumentation for the compound respectively phrasal nature of the A+N Type III construction. Their tests can be grouped into the three larger groups given in (20).

- (20) Tests based on the lexical integrity hypothesis
 A Phrasal extension and phrasal substitution
 B Coordination reduction
 C Anaphoric reference and head noun deletion data

On the face of it, the A and B tests seem to argue for the compound analysis, whereas the C test seems to support the phrasal analysis. In the following, I will show that none of these tests yields decisive results.

4.1 Phrasal extension and phrasal substitution

The first two tests to probe for lexical integrity involve the addition of optional elements or, alternatively, the substitution of smaller exemplars of a specific category with a full blown XP. The logic behind the tests is quite simple: if the A+N construction is phrasal, then the A part as well as the N part are only superficially just A and just N, but stand for some kind of maximal projection which just happens to contain only one element. Instead of just the A and just the N, we should expect a host of other realisations to be possible.

First, the data concerning the phrasal extension of the A component of the construction. If the A is just the single exponent of an AP, then it should behave like any other regular adjectival phrase and allow all kinds of further modification and adjunction, including the addition of degree adverbs etc. If it is not phrasal, these additions should be disallowed. For illustration, see the English data in (21).

- (21) (a) *the very black board*
 (b) **the very blackboard*

Comparing Mandarin A+N type II and A+N type III, we get the same contrast, cf. (22).

- (22) (a) *huáng méiguì*
 yellow rose
 (b) (**hěn/gèng/zuì/zhème/bù*) *huáng méiguì*
 very/more/most/so/not yellow rose
 (c) (*hěn/gèng/zuì/zhème/bù*) *huáng de méiguì*
 very/more/most/so/not yellow SUB rose

Note that this regularity has some exceptions (cf. Paul (2005: 774–75) and references therein as well as Huang (2006)). In particular, adjectives modified with the superlative-forming *zui* ‘most’ are sometimes able to occur without SUB, cf. (23).

- (23) (a) *zui xīn chéngjiù*
 most recent achievement
 ‘the latest achievement’
 (b) *zui gāo jìngjiè*
 most high state
 ‘the highest state (of mind or behaviour)’
 = (37) in Huang (2006)

The adjective-extension data is usually interpreted as supporting the compound analysis of the A+N type III construction. Both the A and the N element of the construction are unchangeable, because they occur here as parts of a compound, and the rules of syntax do not apply within words. The burden of evidence therefore lies on the phrasal analysis.

Paul (2005) argues that a syntactic constraint is responsible for the impossibility of modifying the adjective: the A part of the construction is restricted to a head only. She points to similar constraints on A+N constructions in languages where more than one nominal modification strategy is available. Thus, Sadler & Arnold (1994) discuss the syntactic restrictions on the size of the adjective in English A+N phrases in contrast to the A in N+A constructions, cf. the data in (24).

- (24) (a) *the proud man*
 (b) **the proud of his daughter man*
 (c) *the man proud of his daughter*

Sadler & Arnold (1994) argue that this pattern can be accounted for by assuming that the prenominal A in English A+N phrases is analysed as a small construction. Every extension of the adjective takes place at the A⁰ level. These small constructions do not consist of a head only, cf. (25a), which should be analysed as (25b).

- (25) (a) *very happy man*
 (b) [_{N⁰} [_{A⁰} [very]] [happy]] [_{N⁰} man]]

Whether the ultimate reason for this phenomenon lies in a restriction to heads or in the special properties of small constructions, the availability of different encodings can be accompanied by restrictions on the acceptable syntactic structures in the respective encodings.

A possible explanation for the syntactic constraint lies in the semantics of the construction. As mentioned in section 2, the Mandarin type III constructions only occur with naming function. Plausibly, this is the reason why no further modification or XP substitution is possible. When we further manipulate the adjective or the noun, the semantics necessarily change, and the resulting expression might not be a suitable name. Evidence in this direction comes from German lexicalized A+N phrases, introduced in section 2. These phrases have, similar to A+N compounds, naming function, and they

have two distinct formal properties (cf. Hüning forthcoming): the A cannot be further modified, and the N cannot be extended. The first property directly pertains to the discussion of the Mandarin data, the relevant German data is given in (26).

- (26) (a) *grüner Tee*
 green tea
 'green tea'
 = 1. the subtype of tea [lexicalized phrase]
 2. tea that looks green [standard intersective interpretation]
- (b) *der sehr grüne Tee*
 the very green tea
 'very green tea' = tea that looks very green

The phrase *grüner Tee* 'green tea' has two readings. As a lexicalized phrase, *grüner Tee* 'green tea' is used to refer to a subcategory of tea. In this usage, further modification of the adjective is not possible. The second reading is the standard intersective interpretation of *grün* 'green' and *Tee* 'tea', that is, it refers to something that is green and that is tea. If a modifier of the adjective is present, as in (26b), the only possible interpretation is the standard intersective modification, i.e. 'something that is tea and that is very green'. It is the only possible interpretation, because there is no 'very green' subtype of tea.

The phrasal substitution data is based on a difference between the A+N structures with *de*, type II, and those without. As Duanmu (1998) notes, the type II structure allows the substitution of the N with an XP consisting either of the combination of Numeral-Classifier N or Demonstrative N, cf. (27a). This is not possible if the *de* SUB is not present, cf. the schema (27b).¹⁰

- (27) (a) [M de N] → [M de XP]
 (b) [MN] → *[M XP]
 = (40) in Duanmu (1998)

Data exemplifying the two patterns, using a numeral + classifier string, is given in (28) and (29).

- (28) (a) *sān zhī huáng méiguī*
 three CLF yellow rose
 (b) *sān zhī huáng de méiguī*
 three CLF yellow SUB rose
- (29) (a) *huáng de sān zhī méiguī*
 yellow SUB three CLF rose
 (b) **huáng sān zhī méiguī*
 yellow three CLF rose

As far as I understand the data, it does not tell us much about the A+N type III construction. Rather, it shows that the modification with *de* and the string numeral-classifier are not ordered with respect to each other. The N part of the construction

does not seem to be touched by this, since the Num-CLF-N string, if it is an XP, must certainly be a different type of XP than just the N-part, even if that also is an XP.

4.2 Conjunction reduction

If we investigate the behaviour of the type III constructions with regard to coordination, we again get a clear difference between this construction and the construction with *de* SUB. For the *de*-construction, conjunction reduction is possible, cf. (30).

- (30) (a) *huáng de méiguì he hóng de méiguì*
 yellow SUB rose and red SUB rose
 (b) *huáng he hóng de méiguì*
 yellow and red SUB rose

For the type III construction, conjunction reduction is not possible, cf. (31).

- (31) (a) *huáng méiguì he hóng méiguì*
 yellow rose and red rose
 (b) **huáng he hóng méiguì*
 yellow and red rose

Behind the conjunction reduction test lies the assumption that conjunction reduction is only possible for constituents. Since *huáng* 'yellow' and *hóng* 'red' in (30) can participate in conjunction reduction, they must be constituents, showing that the *A-de* N-construction is a phrasal construction. That these two adjectives cannot participate in conjunction reduction in the A+N type III construction is then in turn taken as evidence for the word status of the construction. However, the relationship between conjunction reduction and constituency is not totally clear-cut.

First of all, the possibility of conjunction reduction does not automatically classify a construction as phrasal. This is already known from the data discussed in Booij (1985), where we see clear cases of word-part conjunction involving even suffixes, cf. the German data in (32).

- (32) *Freund-schaft oder Feind-schaft*
 friend-ship or fiend-ship
 'friendship or hostility'
 Cf. (3) in Booij (1985)

Freundschaft 'friendship' and *Feindschaft* 'hostility' are both formed derivationally by adding the suffix *-schaft* '-ship' to the nouns *Freund* 'friend' and *Feind* 'enemy', respectively. They are not phrasal, but conjunction reduction is possible. The same point can be made with the help of German A+N compounds, cf. (33).

- (33) *Ich habe Rot- und Weißwein vorrätig*
 I have red and white-wine at hand

Second, conjunction reduction is also possible for things that are neither word parts nor constituents, cf. (34), where *a gold Cadillac to Billy Schwartz* is elided.

- (34) *Joan offered and Mary actually gave a gold Cadillac to Billy Schwartz.*
= (38a) in Chaves (2008)¹¹

Finally, it should be mentioned that there is no conjunction in Mandarin Chinese that is comparable in universality to Germanic *and/und*. Mandarin conjunctions like *hé* ‘and’ and *gēn* ‘and’ can only be used with nominal expressions and not with verbs, verb phrases, or clauses.¹² Thus it could well be the case that the impossibility of conjunction reduction for A+N type III constructions is due to other reasons, possibly related to a size constraint, as Paul (2005: 786, fn. 14) speculates.

4.3 Anaphoric reference and head noun deletion

This test uses the availability of head noun deletion under identity as a heuristic for the phrasal-word distinction. The basic schema for this test is given in (35).

- (35) Schema for head noun deletion:
(a) A N ... A N [Two A+N constructions]
(b) A N_i ... A N_i [Head-Head identity can be established]
(c) A N_i ... A N_i [Head of second A+N is deleted/empty]

Instantiations of this schema can be found for English A+N phrases, take for example the sentence *Henrietta likes red shirts, and I like blue shirts*. This sentence contains two A+N constructions with identical heads, and, in accordance with the schema in (35), the second head can be deleted, leading to (36a). However, this pattern in its pure form is not generally available in English but governed by poorly-understood restrictions, cf. (36b).

- (36) (a) *Henrietta likes red shirts, and I like blue.*
(b) **Lucie likes friendly dogs, but I prefer aggressive.*
= (25i) and (28ii) in Huddleston & Pullum (2002: 417)

When using the proform *one*, however, the pattern is freely available, cf. (37).

- (37) (a) *Henrietta likes red shirts, and I like blue ones.*
(b) *Lucie likes friendly dogs, but I prefer aggressive ones.*

In Mandarin Chinese, this test separates the A+N type I constructions from the two other constructions. For the A+N type I construction, this pattern is not available, cf. (38).

- (38) (a) *Wǒ xǐhuān lǜ-chá, hóng-chá yě kěyǐ*
I like green-tea red-tea also possible
‘I like green tea, but black tea is also OK’
(b) **Wǒ xǐhuān lǜ-chá, hóng de yě kěyǐ*
I like green-tea red SUB also possible
Cf. (51) in Paul (forthcoming)

With the other two A+N constructions, this pattern is available, cf. (39) for the A+N type II constructions.

- (39) (a) *Wǒ xǐhuān jiù de shū, xīn de shū bù xǐhuān*
 I like old SUB book, new SUB book NEG like
 (b) *Wǒ xǐhuān jiù de shū, xīn de bù xǐhuān*
 I like old SUB book, new SUB NEG like
 'I like old books, not so much new books.'

This pattern seems to be generally available when *de* SUB is present.¹³ The data for the A+N type III construction is given in (40).

- (40) (a) *Amei bù xǐhuān huáng méiguī hóng méiguī hái kěyǐ*
 Amei NEG like yellow rose red rose still acceptable
 'Amei doesn't like yellow roses, red roses are still OK.'
 (b) *Amei bù xǐhuān huáng méiguī, hóng de hái kěyǐ*
 Amei NEG like yellow rose red SUB still acceptable
 'Amei doesn't like yellow roses, red ones are still OK.'
 = (19) in Paul (2005)

Note that in the A+N with the deleted head, the presence of *de* SUB is obligatory, cf. (41).

- (41) **Amei bù xǐhuān huáng méiguī, hóng Ø hái kěyǐ*
 Amei NEG like yellow rose red still acceptable

This is not important for us, as the test is used to establish the status of the first A+N construction, not the status of the second.

Paul (2005) takes the data to support the analysis of the A+N type III construction as phrasal. This rests on the assumption that the deletion of the second N is the result of a syntactic rule. The antecedent to the deleted head must therefore be visible to syntactic operations. Since wordparts are not visible to syntax, the A+N type III constructions must be considered as phrasal. That is, the different behaviour of type I vs. type III constructions is explained by assuming the following structures:

- (42) **Wǒ xǐhuān* [_{N⁰} *lǜ chá*], [_{NP} *hóng de* [_{N⁰} *chá*]] *yě kěyǐ*
 I like green tea, red SUB also possible
 (43) *Amei bù xǐhuān* [_{NP} [_{A⁰} *huáng* [_{N⁰} *méiguī*]],
 Amei NEG like yellow rose
 [_{NP} *hóng de* [_{N⁰} *méiguī*]] *hái kěyǐ*
 red SUB still acceptable
 'Amei doesn't like yellow roses, red ones are still OK.'

This interpretation endorses the view that words are anaphoric islands, introduced into the literature by Postal (1969). In particular, the configuration we are interested in corresponds to the case of out-bound anaphora, that is, the antecedent is embedded in a word. Out-bound anaphora is, according to Postal, impossible, as illustrated by (44).

- (44) **Max is an orphan and he deeply misses them.* [*them* = his parents]
= (Postal 1969: 206, ex. 3a, his judgement)

If Postal's claim is correct, then the Chinese data above allows no other interpretation. The alternate explanation, therefore, is built on rejecting Postal's claim.

Contra Postal's claim, Ward, Sproat & McKoon (1991) argue that so-called out-bound anaphora are fully grammatical and words therefore do not constitute anaphoric islands. In contrast, pragmatic principles are responsible for the felicity of out-bound anaphora. For illustration, cf. (45), where the antecedent for *it* is *cocaine*, which is embedded in the synthetic compound *cocaine use*.

- (45) *Although casual cocaine use is down, the number of people using it routinely has increased.*
= (22a) in Ward *et al.* (1991)

In German, anaphoric reference to the head of an A+N construction is available regardless of whether the first A+N has compound or phrasal status. In (46), we have the standard pattern where the antecedent is the head of a noun phrase.

- (46) *Peter hat den gelben Ballon wiedergefunden, den roten*
Peter has the yellow-ACC.SG balloon found.again, the red-ACC.SG
sucht er immer noch.
searches he always still
'Peter has found the yellow balloon, but he is still searching for the red one.'

In (47) and (48), the antecedent is the head of an A+N compound, *Grünglas* 'green glass (as a subcategory of glass in the context of recycling)' and *Großstadt* 'big town', respectively. In both cases, the compound-status is evident due to the main accent falling on the adjective and the adjective itself being realized in its shortform, i.e. without any inflectional ending.

- (47) *Ich bin das Grünglas losgeworden, das weiße*
I am the green-glass got.rid.off, the white-NOM.SG
liegt noch im Auto.
lies still in.the car
'I got rid of the green glass, the white glass is still in the car.'
- (48) *Ich liebe Großstädte, in kleinen gehe ich ein.*
I love big-towns, in small-ACC.PL go I in
'I love big cities, I cannot exist in small cities.'

Interestingly, the second adjective must appear in its inflected form, and not in the adjectival short form that appears in compounds, although the basis for the deletion in e.g. (47) is not *das weiße Glas* 'the white-NOM.SG. glass', but *das Weißglas* 'the white-glass'; compare the underlying structure for (47) given in (49).

- (49) ... [_{N0} Grün-glas] ... [_{NP} [_{N0} Weiß][_{N0} glas]] ...
 green-glass white glass

Apparently, if the head noun is deleted, inflection for number and case must be expressed elsewhere, leading to the use of the inflected form of the adjective. Judged from the patterns available in German, the Chinese data does not show anything about the word or phrasal status of the antecedent.¹⁴

While Postal's claim that words are anaphoric islands is stronger than at least one reading of the Lexical Integrity Hypothesis as formulated by Huang (1984), Huang makes it clear in the text that he also subscribes to the view of words as anaphoric islands. In contrast, Selkirk's (1982) *Word Structure Autonomy Condition*, which is one of the sources given by Huang for the LIH, merely prohibits the manipulation of the words themselves by rules of syntax, but leaves it open whether information about the internal structure of the words is available to syntax or not.¹⁵

5 Consequences and implications

Having presented the data and the arguments used by the proponents of the two different analyses to account for the data, the question is what all this tells us about the compound or phrasal nature of the type III construction. For most of the criteria, it is clear that whatever we make of the data has more to do with our convictions regarding morphology and syntax than with anything else. This holds especially for the arguments based on the productivity data. If it is correct that the availability of A+N type III constructions is determined by whether the resulting constructions can serve as names for new categories in a natural and plausible way, and if we allow for a certain leeway in saying that it is not always totally clear why a certain combination does or does not work as such a name in a given language, then most of the restrictions on the productivity can be accounted for. Even large parts of the morphological constraints can be explained in this way. In contrast, if we adopt the view that *phrasal* means *fully productive*, then the data points towards word-status for the A+N type III constructions. To be on the safe side, we therefore best set the data concerning the semantics and the productivity aside, and look at the results from the syntactic tests.

Where do we stand here? I think it is fair to say that the unavailability of further modification of the adjective is not a decisive argument. As argued by Paul (2005), it can be accounted for by assuming syntactic restrictions that are not in any way unusual. In addition, the same restriction exists for German lexicalized A+N phrases, pointing again to an underlying semantic explanation. As far as conjunction reduction is concerned, the interpretation of the data rests on a more fundamental point: if constructions that cannot undergo conjunction reduction are necessarily not-phrasal, then A+N type III constructions cannot be phrasal. If, however, we accept that the availability of conjunction reduction can be restricted, then it the data does not bear on the decision phrase or compound. Finally, for the test involving outbound anaphora, I have shown that it cannot be used to decide the issue.

All in all, it seems that up to here neither the pro-phrase nor the pro-word proponents have a clear case in hand, so it seems reasonable to look at the third alternative, the distinction between morphological words and syntactical words proposed by Feng (2001), in more detail. Feng (2001) uses the unavailability of further modification of the A as his main argument for saying that the A+N type III constructions are words. In turn, the reason why he labels them syntactical words lies in the ordering data given in the first section and repeated here for convenience:

- (50) (a) *dà pánzi*
big plate
'big plate'
- (b) *dà bái pánzi*
big white plate
'a big white plate'
- (c) **bái dà pánzi*
white big plate
Cf. Feng (2001)
- (51) (a) *dà guàr*
big gown
'unlined long gown'
- (b) **dà bái guàr*
big white gown
- (c) *bái dà guàr*
white big gown
'a white unlined long gown'
Cf. Feng (2001)

A type III construction like *dà pánzi* 'big plate' has to obey the general ordering restriction SIZE > COLOUR, a type I construction like *dàguàr* 'unlined long gown' does not. Thus, when the colour adjective *bái* 'white' is added, this results in *dà bái pánzi* on the one hand and *bái dàguàr* on the other hand. While this data provides further evidence for the compound status of the type I constructions, the question is how to explain the data regarding the type III construction. Feng (2001) accounts for it in the following way: the syntactical words are formed by the combination of two zero-level entities in a process of X⁰-adjunction. This is a syntactic process, and therefore the SIZE > COLOUR ordering is obeyed, but the resulting entity is still a zero-level entity, and therefore a word. As already pointed out by Paul (2005: 786, fn. 16), this analysis contains a major flaw in that the combination between *bái* 'white' and *dà pánzi* 'big plate' must be analysed as the combination of one adjective and a compound, and not, as the usage of the ordering restrictions suggests, the combination of two adjectives and a noun. Basically, the only way this could work would be to assume that *dà bái pánzi* 'big white plate' can only be derived from the combination of *dà* and *bái pánzi*, while the derivation from *bái* and *dà pánzi* is blocked, because the syntax can see the internal structure of *dà pánzi* and anticipate that a violation of the ordering restrictions would result. This seems

a very counterintuitive way, and I agree with Paul (2005) that the Chinese adjective ordering data by itself rather speaks for a phrasal status of the whole construction. How these ordering restrictions can be accounted for on the phrasal level has been discussed in detail in Scott (2002), who proposes a series of different functional projections below the DP level to account for the data. Note, though, that even the evidence of adjective ordering according to well-known cross-linguistic preferences constitutes, as most of the criteria discussed, only a one way criterion. If constructions do not obey this ordering, they can still be phrasal. Thus, incidentally, the A+N type II construction does not obey the restrictions, cf. (52).

- (52) (a) *xiǎo de lǜ de huápíng*
 small SUB green SUB vase
 (b) *lǜ de xiǎo de huápíng*
 green SUB small SUB vase
 ‘small green vase’
 = (2a) in Sproat & Shih (1991)

Another example for a departure from the SIZE > COLOUR ordering comes from the German lexicalized A+N constructions that have been mentioned before, cf. (53).

- (53) (a) *großer Zeh*
 big toe
 ‘big toe’
 (b) *großer blauer Zeh*
 big blue toe
 ‘big blue toe’ = a toe that is big and blue
 ≠ a big toe that is blue
 (c) *blauer großer Zeh*
 blue big toe
 ‘blue big toe’ = a big toe that is blue
 ≠ a toe that is big and blue

German A+N phrases obey the restriction, cf. *großer weißer Mercedes* ‘big white Mercedes’ in example (2), but if the A+N phrase is lexicalized, as *großer Zeh* ‘big toe’ in (53-a), the ordering restrictions are violated, cf. (53-b) and (53-c).

But, returning to the issue at hand: is now the distinction between morphological words and syntactical words proposed by Feng (2001) a step forward towards a solution? His specific analysis is not very intuitive or elegant, but, as far as I can see, it delivers the correct results. The idea of X⁰ adjunction is an attempt to make word-internal structure visible to syntax, while treating the construction itself as a word. In this way, lexical integrity, or rather Selkirk’s word structure autonomy are upheld, but at the same time, we have thrown in a bit of structure for the syntax to work on. We must assume that the syntax in some way or other has access to this structure, because otherwise neither the ordering data nor the outbound anaphora data can be explained. In contrast, this kind of structure is not available to the type I compounds, again explaining the lack of

adjective ordering and the impossibility of outbound anaphora. If we want to treat *de*-less constructions as compounds, Feng (2001) is on the right track.

On the other hand, it seems a fair point that the adjective ordering data is the strongest support in favour of a phrasal analysis of the type III constructions. The fact that these constructions obey the adjective ordering restrictions can be explained in an elegant way. Nonetheless, two remarks are in order. First of all, there do not seem to be too many adjectives appearing as the first adjective in A A+N type III constructions. Apart from *dà* ‘big’ and its counterpart *xiǎo* ‘small’, Sproat & Shih (1991) only mention *hǎo*, appearing as the first element of the QUALITY > SHAPE and QUALITY > COLOUR orderings.¹⁶ Second, although A A+N compounds as such appear to be very rare in German, when they do occur, their ordering is not free, cf. (54).

- (54) (a) [*Billig*] [*wein*]
 ‘cheap wine’
 (b) [*Rot*] [*wein*]
 ‘red wine’
 (c) [*Billig*] [*rot*] [*wein*]
 ‘cheap red wine’
 (d) * [*Rot*] [*billig*] [*wein*]
 red cheap wine

Again, we have here the QUALITY > COLOUR ordering, but the result is clearly a compound—obviously so by all the morphological and prosodic criteria available in German. Assuming that we can generalise that the first adjective in an A A+N compound in German and in an A+N N type III construction in Mandarin is always a gradable adjective, the ordering could also be explained on purely semantic grounds: gradable adjectives are subject to a contextually determined standard of comparison, and if material influencing this standard is overtly present, it must be in the scope of the gradable adjective. More data on possible A A+N construction in German and Mandarin is needed to settle this issue.

Apart from the ordering data, what else argues in favour of the phrasal analysis? It seems that all the other criteria are easier to accommodate in a compound analysis while the phrasal analysis has to accept that this construction combines very many properties which are already by themselves not very welcome in phrases: (a) the construction has a very specific semantic requirement that it has to result in a plausible name (b) it is morphologically restricted (c) it does not allow conjunction reduction (d) it has productivity gaps that seem erratic even given the semantic peculiarity of the construction.

6 Conclusion

This paper revolved around the question whether the A+N type III construction in Mandarin is best analysed as a phrase or as a compound. As Mandarin does not have prosodic or morphological clues to settle the issue, data concerning productivity and syntax must be used. On the basis of this data, no decision is possible. For every point

that on the face of it speaks in favour of a compound analysis, counterexamples or alternative explanations could be provided that show that these criteria by themselves are not decisive. Similarly, I argued that the main arguments in favour of the phrasal analysis, data involving adjective ordering and anaphoric reference, could also be explained in a compound analysis and furthermore that German A+N compounds exhibit exactly the same data patterns. Although the data by itself thus does not settle the issue, I believe that an analysis along the lines of the proposal by Feng (2001) is most reasonable. Through the concepts of morphological words and syntactical words we can distinguish between the A+N type I and A+N type III constructions, and with the assumption of an internal syntactic structure of the syntactical words that is visible to syntax but not changeable by syntax we can explain the adjective ordering and the anaphoric reference data, while at the same time sticking to a Selkirkian variant of the principle of lexical integrity.

Notes

1. I would like to thank Angela Grimm, Fabian Heck, and Jochen Trommer for their very helpful comments on the material presented here in its first, raw, form. Thanks also to the two anonymous reviewers, who likewise gave very helpful comments and advice. Finally, thanks to Waltraud Paul for discussing some of the issues involving the Chinese data with me.
2. There is an extensive amount of previous literature on the issue of nominal constructions with and without *de* SUB, see the references in Duanmu (1998) and Paul (2005).
3. More precisely, it is *blackcurrant* for the majority of speakers, cf. Huddleston & Pullum (2002: 1650), and *hotdog* for many speakers, cf. Huddleston & Pullum (2002: 448).
4. A prosodic distinction seems to be possible in other Chinese dialects, see Duanmu 1998. Packard (2000: 238–245) discusses data that shows that compounds that are more opaque in meaning are more likely to be realized with a neutral-toned constituent-part as well as more likely to resist to undergo third-tone sandhi and/or third-tone sandhi reversion than compounds that are more transparent.
5. Sproat & Shih (1991: 574) argue that adjectives that cannot occur as predicatives cannot occur with *de* 'SUB'.
6. This point was raised by one of the reviewers.
7. The classic reference for the basic distinction between two groups of adjectives in Mandarin is Zhu (1980). Here, I follow Paul (2006), whose classification is a slight revision of Zhu's.
8. Both Bolinger (1967) and Paul (2005) are well aware of this.
9. Huang (1984) cites Jackendoff (1972) and Selkirk (1982) as sources for the LIH.
10. Duanmu (1998) refers to Fan (1958: 214) as the origin of the observation.
11. Chaves (2008: 272) mentions that there are also accounts of this type of data as the results of multiple rightward extraposition.
12. As one reviewer pointed out, using a so-called reduced conjunction along the pattern exemplified by (i-a) is also impossible, cf. (i-b).

- (i) (a) *gōngyè hé shāngyè* → *gōngshāngyè*
 industry and commerce
 (ii) (b) **huáng hóng méiguī*
 yellow red rose

13. For a recent analysis of this pattern as N' ellipsis, see Saito, Lin & Murasugi (2008).
14. Anaphoric reference to the headnoun of an A+N compound is, nevertheless, not unrestrictedly possible in German. In particular, its availability depends on semantic transparency reminiscent of the difference in transparency between Mandarin type I and type III constructions. Thus, in (i), anaphoric reference is not possible, presumably because the birdspecies-meaning is not transparent enough.
- (i) *Mein Vater hat in seinem Garten schon mal einen Grünspecht
 My father has in his garden once a green-woodpecker
 gesehen, aber noch nie einen schwarzen ~~Specht~~.
 seen but never ever a black ~~woodpecker~~
- Intended: 'My father once saw a green woodpecker in his garden, but he has never seen a black woodpecker.'
15. Selkirk (1982: 70) explicitly mentions anaphoric reference as an operation that would be allowed under her condition.
16. Feng (2001, p. 168) only gives examples with *dà* 'big' or *xiǎo* 'small' as the first element.

References

- Bolinger, Dwight 1967. Adjectives in English: attribution and predication. *Lingua* 18: 1–34.
- Booij, Geert E. 1985. Coordination reduction in complex words: a case for prosodic phonology. In Harry van der Hulst & Norval Smith (eds.), *Advances in nonlinear phonology*. Dordrecht: Foris. 143–160.
- Chaves, Rui P. 2008. Linearization-based word-part ellipsis. *Linguistics and Philosophy* 31: 61–307.
- Fan, Jiyan 1958. Xing-ming zuhe jian 'de' zi de yufa zuoyong [The grammatical function of 'de' in adjective-noun constructions]. *Zhongguo Yuwen* 5: 213–217.
- Duanmu, San 1998. Wordhood in Chinese. In Jerome L. Packard (ed.), *New approaches to Chinese word formation*. Berlin: Mouton de Gruyter. 135–196.
- Feng, Shengli 2001. Lun hanyu 'ci' de duoweixing [the multidimensional properties of 'word' in Chinese]. *Dangdai Yuyanxue* 3: 161–174.
- Fu, Jingqi 1987. *La structure du syntagme nominal en chinois*. Doctoral dissertation, Paris: University of Paris.
- Huang, C. T. James 1984. Phrase structure, lexical integrity, and Chinese compounds, *Journal of the Chinese Language Teachers Association* 19: 53–78.
- Huang, Shi-Zhe 2006. Property theory, adjectives, and modification in Chinese. *Journal of East Asian Linguistics* 15: 343–369.
- Huddleston, Rodney & Geoffrey K. Pullum 2002. *The Cambridge Grammar of the English Language*. Cambridge: CUP.
- Hüning, Matthias forthcoming. Adjective+noun constructions between syntax and word formation in Dutch and German. In Sascha Michel & Alexander Onysko (eds.), *Cognitive approaches to word formation*, Berlin: Mouton de Gruyter.
- Jackendoff, Ray 1972. *Semantic Interpretation in Generative Grammar*. Cambridge, Massachusetts: MIT Press.
- Li, Charles N. & Sandra A. Thompson. 1981. *Mandarin Chinese. A Functional Reference Grammar*. Berkeley: University of California Press.
- Packard, Jerome L. 2000. *The morphology of Chinese: a linguistic and cognitive approach*. Cambridge: CUP.

- Paul, Waltraud 2005. Adjectival modification in Mandarin Chinese and related issues. *Linguistics* 43: 757–793.
- Paul, Waltraud 2006. Zhu Dexi's two classes of adjectives revisited. In Christoph Anderl & Halvor Eifring (eds.), *Studies in Chinese language and culture. Festschrift in honour of Christoph Harbsmeier on the occasion of his 60th birthday*. Oslo: Hermes Academic Publishing. 303–315.
- Paul, Waltraud forthcoming. Adjectives in Mandarin Chinese: the rehabilitation of a much ostracized category. In Patricia Cabredo-Hofherr & Ora Matushansky (eds.), *Formal analysis of adjectives*, Amsterdam: Benjamins.
- Postal, Paul M. 1969. Anaphoric islands. In Robert Binnick et al. (eds.), *Papers from the 5th regional meeting of the Chicago Linguistic Society*. Chicago: Chicago Linguistic Society, University of Chicago. 209–39.
- Sadler, Louisa & Douglas J. Arnold 1994. Prenominal adjectives and the phrasal/lexical distinction. *Journal of Linguistics* 30: 187–226.
- Saito, Mamoru, Lin, T.-H. Jonah & Keiko Murasugi 2008. 'N'-ellipsis and the structure of noun phrases in Chinese and Japanese. *Journal of East Asian Linguistics* 17: 247–271.
- Scott, Gary-John 2002. Stacked adjectival modification and the structure of nominal phrases. In Guilielmo Cinque (ed.), *Functional structure in DP and IP. The cartography of syntactic structures*, Vol. 1. Oxford: Oxford University Press. 91–120.
- Selkirk, Elisabeth O. 1982. *The syntax of words*. Cambridge: MIT Press.
- Sproat, Richard & Chilin Shih 1991. The cross-linguistic distribution of adjective ordering restrictions. In Carol Georgopoulos & Roberta Ishihara (eds.), *Interdisciplinary Approaches to Language. Essays in Honor of S.-Y. Kuroda*. Dordrecht: Kluwer. 565–593.
- Ward, Gregory, Sproat, Richard & Gail McKoon 1991. A pragmatic analysis of so-called anaphoric islands. *Language* 67: 439–474.
- Wasow, Thomas 1977. Transformations and the lexicon. In Peter Culicover, Thomas Wasow & Adrian Akmajian (eds.), *Formal Syntax*. New York: Academic Press. 327–360.
- Zhu, Dexi 1980. Xiandai hanyu xingrongci yanjiu [A study of adjectives in Modern Chinese]. In *Xiandai hanyu yufa yanjiu* [Research on Modern Chinese grammar]. Beijing: Shangwu yinshuguan. 3–41. Reprint of a 1956 paper.

Author' address: (Martin Schäfer)
Friedrich-Schiller-Universität Jena
Institut für Anglistik/Amerikanistik
Ernst-Abbe-Platz 8
07743 Jena
Germany
E-mail: post@martinschaefer.info