A Constraint-based Analysis of A-NOT-A Questions in Mandarin Chinese

Wenjie WangSanghoun SongFrancis BondNanyang Technological UniversitySingapore{wwang5, sanghoun, fcbond}@ntu.edu.sg

1 Basic Properties

The present study provides a constraint-based analysis of A-NOT-A questions in Mandarin Chinese within the framework of HPSG (Pollard and Sag, 1994) and MRS (Copestake et al., 2005). Hereafter, the two components in the A_1 -NOT- A_2 structure are labelled as A_1 and A_2 .

Using the A-NOT-A structure is one of the ways to express polar questions in Mandarin Chinese. The specific forms of A-NOT-A questions are exemplified below. Note that all variations presented in (1) convey almost the same meaning: "Does Zhangsan like dogs (or not like dogs)?"¹

- (1) a. Basic: A-NOT-A 张三 喜欢 不 喜欢 狗 ?
 Zhāngsān xǐhuān bù xǐhuān gǒu ?
 Zhangsan like NOT like dog PU
 - b. Contracted: A'-NOT-A 张三 喜 不 喜欢 狗 ?
 Zhāngsān xǐ bù xǐhuān gǒu ?
 Zhangsan like NOT like dog PU
 - c. Phrasal: AB-NOT-AB 张三 喜欢 狗 不 喜欢 狗 ?
 Zhāngsān xǐhuān gǒu bù xǐhuān gǒu ?
 Zhangsan like dog NOT like dog PU
 - d. Phrasal: AB-NOT-A 张三 喜欢 狗 不 喜欢 ? Zhāngsān xǐhuān gǒu bù xǐhuān ? Zhangsan like dog NOT like PU

(1) shows that A_1 and A_2 are reduplicates of each other. Partial reduplication can be performed, with the verb being reproduced without the complement, as shown in (1a) and (1d). Additionally, (1b) shows that A_1 can also be reduplicated with only its first character/syllable, while A_2 allows only one type of partial reduplication (i.e., deletion of complement).

All lexical types capable of behaving as a syntactic head of predicates in Mandarin Chi-

nese, such as verbs, adjectives, and prepositions, can participate in the A-NOT-A structure (Tseng, 2009).² Two more examples in which adjectives and prepositions are used are provided in (2).

张三 不 高 ? (2) a. 高 Zhāngsān gāo bù gāo ? Zhangsan tall NOT tall PU 'Is Zhangsan tall (or not tall)?' 张三 在不 在家 ? b. ? Zhāngsān zài bù zài jiā Zhangsan at NOT at home PU

'Is Zhangsan at home (or not at home)?'

Mandarin Chinese employs two negative operators (π bù and $\not\gtrsim$ méi), the choice of which hinges on the aspectual property of the verbal item that they are attached to, as exemplified in (3). Both of them can participate in the *A*-NOT-*A* structure with slightly different co-occurrence constraints.

(3)	a.	去不去? qù bù qù? go NOT go PU
		'Are you going?'
	b.	去没 去? qù méi qù? go NOT go PU
		'Have you gone (somewhere)?'

2 Basic Constraints

2.1 Polar Questions

In the system of expressing polar questions in Mandarin Chinese, A-NOT-A questions have a sibling, in which a sentence-final particle $\frac{n}{2}$ ma is used (henceforth, MA-questions). For example, (4) has a similar meaning to (1).

¹Due to the page limit, the abstract does not discuss with the final type.

²Frequency adverbs such as $\ddagger cháng$ "frequently" can also be the *A* element, although its reduplicated form $\ddagger \ddagger i$ is not accepted. All other adverbs are not accepted.

(4) 张三 喜欢 狗 吗?
Zhāngsān xǐhuān gǒu ma?
Zhangsan like dog MA PU
'Does Zhangsan like dogs?'

If these two forms of polar questions are were treated as allostructures of each other, the semantic representation should be almost the same in order for one form to be paraphrased into the other form. However, there are at least three reasons for believing that they are not equivalent:

First, they differ semantically. When a universal quantifier 都 $d\bar{o}u$ is used, a scope ambiguity happens with MA-questions but not with A-NOT-A questions, as shown in (5) (McCawley, 1994).

? (5) a. 他们 都 喜欢 不 喜欢 开车 tāmen dōu xǐhuān bù xĭhuān kāichē ? all like NOT like drive ΡU they 'Do they all like to drive?' 他们 都 喜欢 开车 吗 ? b.

tāmen dōu xǐhuān kāichē ma ? they all like drive MA PU 'Do they all like to drive?' or 'Do all of them like to drive?'

Second, they are pragmatically different. While the asker in MA-questions has a stance to the expressed preposition (e.g., confirmation or denial), the asker in A-NOT-A questions does not (Liing, 2014). Hence, the two types of polar questions are not necessarily interchangeable.

Third, they differ in terms of information structure. In MA-questions, focus can be assigned to any constituent. For instance, in (4), either the subject 张三 Zhāngsān, the object 狗 gǒu, or the verb 喜欢 xihuan can be evaluated as containing focus. Should focus be required, the asker employs a specific prosodic clue and/or the focus marker 是 shì. (6) presents that different constituents in MA-questions can be freely clefted.

是 喜欢 李四 吗 ? 张三 (6)a. shì Zhāngsān xǐhuān Lǐsì ma ? SHI Zhangsan like Lisi MA PU 'Is it Zhangsan (and not anyone else) who likes Lisi?' 李四 吗 张三 是 喜欢 ? b. Zhāngsān shì xǐhuān Lǐsì ma ? MA PU Zhangsan SHI like Lisi

'Is it that Zhangsan likes Lisi?'

c. 张三 喜欢 的 是 李四 吗 ? Zhāngsān xǐhuān dè shì Lǐsì ma ? Zhangsan like DE SHI Lisi MA PU

'Is it Lisi whom Zhangsan likes?'

This is mainly due to the fact that the scope of ma is not explicitly observable from the sentence itself. By contrast, A-NOT-A does not signal focus to any other elements but the structure itself (i.e., no ambiguity). The subject and the object in A-NOT-A questions cannot pass the cleft test exemplified in (6). In other words, A-NOT-A always bears focus (i.e., predicate focus).

2.2 Headedness

The presence of two semantically identical elements (even if only partially reduplicated) in A-NOT-A makes it difficult to convincingly determine whether A_1 or A_2 should be the head. As the components within A-NOT-A cannot be individually shifted or modified, nor can other elements be inserted between, headedness tests that make use of methods such as modification or movement cannot be easily applied.

This "monolithic" property of A-NOT-A means it could be seen as a single morphological word, and therefore the entire phrase is the head, and not its sub-components. Such an analysis will thus require we approach A-NOT-A from the lexicon, and include the possible A-NOT-A forms of lexical entries that can serve as the A elements. This means a lexical entry will have three variants: (i) its normal form, (ii) its basic A-NOT-A form, and (iii) its single-character contracted A-NOT-A form.

2.3 Co-occurrence Constraints

2.3.1 Sentence-final Particles

A-NOT-A questions are not permitted to occur with certain sentence-final particles. In the cases of 了 le, 吗 ma, 吧 ba, 哦 o and 耶 ye, it is because only propositions can be used with these sentence-final particles, whereas A-NOT-A is a question.

2.3.2 Aspectual Markers

As is well-known, Chinese is an aspect-based language, in which aspect is linguistically and necessarily expressed, and plays an important role in syntax. The aspect hierarchy of Chinese is roughly sketched out in (7).



Grammatical aspect in Chinese is largely expressed by verbal markers. There are three aspectual markers in Mandarin Chinese: $\Im \ le$, $\nexists \ zhe$, and $\forall \ guo$, which indicate the perfective, durative, and experiential aspects respectively. Since each verb lexically selects these markers, not all these three items can be necessarily attached to all verbs. For example, $\pm qu$ 'go' does not canonically co-occur with *zhe*. These markers are collectively known as LE-ZHE-GUO or LZG, and they are hierarchically constrained as described in (8) in Type Definition Language.

The LE-ZHE-GUO markers are also restricted in their co-occurrence with A-NOT-A, either with the entire A-NOT-A phrase, or with the individual A elements. The first two markers are not allowed to co-occur with A-NOT-A at all, while guo can only occur with A-NOT-A if the NOT element is \mathcal{Z} méi.

2.4 Character and Predicate

The A elements in A-NOT-A are full or partial reduplicates of each other. One such form is that only the first character of A_1 is reduplicated. With this in mind, we introduced four new feature types to the lexicon entries, as underlined in (9):



The feature types WCHAR and FCHAR specify all characters and the first character of a lexical entry, respectively. The feature WCHAR is identical to the STEM of the lexical entry. Next, the LENGTH specifies that an entry has one or more-than-one character. Finally, the luk feature BOUND specifies if an entry is a bound or nonbound form.³ This is to ensure that one-character A_1 forms of a multi-character word are not used outside of A-NOT-A, as they are not independent morphemes. The P-KEY feature is identical to the PRED feature so as to block homographs from co-occurring as the A elements. An example of such a homograph is 撒 sā / sǎ, which can mean 'let go' and 'scatter', respectively. These two will have different PRED values: _撒_v_1_rel and _撒_v_2_rel. Finally, the SPART feature indicates the type of sentence-final particle that can co-occur with the structure.

To provide a clearer idea, the entries in (10) illustrate the bound and non-bound forms of 喜欢, respectively. As they are identical to each other apart from length, they take the same PRED value.

(10) a.
$$\begin{bmatrix} \bar{a} & & \\ STEM & \Box \langle \cdot \bar{a} \cdot \rangle \\ BOUND & + & \\ CHAR & \begin{bmatrix} FCHAR & \cdot \bar{a} \cdot \\ WCHAR & \Box \\ LENGTH & one \end{bmatrix}$$

PRED $-\bar{a} \cdot \bar{x} \cdot v \cdot rel$
b.
$$\begin{bmatrix} \bar{a} \cdot \bar{x} & \\ STEM & \Box \langle \cdot \bar{a} \cdot \bar{x} \cdot \rangle \\ CHAR & \begin{bmatrix} FCHAR & \cdot \bar{a} \cdot \\ WCHAR & \Box \\ LENGTH & more - than - one \end{bmatrix}$$

PRED $-\bar{a} \cdot \bar{x} \cdot v \cdot rel$

3 Types of *A*-NOT-*A* Questions

3.1 Supertype

The present analysis uses the NOT element as the "origin" of the A-NOT-A structure, which will then select the A elements. A generic A-NOT-A lexical type A-NOT-A-ADV-LEX is defined for this NOT element, as shown in (11):

³The *luk* constraint consists of three components, such as +, -, and *na* (not-applicable).



Here, the element of MOD goes for A_1 , the element of COMPS goes for A_2 , and both take +vjp (verb, adjective or preposition) as their head type. Both A elements are semantically identical, so they take the same SUBJ and COMPS, and share the same ASPECT and P-KEY values. A_1 , being the head of the structure, bears the sentential force (SF) of ques. The focus meaning is represented via Individual CONStraint (Song, 2014). Its I-KEY feature points to the ICONS element, which indicates that the A-NOT-A structure is the focus of the sentence. Thus, IARG1 in ICONS is identical to INDEX of A_1 . In addition, A_2 has the constraint [BOUND –], as bound forms cannot participate as A_2 .

As mentioned before, the NOT element can be either π bù or $\not\gtrsim$ méi, depending on the A elements' aspectual property, as represented below:



The aspectual properties of their A elements are indicated in their respective COMPS' ASPECT constraints. When the NOT element is $b\dot{u}$, the A elements cannot co-occur with any of the LE-ZHE- GUO markers (*no-lzg*), whereas when the NOT element is $m\acute{e}i$, it can co-occur with $gu\dot{o}$.

As we have seen in §1, there are a few patterns for the A-NOT-A structure. With the generic A-NOT-A lexical type we defined in (11), we create two sub-types for A-NOT-A and A'-NOT-A, as shown in (3.2) and (3.3).

3.2 А-NOT-А

The sub-type for the basic form is as follows:

$$\begin{bmatrix} a \text{-not-}a \text{-}basic-adv \text{-}lex \\ MOD & \left\langle \begin{bmatrix} \text{LIGHT} & + \\ \text{WCHAR} & \square \\ \text{BOUND} & - \end{bmatrix} \right\rangle \\ COMPS & \left\langle \begin{bmatrix} \text{LIGHT} & + \\ \text{WCHAR} & \square \end{bmatrix} \right\rangle \end{bmatrix}$$

The basic form of A-NOT-A contains two identical A elements, as shown in (14):

(14) 张三 喜欢 不 喜欢 狗 ?
Zhāngsān xǐhuān bù xǐhuān gǒu ?
Zhangsan like NOT like dog PU

As such, both MOD (A_1) and COMPS (A_2) have identical WCHAR values. The MOD is constrained to [BOUND –] to block it from parsing the contracted form. Lastly, both MOD and COMPS are constrained with [LIGHT +] such that the A-NOT-A structure will be treated as a single lexical item instead of as a phrase (cf. Abeillé and Godard (2001)). The constraints presented so far will account for the following ungrammatical sentences:

3.3 А'-NOT-А

The sub-type for the contracted form is as follows:

(16)
$$\begin{bmatrix} a \text{-not-a-contracted-adv-lex} \\ \text{MOD} & \left\langle \begin{bmatrix} \text{LIGHT} & + \\ \text{WCHAR} & \square \\ \text{BOUND} & + \\ \text{LENGTH} & one \end{bmatrix} \right\rangle \\ \text{COMPS} & \left\langle \begin{bmatrix} \text{LIGHT} & + \\ \text{FCHAR} & \square \\ \text{LENGTH} & more-than-one \end{bmatrix} \right\rangle \end{bmatrix}$$



Figure 1: A sample derivation

In the A'-NOT-A variant, only the first character of A_1 is reduplicated, as shown in (17):

(17) 张三 喜 不 喜欢 狗 ? Zhāngsān xǐ bù xǐhuān gǒu ? Zhangsan like NOT like dog PU

As such, the LENGTH value of MOD (A_1) is constrained to *one*, while its WCHAR – being a single-character word – is identical to the FCHAR of COMPS (A_2) . As it is a bound form, we constrained it to [BOUND +]. In order to block it from parsing A-NOT-A sentences where the A elements are both single-character words, COMPS is given an additional constraint of *more-than-one* to its LENGTH feature. Finally, as with the basic form, both MOD and COMPS are indicated as [LIGHT +] to treat it as a single lexical item instead of a phrase.

3.4 AB-NOT-AB

The constraints for this type are as follows:



The AB-NOT-AB form's A elements are restricted to being verbs, and they are phrases instead of words. As with the basic form, the WCHAR value of the two A elements' verb heads are identical. Unlike the basic form, however, the AB-NOT-AB structure is treated as a phrase, and is thus constrained to [LIGHT –].⁴

4 Sample Derivation

In the sample derivation in Figure 1, we see the NOT element (the ADV) selecting for MOD (A_1) and COMPS (A_2) . It first combines with its COMPS via the *head-comp-phrase* rule, and then with the MOD via the *head-adj-scop-phrase* rule. As we indicate the *A*-NOT-*A* structure to be [LIGHT +], it combines to form only a V (instead of VP). The SUBJ and COMPS of both *A* elements are identical, and the *A*-NOT-*A* structure combines with the object *gŏu* via *head-compphrase*, before finally combining with *Zhāngsān* via *subj-head-phrase*.

The semantic relations are indicated in the MRS. A_1 and A_2 are given the same indexes: ARG0 for the verb itself, ARG1 for the subject *Zhāngsān*, and ARG2 for the object *gŏu*. This

⁴The current analysis does not constrain the objects B elements to be identical. The current analysis sometimes provides unwanted overgeneration. These are left to future work.

means that they share the same argument structure. The second element in the HCONS list is responsible for the scope of negative operator $b\dot{u}$: HARG is co-indexed with the ARG1 of the scopal modifier (i.e. 13), and LARG is co-indexed with the label of of A_1 (i.e. 10). The element in the ICONS list is specified as *focus*, and the values of IARG1 and IARG2 are both co-indexed with the verb's INDEX. This means that the A-NOT-A structure is associated with focus within the clause. Finally, the semantic head 2 has [SF *ques*], which indicates that the sentence is interrogative.

References

- Anne Abeillé and Daniele Godard. 2001. A Class of "Lite" Adverbs in French. In Joaquim Camps and Caroline R. Wiltshire, editors, *Romance Syntax, Semantics and L2 Acquisition: Selected papers from the 30th Linguistic Symposium on Romance Languages, Gainesville, Florida, February 2000, pages 9–26. John Benjamins Publishing Company, Amsterdam/Philadelphia.*
- Ann Copestake, Dan Flickinger, Carl Pollard, and Ivan A. Sag. 2005. Minimal Recursion Semantics: An Introduction. *Research on Language & Computation*, 3(4):281–332.
- Woan-Jen Liing. 2014. *How to Ask Questions in Mandarin Chinese*. Ph.D. thesis, City University of New York.
- James D McCawley. 1994. Remarks on the Syntax of Mandarin Yes-No Questions. Journal of East Asian Linguistics, 3(2):179–194.
- Carl Pollard and Ivan A. Sag. 1994. *Head-Driven Phrase Structure Grammar*. The University of Chicago Press, Chicago, IL.
- Sanghoun Song. 2014. A Grammar Library for Information Structure. Ph.D. thesis, University of Washington.
- Wen-Hsin Karen Tseng. 2009. A Post-syntactic Approach to the A-not-A Questions. UST Working Papers in Linguistics, Graduate Institute of Linguistics, 5(National Tsing Hua University).