Two Strategies of Sameness: Chinese *tong* and *xiangtong*

Yenan Sun
University of Chicago

1. Introduction

While 'identity' is a basic concept for humans, it is often striking how the relevant expressions in natural languages, for instance *same*, represent a big puzzle for semanticists:

(1) Darci saw a flower. Betty saw the **same** flower.

One popular view (Lasersohn 2000, Alrenga 2007a, Charnavel 2015, a.o.) is that *same* involves measuring similarity between two individual x and y in terms of contextually relevant properties and corresponds to **maximal similarity** ('$\forall P \in C[P(x) \leftrightarrow P(y)]$'). On this view, **identity** ('x=y') is a special case of maximal similarity, where the context takes into account every single property. This paper argues that the distinction between identity and maximal similarity cannot always be reduced to pragmatics and is visible formally, at least in some languages. In particular, in Mandarin Chinese the distinction is lexicalized into the determiner-like *tong* (preceding Num-CL, incompatible with demonstratives) and the adjective *xiangtong* (taking the modificational marker *de*)[1], as in (2)[2]:

(2) Daiyu kandao-le yi duo hua… 'Daiyu saw a flower…'
    a. Baochai kandao-le **tong** yi   duo hua
      Baochai see-PERF *tong* one CL flower
      'Baochai saw the same flower.'
    b. Baochai kandao-le yi   duo **xiangtong**-de   hua
      Baochai see-PERF one CL   *xiang.tong*-MOD flower
      'Baochai saw a maximally similar flower.'

This claim differs from the previous account (Liao & Wang 2014), which argued that the distinction in (2) is token-identity vs. type-identity. In the next section, I show that their account faces some problems (Sec 2.1) and motivate the new generalization that the semantic difference between *tong* and *xiangtong* is identity vs. maximal similarity (Sec 2.2). Section 3 presents a compositional analysis. Section 4 discusses the difference between *tong* and *xiangtong* on the internal reading as a further support. Section 5 concludes.

Before going into more data, I'd like to clarify another pair of familiar terms in the literature: *deictic* reading vs. *internal* reading (Dowty 1985, Carlson 1987, Barker 2007). The readings

---

[1] For a detailed overview of syntactic differences between *tong* and *xiangtong*, see Liao & Wang (2014).
[2] Abbreviations in this paper: CL = classifier, PERF = perfective marker, MOD = modificational marker.

in (2) are called *deictic* because the interpretation of *tong*/*xiangtong* is relevant to a referent introduced in the previous discourse (by the indefinite phrase *yi duo hua* 'a flower'). They also both have the *internal* reading, in which no apparent discourse anaphora is involved:

(3) nvhai-men kandao-le {**tong** yi   duo hua   /**xiangtong**-de   hua}.
     girl-PL   see-PERF   *tong* one CL flower *xiang.tong*-MOD flower
     'The girls saw {the same flower/maximally similar flowers}.'

This paper mostly focuses on the data of deictic *tong*/*xiangtong* as in (2) and will briefly discuss sentences like (3) in section 4, as an extension of the analysis.

2. The revised generalizations on *tong* and *xiangtong*
2.1 The previous analysis
Liao & Wang (2014) correctly noted that there is a meaning difference between *tong* and *xiangtong*, cf. (4)(=2). When a native speaker is asked the question 'Is there a (physical) flower that both Baiyu and Baochai saw?' – for (4a) the answer would be 'yes'; yet for (4b) the answer would be 'no'.

(4) Daiyu   kandao-le   yi   duo hua…
    Daiyu   see-PERF   one CL flower
    'Daiyu saw a flower…'
   a. Baochai kandao-le **tong** yi   duo hua
      Baochai see-PERF *tong* one CL flower
      'Baochai saw the same flower.'
   b. Baochai kandao-le yi   duo **xiangtong**-de   hua
       Baochai see-PERF one CL   *xiang.tong*-MOD flower
      'Baochai saw a flower of the same type/a maximally similar flower.'

They claimed this is a difference between token-identity vs. type identity, which at first sight is compatible with the judgments, as reflected in the translation. However, there are two pieces of evidence against their analysis (or descriptions).

First, *tong* clearly is not limited to token-identity as claimed. When *tong* is followed by a type-level classifier such as *zhong* 'kind' as in (5), the relevant reading involves identity of type-denoting expressions ('type' defined as a type e entity in the nominal domain; Dayal 2004, Alrenga 2007b).

(5) Baochai kandao-le **tong** yi   zhong hua
    Baochai see-PERF *tong* one CL_{type} flower
    'Baochai saw the same type of flower'

Though Liao & Wang mentioned (footnote 4 in their paper) that based on the definition of the type-token distinction in Vergnaud & Zubizarreta (1992), a certain level of recursion is allowed – so that (5) expresses identity of prototypical tokens. However the worry is that the original examples of 'proto-typical token' in Vergnaud & Zubizarreta all involve predicates that can only apply to kind-level objects (e.g. 'extinct' as in (6)), and it is not clear how the 'proto-typical token' can occur in the episodic contexts like (5) without admitting it is a

'type' in the first place under their definition (i.e. 'some mental entity that can be instantiated by tokens'), considering the common practice to derive the meaning of (5) is to apply Chierchia (1998: 364) 'Derived Kind Prediction'.

(6) The dodo is extinct. (The proto-typical dodo token is extinct)

Second, *xiangtong* does not always involve type-identity. Imagine a scenario in which John and Bill are eating mushrooms from a plate: there are three mushrooms in total and they are undistinguishable from each other in appearance. Now John and Bill each ate one mushroom and (7) can be said:

(7) Yuehan chi-le   yi zhi mogu.   Bier chi-le   yi zhi   (wanxuan)
    John   eat-PERF one CL_{token} mushroom Bill eat-PERF one CL_{token} completely
    xiangtong-de   mogu
    xiangtong-MOD   mushroom
    'John ate a mushroom. Bill ate an completely identical mushroom.'

Crucially (7) can be followed by a comment like (8):

(8) danshi zhiyou Bier zhongdu-le,   yinwei   [tamen chi]-de   bu shi
    but   only   Bill get.posioned-PERF because they eat-MOD   not be
    tong yi zhong   mogu.
    tong one CL_{type} mushroom
    'But only Bill got poisoned, because what they ate were not the same type of mushroom.'

In this case, no identity between types is involved since one mushroom is poisonous while another is not thus they must belong to different types. In fact the identity of types is negated in the comment.

In short, token-identity vs. type-identity cannot be the crucial difference between *tong* and *xiangtong*. The next subsection shows that the new generalization, namely the semantic distinction between *tong* and *xiangtong* is identity vs. maximal similarity, fits the data better.

2.2 The new generalization
This section recasts the data mentioned above and argues that the correct descriptive generalization on the meaning difference between *tong* and *xiangtong* should be identity vs. maximal similarity. Three further pieces of evidence are provided.

This paper argues that *tong* always involves identity of references between entity-denoting expressions, and whether it is identity of tokens or types depends crucially on whether the classifier is token-level or type-level, as in (9):

(9) Daiyu kandao-le yi duo   hua.   Baochai kandao-le **tong** yi duo/zhong   hua
    Daiyu see-PERF one CL_{token} flower   Baochai see-PERF *tong* one CL_{token}/CL_{type} flower
    'Daiyu saw a flower. Baochai saw the same {piece/type} of flower.'

In contrast, *xiangtong* always involves maximal similarity between entities (can be either tokens or types, if the classifier is not specified) and crucially it explains (10) (=7) is

felicitous even if the two mushroom tokens are not of the same type: maximal similarity (defined as 'sharing all the contextually-relevant properties') is sensitive to context, and it is possible for two mushrooms to share every property in appearance but not to be the same type (in fact, it is a very common case for mushrooms – that is why people often get poisoned after eating them)[3].

(10) Yuehan chi-le   yi zhi mogu.   Bier chi-le   yi zhi (wanxuan)
    John   eat-PERF one CL_{token} mushroom Bill eat-PERF one CL_{token} completely
    xiangtong-de   mogu
    xiangtong-MOD   mushroom
    'John ate a mushroom. Bill ate a mushroom that is maximally similar (in appearance).'

There are three further pieces of evidence indicating that the *identity* vs. *maximal similarity* distinction is on the right track.

2.2.1 Scalarity
Maximal similarity involves a universal quantification over all the (relevant) properties[4], while identity is a strict, non-quantificational relation. One potential diagnostics is that since the former provides an identifiable standard value on a scale (scalarity), it can be modified by 'almost' (Lee & Horn 1994, Amaral 2006, Alrenga 2010), parallel to universal quantifier phrases and adjectives with at least partially closed end scale (11). In contrast, identity relation is not scalar in the first place so that it cannot be modified by 'almost'.

(11) a. John likes almost {every girl/all the girls/*some girls}.
    b. The bottle is almost {full/empty/??heavy/??light}

The following contrast confirms it: in argument place, *jihu* 'almost' is bad with *tong* (regardless of the kinds of the classifiers) but good with *xiangtong*:

(12) Daiyu kandao-le yi duo hua. 'Daiyu saw a flower…'
    a. Baochai kandao-le yi duo   jihu   xiangtong-de   hua
      Baochai see-PERF   one CL_{token} almost *xiang.tong*-MOD flower
      'Baochai saw an almost identical flower.'
    b. *Baochai kandao-le jihu   tong   yi duo/zhong hua
       Baochai see-PERF almost *tong*   one CL_{token}/CL_{type} flower
      'Baochai saw almost the same piece/kind of flower.'

2.2.2 Contextual restriction
While identity is a strict relation, maximal similarity usually comes with domain restriction, encoded as a context-dependent variable C in the semantics ('$\forall P \in C[P(x) \leftrightarrow P(y)]$'). It is confirmed by the contrast that *xiangtong* can have an adverbial phrase like 'in terms of color/in color' to specify the contextual restriction of comparison, while *tong* cannot:

---

[3] I thank Itamar Francez for suggesting this scenario.
[4] Although how exactly maximal similarity is modeled can vary (e.g. empty set on the dissimilarity scale in Alrenga 2007a), something consistent is that it involves scalarity.

(13) Daiyu kandao-le yi duo hua… 'Daiyu saw a flower…'

    a. Baochai kandao-le yi   duo   (zai.yanse.shang) xiangtong-de   hua.

       Baochai see-PERF one CL_{token} in.color   *xiang.tong*-MOD flower

       'Baochai saw a flower that is maximally similar (in color).'

    b. Baochai kandao-le (*zai.yanse.shang) tong yi   duo/zhong   hua.

       Baochai see-PERF   in.color   *tong* one CL_{token}/CL_{type} flower

       'Baochai saw the same piece/kind of flower (*in color).'

2.2.3 Asymmetry between singular and plural entities

There is an interesting asymmetry between singular and plural entities when they are modified by *xiangtong*. I already show that coreference of token-denoting expressions (or identity) is only possible with *tong*, thus for the reading 'there is one house that both John and Bill cleaned' is intended, only *tong* but not *xiangtong* in (14) can be used:

(14) Yuehan dasao-le yi   jian   fangzi. Bier dasao-le   {tong yi   jian   fangzi

     John clean-PERF one CL_{token} house Bill clean-PERF   *tong* one CL_{token} house

     /#xiangtong-de   fangzi}

      xiang.tong-MOD house

     'John cleaned a house. Bill cleaned the same house/#an identical house.'

      (Int: 'There is one house that both John and Bill cleaned')

But when both John and Bill cleaned a plurality of houses, for instance in (15), then when the reading 'there are three houses that both John and Bill cleaned' is intended, the sentence with *xiangtong* suddenly improves (for at least some speakers):

(15) Yuehan dasao-le san   jian fangzi. Bier dasao-le   {tong yi   xie   fangzi

     John clean-PERF three CL_{token} house Bill clean-PERF   *tong* one CL_{PL} house

     /xiangtong-de   fangzi}

      xiang.tong-MOD house

     'John cleaned three houses. Bill cleaned the same houses/identical houses.'

      (Int: 'There are three houses that both John and Bill cleaned')

I consider that the asymmetry arises because for plural entities, there is one extra way (which is not available for atomic entities) of measuring similarity between them: by looking at to what extent the parts they contain overlap. When two plural entities have completely overlapping parts, the two can be considered as maximally similar to each other. In other words, 'having x as the part' can be considered as a special property of a plural entity, and in the case of (15), the plural entity that John cleaned is maximally similar to the plural entity that Bill cleaned in that the two entities share all the relevant properties, namely both have *m, l, g* as their parts (suppose the three houses they cleaned have names: Morewood, Lakeside, and Greenland).
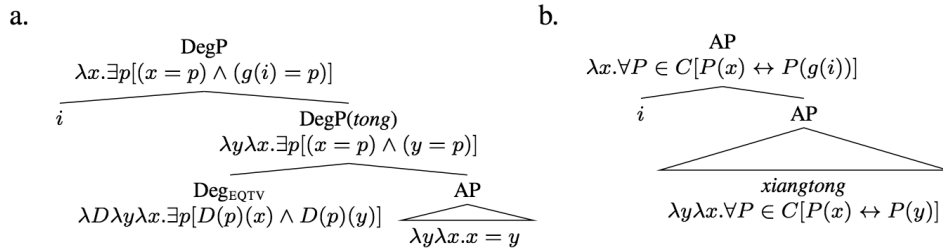
Note the identity reading of (15) does not mean that *xiangtong* expresses identity in the way that *tong* does: if it were the case, it would be a mystery why such a reading is never possible for atomic entities as in (14).

To sum up, this section motivates the new descriptive generalization on the semantic difference between *tong* and *xiangtong* such that it is a distinction between identity vs. maximal similarity. In the next section I present a compositional analysis to capture both the syntactic and semantic differences.

3. Analysis

I propose that the determiner-like *tong* should be decomposed as a comparative structure, namely a DegP headed by an equative head (16a); while *xiangtong* is a relational adjective without the DegP structure (16b):

(16)

a.

$$\text{DegP}$$
$$\lambda x.\exists p[(x = p) \wedge (g(i) = p)]$$

i $\text{DegP}(tong)$
$$\lambda y \lambda x.\exists p[(x = p) \wedge (y = p)]$$

Deg_{EQTV} AP
$$\lambda D \lambda y \lambda x.\exists p[D(p)(x) \wedge D(p)(y)]$$ $$\lambda y \lambda x. x = y$$

b.

$$\text{AP}$$
$$\lambda x.\forall P \in C[P(x) \leftrightarrow P(g(i))]$$

i AP

xiangtong
$$\lambda y \lambda x.\forall P \in C[P(x) \leftrightarrow P(y)]$$

Putting aside the complex structure of *tong* for now, both *tong* and *xiangtong* denote a two-place relation: either it is an identity relation, or a maximal similarity relation. In the deictic reading there is a referential index *i* introduced in the previous discourse, and it can saturate the relation as some kind of implicit argument. Ultimately (16a) denotes an indexical property 'being *g(i)*' while (16b) denotes an adjectival (or intersecitve) property 'being maximally similar to *g(i)*'.

Now looking back at the decomposition of *tong*: the Deg° is an equative in that it takes a relation *D* and fixes one of its arguments to some parameter *p* and returns a new relation that holds of two individuals just in case the property 'being *D* to *p*' holds of them.[5]

It might strike you that the denotation of *tong* is truth-conditionally equivalent to the plain identity relation '*λyλx. x = y*': when there is some *p* such that *x = p* and *y = p*, then of course it is the case that *x = y*. Besides the cross-linguistic evidence that 'same' in many languages involves comparative syntax (Heim 1985, Beck 2000, Charnavel 2015, Oxford 2010, Hanink 2017), I will show in Section 4 that the existence of DegP is crucial to account for the scope-taking behavior of *tong*, in contrast to the non-scope-taking adjective *xiangtong*.

The analysis of *tong* and *xiangtong* in (16) also explains the distributional fact that the former is determiner-like and the latter is a typical adjective: *tong* (after saturating its implicit argument), denoting an indexical property, takes up the specifier of *yi* 'one', which is not the ordinary numeral here but rather a strong (or anaphoric) definite article, in the sense of Schwarz (2009)[6]. The nominal phrase *tong yi duo hua* 'the same piece of flower', as in (17),

---

[5] I adopt the analysis of phrasal comparative heads as in Kennedy (1997), with minor adjustments since the AP here is not projected by a scalar adjective.

[6] See independent discussions on indexical positions in Schwarz (2009), Elbourne (2005), especially the recent ones on Chinese in Jenks (2018). In those works, the indexical position is usually in the extended projection of

is derived in (18).

(17) Daiyu kandao-le yi   duo hua.   Baochai kandao-le **tong** yi   duo/zhong   hua
Daiyu see-PERF one CL_{token} flower Baochai see-PERF *tong* one CL_{token}/CL_{type} flower
'Daiyu saw a flower$_i$. Baochai saw the same {piece/type} of flower.'

(18)

$$\iota z[flower(z) \wedge AT_{token}(z) \wedge (\exists p[(g(i) = p) \wedge (z = p)])]$$

DegP: $\lambda x.\exists p[(g(i) = p) \wedge (x = p)]$

D': $\lambda Q.\iota z[flower(z) \wedge AT_{token}(z) \wedge Q(z)]$

i DegP(*tong*)

D: *yi* $\lambda P\lambda Q.\iota z[P(z) \wedge Q(z)]$

ClP: $\lambda x.flower(x) \wedge AT_{token}(x)$

duo hua

The anaphoric definite *yi*, compared to the (weak) definite article, has an extra argument for the indexical property. There are reasons to believe *yi* 'one' has a special function here since it is the only numeral that can follow *tong* and can never be omitted. The nominal phrase in (18) thus picks out the unique flower token with the property of 'being *g(i)*'. The information that the entity is a token or a type is encoded as follows: adopting Trinh (2011), Jenks (2018), classifiers constrain the predicate *hua* 'flower' (which ranges over a polysemous domain including both flower tokens and flower types, as in Dayal 2004) to one of the domains, as illustrated in (19). If the type-level classifier *zhong* is used, it picks out the unique flower type with the property of 'being *g(i)*' (and a salient flower type can be a discourse referent *i* as well).

(19) a. $[[NP(hua)]] = \lambda x.$ flower(x)
     b. $[[CLP(duo\ hua)]] = \lambda x.$ flower(x) $\wedge$ $AT_{token}(x)$
     c. $[[CLP(zhong\ hua)]] = \lambda x.$ flower(x) $\wedge$ $AT_{type}(x)$

In contrast, *xiangtong* in (16b) ultimately denotes an intersective property – since multiple entities can be maximally similar to a certain entity, no uniqueness is guaranteed and it behaves just like a typical adjective, compatible with both definite and indefinite interpretations. The nominal phrase *yi duo xiangtong-de hua* 'a maximally similar flower' in (20) is derived as in (21).

(20) Daiyu kandao-le   yi   duo   hua. Baochai kandao-le   yi   duo   **xiangtong**-de   hua
     D.   see-PERF   one CL_{token} flower B.   see-PERF   one CL_{token} *xiang.tong*-MOD flower
     'Daiyu saw a flower$_i$. Baochai saw a maximally similar flower.'

(21) a. $[[xiangtong$-$de\ hua]] = \lambda z.$ flower (z) $\wedge$ $(\forall P \in C[P(g(i)) \leftrightarrow P(z)])$
     b. $[[yi\ duo\ xiangtong$-$de\ hua]]$
        $= \lambda z.$ flower (z) $\wedge$ $|z| = 1$ $\wedge$ $AT_{token}(z)$ $\wedge$ $(\forall P \in C[P(g(i)) \leftrightarrow P(z)])$

To sum up, this section proposes a compositional analysis of *tong* and *xiangtong*, which

---

D. I'll be vague about the exact position that the DegP headed by *tong* should be but the relevant point here is that it should at least precede Num-CL.

encodes the identity vs. max.similarity difference and the determiner-like vs. adjective distribution. In the next section I turn to examples of internal *tong* and *xiangtong*, and argue that the difference can be derived from the current analysis as well.

4. Internal readings and scope-taking

While Liao & Wang used many examples in which the internal reading of *tong/xiangtong* is relevant, they didn't note that unlike *tong* in (22), *xiangtong* cannot achieve the internal reading in a singular form: when the numeral *yi* 'one' (and the classifier) is specified as in (23a), the only available reading is that the boys all cleaned a house that is maximally similar to a previously mentioned house – thus (23a) sounds odd when uttered out of blue and cannot give rise to the internal reading like (23b):

(22) nanhai-men dasao-le   tong yi   jian   fangzi.
       boy-PL   clean-PERF *tong* one CL$_{token}$ house
       'The boys cleaned the same house.'   (Internal reading **possible**)
(23) a. nanhai-men dasao-le   yi   jian   xiangtong-de   fangzi.
          boy-PL   clean-PERF   one CL$_{token}$ *xiang.tong*-MOD house
        'The boys cleaned a maximally similar house'   (Internal reading **impossible**)
      b. nanhai-men dasao-le   xiangtong-de   fangzi.
          boy-PL   clean-PERF   *xiang.tong*-MOD house
        'The boys cleaned maximally similar houses.'   (Internal reading **possible**)

Note if *tong* and *xiangtong* are both relational terms, one being the identity relation while the other being the maximal similarity relation, it would be intriguing why (22) is possible while (23a) is not: the essence of internal reading is that without the discourse anaphora, the value of one argument of the relevant relation is settled within the sentence, as illustrated in (24):

(24) a. 'Every one of the boys cleaned a house that is exactly the house that any other boy cleaned'
      b. 'Every one of the boys cleaned a house that is maximally similar to the house that any other boy cleaned'

The similar contrast is observed in English as well: *same* can license the internal reading in a singular form while it is not possible for most relational adjectives like *similar, identical, hostile*, etc. (Matushanksy & Ruys 2007, Barker 2007, Brasoveanu 2011, Charnavel 2015)

(25) a. The girls saw the same flower.
      b. #The girls saw a similar flower/an identical flower/a hostile boy.
         (Internal reading **impossible**)

While this kind of reading of *same* has been analyzed differently in the literature (Barker 2007, Dotlacil 2010, Brasoveanu 2011, Charnavel 2015), I derive the internal reading of *tong* from the equative head of the DegP, which can take scope (*ala* Barker 2007). More specifically, the parasitic scope occurs such that both the Deg head and a plurality in the sentence (e.g. *nanhai-men* 'the boys in (22)) are scoped out and the scope of the former is

parasitic on that of the latter: [7]

(26)
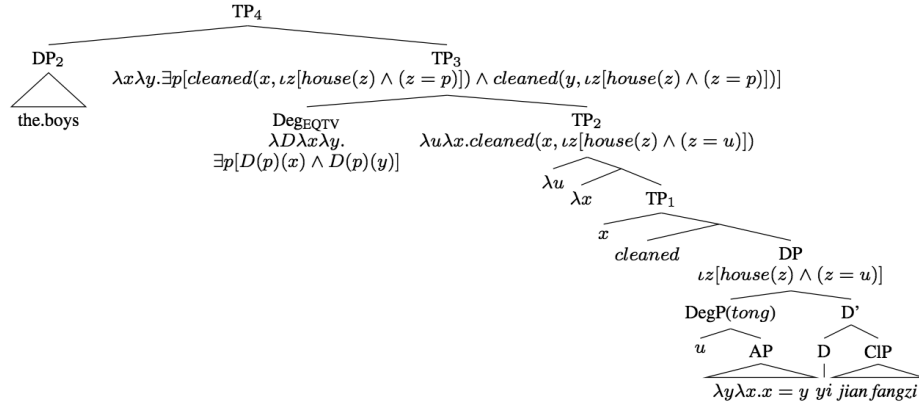
The relation created by the QR, namely [[TP$_2$]], is a relation between an individual x and the entity u such that the house x cleaned is exactly u. This relation can saturate the equative head and yield a new relation ([[TP$_3$]]) between individuals x and y such that the house x cleaned is exactly the house that y cleaned.

The potential type mismatch between [[TP$_3$]]$_{<e,et>}$ and [[DP$_2$]]$_e$ triggers the application of *Hmg* (homogeneity, based on Beck 2000, 2001; Schwarzschild 1996). This operation freely transfers any symmetric relation R into a property of a plural individual X such that R holds between all the atomic parts of X[8]:

(27) Operation *Hmg*: For any symmetric relation R, $[[R^{Hmg}]] = \lambda X. \forall x, y \leq X[R(x)(y)]$.

Since [[TP$_3$]]$_{<e,et>}$ is such a symmetric relation, applying *Hmg* as in (28a) distributes this symmetric relation between all the atomic parts of plural subject 'the boys' (its denotation is represented as a plural individual B here), deriving the internal reading as in (28b).

(28) a. $[[TP_3{}^{Hmg}]]$
$= \lambda X. \forall x, y \leq X[\exists p[\text{cleaned}(x, \iota z[\text{house}(z) \wedge (z{=}p)]) \wedge \text{cleaned}(y, \iota z[\text{house}(z) \wedge (z{=}p)])]]$
b. $[[TP_4]]$
$= \forall x, y \leq B[\exists p[\text{cleaned}(x, \iota z[\text{house}(z) \wedge (z{=}p)]) \wedge \text{cleaned}(y, \iota z[\text{house}(z) \wedge (z{=}p)])]]$

In short, the internal reading of *tong* is possible due to the scope-taking Deg head and the *Hmg* operation.

---

[7] To avoid making the tree oversized, I omit the contribution of classifiers in the formula since they are not relevant here.
[8] It seems to be a very common operation in languages, for instance *get married* is a symmetric predicate and it can switch between a 2-place predicate as in (ia) and a 1-place predicate taking a plural individual as in (ib).
  (i) a. John got married with Mary.
     b. They got married.

Now why the internal reading of *xiangtong* cannot be licensed unless the host it modifies is plural, as in (29) (=23)?

(29) nanhai-men dasao-le   (#yi jian)   xiangtong-de   fangzi.
     boy-PL   clean-PERF   one CL_{token} *xiang.tong*-MOD house
     'The boys cleaned maximally similar houses   (Internal reading)

It is expected under our account since *xiangtong* is an ordinary relational adjective thus cannot take scope like *tong*, namely to QR and yield a relation between an individual *x* and an entity *u* that such *x* saw a flower that is maximally similar to *u*. The reason that only when the host it modifies is in plural form it can license the reading is because *Hmg* can apply to *xiangtong* directly, as in (30), considering *xiangtong* itself is a symmetric relation:

(30) a. $[[xiangtong\text{-}de^{Hmg}]] = \lambda X. \forall x,y \leq X[\forall P \in C[P(x) \leftrightarrow P(y)]]$
     b. $[[xiangtong\text{-}de^{Hmg} fangzi]] = \lambda X. house^*(X) \wedge (\forall x,y \leq X[\forall P \in C[P(x) \leftrightarrow P(y)]])$

In this case, (30a) is a property must be held of an entity that contains parts, namely a plural entity, which accounts for the obligatory absence of numeral *yi*. In other words, the internal reading of *xiangtong* in (29) is actually a special case of reciprocal reading of *xiangtong*[9]: the boys cumulatively cleaned the houses that are maximally similar to each other. As expected, when an obligatorily distributive operator *gezi* 'each' is inserted, the internal reading is not available for (31) (when uttered out of blue):

(31) nanhai-men gezi   dasao-le   xiangtong-de   fangzi
     boy-PL   each   clean-PERF   *xiang.tong*-MOD house
     'The boys each cleaned maximally similar houses.'   (Internal reading unavailable)

To conclude, the complex structure of *tong* exactly accounts for its exceptional scope-taking behavior, compared to the ordinary relational terms like *xiangtong*.

5. Conclusions
This paper, based on the data of Mandarin Chinese, advocates the empirical distinction between identity and maximal similarity, which the grammar seems to be sensitive to in various ways. However, after emphasizing this distinction throughout the paper, there is nevertheless obvious morphological overlapping between Chinese *tong* and *xiangtong*, suggesting the potential connection between identity and maximal similarity. I do not have a full story for it now but the initial hypothesis is that while identity is an equation between type *e* individuals, maximal similarity can be seen as an equation between the type-shifted (Partee 1986) individuals, namely the sets of properties that individuals have (type $<et,t>$). In this way, the equation (identity) is the more basic meaning and maximal similarity is coerced from it.

Another interesting question is to ask is that whether such a distinction is systematically made in the grammar of other languages. Due to limited space, I cannot extensively discuss

---

[9] This point is similar to what Beck (2000) argued for the plural NP dependent reading of German *verschieden* 'different' such that it is actually a special case of reciprocal reading.

the facts but a quick look to English shows that the familiar *identity* vs. *maximal similarity* distinction can be found as well, i.e. between NP-internal *same* (as in 'the same flower) and predicative *same* (as in 'flower that is the same'):

Obligatorily definite or not:
(32) a. Darci saw a flower. Betty saw {the/*a} same flower.
    b. Darci saw a flower. Betty saw {the/a} flower that is the same.

Can be (freely) modified by 'almost':
(33) a. Darci saw a flower. ??Betty saw almost the same flower.
    b. Darci saw a flower. Betty saw a flower that is almost the same.

Overt contextual restriction:
(34) a. Darci saw a flower. *Betty saw the same flower in color.
    b. Darci saw a flower. Betty saw a flower that is the same in color.

Internal reading in a singular form or not:
(35) a. The boys cleaned the same house.
    b. The boys cleaned {#a house that is the same/houses that are the same}.

Note that the predicative *same* patterns with the typical adjectives such as *identical* and *similar* in terms of the above tests, suggesting they all belong to the similarity-based side. It is interesting to see that in both Chinese and English, those similarity-based terms are lexicalized into (typical) adjectives while identity terms such as *tong* and NP-internal *same* seem to be more functional.

References

Alrenga, Peter. 2007a. Dimensions in the semantics of comparatives. Doctoral Dissertation, University of California, Santa Cruz.
Alrenga, Peter. 2007b. Tokens, types, and identity. In *Proceedings of the North East Linguistic Society (NELS)*, volume 38.
Alrenga, Peter. 2010. Comparisons of similarity and difference. *Adjectives. Formal Analyses in Syntax and Semantics*.
Amaral, Patricia. 2006. On the semantics of almost. In *Annual Meeting of the Linguistics Society of America, University of New Mexico*.
Barker, Chris. 2007. Parasitic scope. *Linguistics and Philosophy* 30:407–444.

Beck, Sigrid. 2000. The semantics of *different*: Comparison operator and relational adjective. *Linguistics and Philosophy* 23:101–139.

Beck, Sigrid. 2001. Reciprocals are Definites. *Natural language semantics* 9:69–138.

Brasoveanu, Adrian. 2011. Sentence-internal different as quantifier-internal anaphora. *Linguistics and philosophy* 34:93–168.

Carlson, Greg N. 1987. Same and different: Some consequences for syntax and semantics. *Linguistics and Philosophy* 10:531–565.

Charnavel, Isabelle. 2015. Same, different and other as comparative adjectives–a uniform analysis based on French. *Lingua* 156:129–174.

Chierchia, Gennaro. 1998. Reference to kinds across language. *Natural language semantics* 6:339–405.

Dayal, Veneeta. 2004. Number marking and (in) definiteness in kind terms. *Linguistics and philosophy* 27:393–450.

Dowty, David. 1985. A unified indexical analysis of same and different: A response to Stump and Carlson. In *University of Texas Workshop on Syntax and Semantics, Austin, Texas*.

Elbourne, Paul D. 2005. *Situations and individuals*, volume 90. Mit Press Cambridge, MA.

Hardt, Daniel, and Line Mikkelsen. 2015. Same but different. *Linguistics and Philosophy* 38:289–314.

Ionin, Tania, and Matushansky, Ora. 2006. The composition of complex cardinals. *Journal of Semantics* 23:315–360.

Jenks, Peter. 2018. Articulated definiteness without articles. *Linguistic inquiry* 49: Jenks, P., 2018. Articulated definiteness without articles. *Linguistic Inquiry* 49:501-536.

Kennedy, Christopher. 1997. Projecting the adjective: The syntax and semantics of gradability and comparison. Doctoral Dissertation, University of California, Santa Cruz.

Lasersohn, Peter. 2000. Same, models and representation. In *Semantics and linguistic theory* 10: 83–97.

Lee, Young-Suk, and Laurence Horn. 1994. Any as Indefinite plus. *Manuscript, Yale University.*

Liao, Wei-wen Roger, and Yuyun Wang. 2014. The same difference: Comparative syntax-semantics of English same and Chinese *tong*/*xiang-tong*. *Chinese Syntax in a Cross-Linguistic Perspective*:128.

Matushansky, Ora. 2010. Same problem, different solution. *Ms., University of Utrecht*.

Matushansky, Ora, and EG Ruys. 2007. *Same* in Russian. *Foral Seatics i Moscow* 3.

Oxford, Will. 2010. *Same*, *different*, and *other*: The microsyntax of identity adjectives. *Ms. University of Toronto*.

Partee, B. 1986. Noun phrase interpretation and type-shifting principles (reprinted in B. Partee (2004), Compositionality in Formal Semantics): 203–230.

Schwarz, Florian. 2009. *Two types of definites in natural language*. University of Massachusetts Amherst.

Trinh, Tue. 2011. Nominal reference in two classifier languages. *Sinn und Bedeutung 15*: 629.

Vergnaud, Jean-Roger, and Maria Luisa Zubizarreta. 1992. The definite determiner and the inalienable constructions in french and in english. *Linguistic inquiry* 23:595–652.