Mandarin SHENME as a superweak NPI

Jing Lin, University of Amsterdam
Fred Weerman, University of Amsterdam
Hedde Zeijlstra, University of Göttingen

Abstract: In the past thirty years, Frans Zwarts has written several papers providing crucial insight in licensing contexts for Negative Polarity Items (NPIs), presenting a more nuanced picture than Ladusaw's (1979) downward entailing (DE) requirement. Zwarts demonstrated (1981) that a number of Dutch NPIs appear only in a subset of DE-contexts, and proposed (1995) non-veridicality as a logico-semantic property that licenses so-called superweak NPIs. Such superweak NPIs, however, have hardly been attested. We show that Mandarin shenme *('a (thing)') is a prototypical superweak NPI. We explain its ungrammaticality in veridical contexts by arguing that* shenme *exhibits a lexical referential deficiency. Acquisitional data, furthermore, suggest that children initially analyze* shenme *as a WH-quantifier but acquire the referential deficiency underlying its NPI status after the age of four.*

1. Introduction

Negative Polarity Items (NPIs) are lexical items or idioms that may only appear in some kind of negative contexts. See (1) for such a distribution of an NPI, i.e., English *any*.[17] The NPI is marked in italics.

(1) a. It is *(not) the case that John saw *any* robins.

   b. Nobody/*Somebody saw *any* robins.

   c. Few/*Many people saw *any* robins.

Given this distribution, Ladusaw (1979) proposes that NPIs such as English *any* are restricted to Downward Entailing (DE) contexts only – contexts satisfying an entailment relation from set to subset:[18] under the scope of sentential negation as in (1a), under the scope of negative indefinites as in (1b) and under the scope of semi-negative quantifiers

---

[17] See also Ladusaw (1979) for the distribution of the NPI *any*.

[18] A function F is Downward Entailing, iff for every arbitrary X and Y, it holds that $X \subseteq Y \rightarrow F(Y) \subseteq F(X)$ (adapted from Zwarts 1993).

as in (1c).[19] However, not all NPIs are licensed in exactly the same set of DE-contexts. As first noticed in Zwarts (1981), Dutch *mals* ('soft') and *ook maar* ('at all'), for example, may only appear in certain kinds of DE-contexts (see (2) and (3)).[20]

(2) a. Deze   kritiek   is   niet *mals*   .
      this   critique is   not mild
      'This critique is not mild.'

    b. *Geen   kritiek   is   *mals*   .
      no   critique is   mild
      Intended: 'No critique is mild.'

    c. *Weinig kritieken zijn *mals*   .
      few   critiques are mild
      Intended: 'Few critiques are mild.'

(3) a. Het lukt   niet om *ook maar*   één vis te   vangen .
      it   works   not to at all   one fish to   catch
      'It is not (even) possible to catch one fish at all.'

    b. Ik   heb   nooit   *ook maar*   één vis gevangen   .
      I   have   never   at all   one fish caught
      'I have never caught a fish at all.'

    c. *Weinig mensen hebben *ook maar*   één vis gevangen   .
      few   people have   at all   one fish caught
      Intended: 'Few people have caught a fish'.

On the other hand, the distribution of NPIs like English *any* and Dutch *enig* ('any') even extends beyond DE-contexts, i.e., in polar questions and in complement clauses of intensional verbs as shown in (4) and (5).[21]

---

[19] Other DE-contexts are conditional clauses, restrictive clauses of a universal quantifier and comparative clauses, etc.

[20] For the distribution of the NPI *ook maar* ('at all'), see Zwarts (1993) and Giannakidou (1997); for the distribution of the NPI *bijster* ('very'), see Zwarts (1993) and Van der Wouden (1997).

[21] For the distribution of the NPI *any* beyond DE-contexts, see Giannakidou (1998, 1999); for the distribution of the NPI *enig* ('any'), see Hoeksema (2010) and Giannakidou (2010).

(4) a. Did you see *any* students?

    b. I guess you saw *any* students.

(5) a. Heb   je   *enig* probleem   met NPIs ?

       have   you any problem   with NPIs

       'Do you have any problem with NPIs?'

    b. Ik   geloof   dat je   *enig* probleem   hebt   met NPIs   .

       I   believe   that you any problem   have   with NPIs

       'I believe that you have any problem with NPIs.'

As to explain NPIs' distribution in different kinds of contexts, Zwarts (1993) proposes that NPIs come about in different strengths depending on the negativity of their licensing conditions. Superstrong NPIs (Dutch *mals*) may only appear in stronger negative contexts such as under the scope of *niet* ('not'), i.e., anti-morphic contexts;[22] strong NPIs (Dutch *ook maar*) are restricted to strong negative contexts like under the scope of *niemand* ('nobody') or *zonder* ('without'), i.e., anti-additive contexts;[23] weak NPIs (Dutch *ooit* ('ever') and English *any*) are merely licensed in weak negative contexts – DE-contexts.[24] This suggests that NPIs such as Dutch *enig* are even weaker than *ooit* or *any*, since they may also appear in weaker negative contexts compared to DE-contexts. As to capture the distribution of such weaker NPIs, Zwarts (1995) introduces non-veridicality and claims that non-veridical contexts license NPIs such as *enig*. As non-veridical contexts are the weakest type of negative contexts, we refer to those NPIs of such weaker strength as superweak NPIs in this paper (see also Hoeksema 2012).

    Nonetheless, the literature hardly attests any superweak NPIs that are excluded from all veridical contexts. By examining the distribution of Mandarin indefinite *shenme* ('a (thing)') in spoken Mandarin, however, we show that *shenme* is a prototypical NPI of the superweak strength, allowed only in non-veridical contexts. Assuming that *shenme* is lexically deficient in referring, developed from Giannakidou (1998, 1999) and Lin (1996, 1998), we provide an explanation for *shenme*'s grammaticality in non-veridical contexts only. Moreover, by presenting acquisitional data collected in a corpus study in the CHILDES database (MacWhinney 2009), we show that Mandarin children acquire the

---

[22] A function *F* is *anti-morphic*, iff for every arbitrary X and Y, it holds that $F(X \cup Y) \Leftrightarrow F(X) \cap F(Y)$ and $F(X \cap Y) \Leftrightarrow F(X) \cup F(Y)$ (adapted from Zwarts 1993).

[23] A function *F* is *anti-additive*, iff for every arbitrary X and Y, it holds that $F(X \cup Y) \Leftrightarrow F(X) \cap F(Y)$ (adapted from Zwarts 1993).

[24] See Hoeksema (1999) for the distribution of *ooit* ('ever') in DE-contexts only.

superweak NPI by initially analysing it as an interrogative indefinite and then reanalyse it as a nonreferential existential quantifier after the age of four.

The paper is structured as follows. Section 2 introduces the definition of non-veridical contexts. Section 3 examines the restrictions in distribution of *shenme* to different nonveridical contexts in Mandarin Chinese, which leads to the conclusion that *shenme* is a superweak NPI. We establish an analysis of *shenme* in Section 4 that accounts for why this indefinite is only banned from veridical contexts, i.e., being a superweak NPI. Section 5 focuses on language acquisition. We discuss data collected in CHILDES and propose an explanation for how Mandarin children acquire the superweak NPI such that they obtain the knowledge of *shenme*'s referential deficiency in its lexical semantics. Section 6 concludes.

2. Non-veridical contexts

Zwarts (1986, 1993) defines (non)veridicality in terms of truth, see below.

(6)   (Non)veridicality for propositional operators
         A propositional operator F is veridical, iff Fp entails p: Fp $\models$ p; otherwise F is non-veridical.

Informally, a veridical context is a context in which the truth of a proposition can be entailed. Complement clauses of factive verbs, for instance, are veridical, since the truth of the proposition (7b) is entailed by (7a).

(7)   a.   I know you are busy.
         b.   You are busy.

On the other hand, a non-veridical context is a context in which the truth of a proposition cannot be entailed. Contexts that exhibit non-veridicality are polar questions, imperatives, complement clauses of intensional verbs, imperfectives, etc. Polar questions are non-veridical because sentences like (8a) do not entail the truth of (8c). In the same vein, complement clauses of intensional verbs are non-veridical as well: (8b) does not entail the truth of (8c), either.

(8)   a.   Are you busy*?*
         b.   I guess you are busy.
         c.   You are busy

Non-veridical contexts form a weaker type of negative contexts than DE-contexts. As proven in Zwarts (1993), DE-contexts and non-veridical contexts stand in a subset relationship with each other. All DE-contexts are non-veridical but not the other way around.[25] This can also be presented by means of a hierarchy (adapted from Zwarts 1995, see also Van der Wouden 1994, 1997 and Hoeksema 2012).
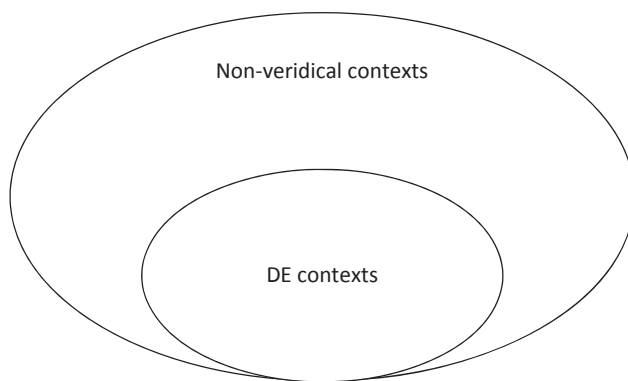

Figure 1: The subset relationship between DE and non-veridical contexts

3. *Shenme* as a superweak NPI

Traditional Chinese grammars categorised *shenme* as an interrogative pronoun with some non-interrogative functions (Li 1924, Lü 1982, Ding 1961 and Zhao 1979, among others). This is because besides its interrogative interpretation as shown in (9a),[26] *shenme* may also appear in some non-interrogative sentences functioning as a pronoun of *XuZhi* ('vague reference') in (9b), or that of *RenZhi* ('free choice reference') in (9c) or that of *BudingZhi* ('unspecific reference') in (9d).

---

[25] Such a subset relationship also applies to anti-morphic contexts, anti-additive contexts and DE-contexts. Anti-morphic contexts are anti-additive but not the other way around, and anti-additive contexts are DE but not vice versa.

[26] The term *interrogative sentences* in this paper does not cover polar questions but refers only to those interrogative sentences that are introduced by an interrogative pronoun in a traditional sense, i.e., *shenme*.

(9) a. Ni zuotian   mai le *shenme* (ne) ?[27]
      you yesterday   buy PRF *shenme* Q-marker
      'What did you buy yesterday?'

   b. Ta haoxiang shi zai xie   *shenme* .
      s/he probably COP at write   *shenme*
      'S/he is probably writing something.'

   c. *Shenme* shuiguo wo dou ai   chi .
      shenme fruit   I   all love eat
      'I love to eat all fruit.'

   d. wo lai   mai xie   shu   he bi   *shenme* de .
      I   come   buy some   books and pens *shenme* PAR
      'I come here to buy some books, pens and other things like that.'

The facts in (9b) to (9d) led some scholars to conclude that *shenme* can appear as a polarity item in some non-interrogative contexts (Huang 1982, Cheng 1994, 1995). This is supported by the following examples, where a non-interrogative reading of *shenme* is unavailable in simple affirmative clauses as in (10a), in perfectives as given in (10b) or in complement clauses of a factive verb, see (10c) (Lin 1996, 1998, Li 1992 and Xie 2007).[28]

---

[27] In Mandarin Chinese, a Q-marker may be either overt or covert in interrogative sentences (e.g. Ni 2005).

[28] Cheng (1994, 1995) also observes that *shenme* is ungrammatical in the subject position of Mandarin X-NEG-X questions, a specific type of polar questions in Mandarin Chinese (see further footnote 12). This is shown by the examples below.

(i) **Shenme* huai   mei huai   (ne) ?
     shenme broken   NEG broken   Q-marker
     Intended: 'Is there anything broken or not?'

(ii) **Ta xiang   cha   qingchu *shenme* huai   mei huai .
     s/he want check   clearly   *shenme* broken   NEG broken
     Intended: 'S/he wants to check carefully if there is anything broken or not.'

The reason why *shenme* is not allowed in the subject position of Mandarin X-NEG-X questions is syntactic in nature, and therefore differs from why *shenme* may not appear in (10a) to (10c). It here concerns the scope of the Mandarin X-NEG-X operator. The reader is referred to Cheng (1989) and Huang (1982, 1990) for the syntactic structure of X-NEG-X questions in Mandarin Chinese and to Li (1992) for an explanation for *shenme*'s ungrammaticality in (i) and (ii).

(10) a.   Ta   zuotian   chi le *shenme* pingguo .

          s/he   yesterday   eat PRF *shenme* apple

          *'S/he ate an apple yesterday.'

          'What kind of apples did s/he eat yesterday?'

     b.   Ta cengjing tou   guo *shenme* shoushi .

          s/he once   steal   PRF *shenme* jewellery

          *'S/he has once stolen some jewellery.'

          'What kind of jewellery has s/he ever stolen?'

     c.   Laoshi   zhidao   ta shuo   le *shenme* hua .

          teacher know   s/he say   PRF *shenme* word

          *'The teacher knows that he said something.'

           'What does the teacher know that s/he said?'

           'The teacher knows what s/he said.'

Adopting the polarity perspective of Huang (1982) and Cheng (1994, 1995), we show in this section that *shenme* – irrespective of its (non)interrogative interpretation – is an NPI of the superweak strength, restricted to non-veridical contexts only. Compared to the previous approaches, our treatment of *shenme* as a superweak NPI affords a unified understanding of both its interrogative and non-interrogative functions. Moreover, it introduces the Mandarin quantifier into the landscape of NPIs proposed in Zwarts (1993), providing evidence for non-veridicality as a licensing property for prototypical superweak NPIs (Zwarts 1995).

     We start with an overview of linguistic contexts that can license *shenme* by reviewing the literature.

Linguistic context	Reference
Under the scope of a negative marker	Cheng (1994, 1995), Lin (1996, 1998)
Restrictive clauses of a universal quantifier	Cheng (1994, 1995), Lin (1996, 1998)
Conditional clauses[29]	Cheng (1994, 1995), Lin (1996, 1998)
BEFORE-clauses	Lin (1996, 1998), Xie (2007)
Interrogative contexts	Li (1924), Lü (1982, 1985)
Imperfectives[30]	Li (1992), Lin (1996, 1998), Xie (2007)
Imperatives	Li (1992), Lin (1996, 1998)
Modal contexts	Li (1992), Lin (1996, 1998), Xie (2007)
Polar questions[31]	Cheng (1994, 1995), Lin (1996, 1998)
Complement clauses of non-factive verbs[32]	Lin (1998)
Under the scope of inference-marker *le*	Lin (1998)

Table 1:[33]   Linguistic contexts sanctioning *shenme*

---

[29] This category includes *donkey*-sentences as well, see Cheng & Huang (1996) and Li (2002).

[30] This category includes futural, progressive and habitual aspects.

[31] Polar questions here include X-NEG-X questions that are typical in Mandarin Chinese. An X-NEG-X question contains, as its name predicts, an X-NEG-X construction, which is absent in a generic polar question. In such an X-NEG-X construction, X refers to a lexical element of any morphological category, such as a noun, a verb or an adjective and NEG refers to a negative maker, i.e., *bu* or *mei*. Some examples of Mandarin X-NEG-X examples are given below.

(i)   Ni   jintian   wan mei wan (ne) ?
       you   today   late NEG late Q-marker
       'Were you late today, or not?'

(ii)   Ta   erzi shi bu   shi hen congming (ne) ?
       His/her   son COP NEG COP very clever   Q-marker
       'Is his/her son very clever or not?'

(iii) Nimen   zuowan   shui mei shui (ne) ?
       you   last night sleep   NEG sleep   Q-marker
       'Did you sleep last night or not?'

[32] As pointed out by a reviewer, the set of non-factive verbs is heterogeneous since it does not only refer to intensional verbs such as *believe* or *guess*, but also contains various negative predicates like *impossible* or *fail.* As we will show in (13a), *shenme* can be licensed by all such predicates in Mandarin Chinese. Note that we do not classify verbs that entail the truth of their embedded propositions (e.g. 'prove') as non-factives.

[33] Examples of *shenme* occurring in each of the linguistic context listed in Table 1 are given in Appendix 1.

In table 1 we listed a total of 12 linguistic contexts in which *shenme* is allowed to appear according to the literature. We now categorise these 12 contexts depending on their degree of negativity. Negative contexts introduced by a sentential negative marker are anti-morphic. Restrictive clauses of a universal quantifier, conditional clauses and *BEFORE*-clauses are typical DE contexts. *DONKEY*-sentences in Mandarin Chinese are DE as well because the entailment relationship from (12a) to (12b) holds.

(12) a. Ni xiangyao *shenme*   shuiguo, wo jiu   gei ni   mai *shenme* shuiguo
         you want   *shenme*   fruit   I   then   for you buy *shenme* fruit
         'Whatever fruit you would like to have, I will buy it for you.'
     b. Ni   xiangyao *shenme* pingguo, wo jiu gei ni   mai *shenme* pingguo
         you want   *shenme* apple   I   thenfor you buy *shenme* apple
          'Whatever apple you would like to have, I will buy it for you.'

Interrogative sentences, imperfectives, imperatives and modal contexts introduced by epistemic modal adverbs are prototypical non-veridical contexts. Section 2 already illustrates that polar questions and complement clauses of intensional verbs exhibit non-veridicality. According to Lin (1998), sentences marked by the inference marker *le* in Mandarin Chinese are non-veridical as well, since as an indicator for circumstantial inference, *le* expresses epistemic modality and "may allow a speaker to infer that something must have happened only on the basis of his/her observation of the environment without witnessing the event or changing state" (Lin 1998: 223). To summarize, all the 12 linguistic contexts listed in Table 1 exhibit at least non-veridicality.

    In addition to these linguistic contexts, however, we also observe that *shenme* can appear in the following kinds of negative contexts. They are negative contexts introduced by an inherently negative verb or a negative quantifier as shown in (13a) and (13b) respectively, and those introduced by a negative universal quantifier as given in (13c). These negative contexts are all DE; the contexts illustrated in (13a) and (13b) are anti-additive as well.

 (13) a. ta   fouren shuo   guo *shenme* fashehui   de   hua .[34]
        he   deny   say   PRF *shenme* antisocial   MOD   word
        'He denied having said any antisocial word.'

---

[34] This sentence can also be assigned an interrogative interpretation if uttered with a rising intonation and/or in the presence of a Q-marker *ne*: 'What is that antisocial word that he denied to have said?'.

   b.   meiren   shuo   guo *shenme* fanshehui   de   hua .

       nobody say   PRF *shenme* antisocial   MOD   word

       'Nobody said any antisocial word.'

   c.   bushimeigeren   dou shuo   le *shenme* .[35]

       not everybody   all   say   PRF *shenme*

       'Not everybody said anything/something.'

As introduced already in Section 2, anti-morphic contexts, anti-additive but not anti-morphic contexts and DE contexts are all non-veridical. This means that all the attested linguistic contexts in which *shenme* can appear are non-veridical. But is the distribution of *shenme* also restricted to non-veridical contexts? As illustrated in (10), it is infelicitous to use *shenme* in simple affirmative clauses, in perfectives or in complement clauses of a factive verb. The fact that these contexts are all veridical confirms *shenme*'s restricted distribution to non-veridical environments only.[36] We therefore conclude that *shenme* is a superweak NPI that requires at least non-veridical contexts as felicitous licensing conditions.

As to provide empirical evidence for *shenme*'s status as a superweak NPI, we did a corpus investigation by employing a subcorpus of the *PKU-CCL YuLiaoKu* (the PKU-CCL Corpora), in particular *KouYu* ('spoken Mandarin'). The corpus results are summarized in the table below, which presents a quantitative overview of how the indefinite is distributed in spoken Mandarin.

Non-veridical	Count (percentage)	Veridical	Count (%)
Anti-morphic	86 (9.21%)	*perfectives*	1 (0.11%)
Anti-additive but not anti-morphic	1 (0.11%)		
DE but not anti-additive	130 (13.92%)		
Non-veridical but not DE	716 (76.66%)		
TOTAL	**933 (99.89%)**	**TOTAL**	**1 (0.11%)**

---

[35] This sentence can also be assigned an interrogative interpretation if uttered with a rising intonation and/or in the presence of a Q-marker *ne*: 'What is that that not everybody has mentioned?'.

[36] As already mentioned all these contexts may be assigned an interrogative interpretation if uttered with a rising intonation and/or in presence of a Q-marker *ne*; however, when an interrogative interpretation is achieved, these contexts are no longer veridical but non-veridical. Thus, the restricted distribution to non-veridical contexts only, proposed for the Mandarin indefinite is not violated.

Table 2: Distribution of *shenme* in spoken Mandarin[37]

As can be seen from the results above, *shenme* appears in non-veridical contexts in spoken Mandarin at more than 99% of the times. This shows that Mandarin speakers do indeed analyse the indefinite as a superweak NPI, banned from veridical contexts only.

4. Explaining *shenme* as a superweak NPI

From a distributional perspective, we showed in Section 3 that *shenme* is an NPI that is licensed by all non-veridical contexts. This section provides an explanation for why *shenme* is a superweak NPI, surviving in non-veridical contexts only (c.f. Zwarts 1993).

We here adopt Giannakidou (2002): NPIs that are subject to non-veridical licensing may become NPIs because they are referentially deficient. Before we show how this analysis explains *shenme*'s restricted distribution to non-veridical contexts only, we briefly demonstrate referentiality and semantic contexts that require obligatory referring and those that do not.

Referentiality can be informally understood as the ability to refer. Most NPs, for instance, exhibit this ability and can therefore be employed to refer. In examples given in (13), indefinite NPs (marked in italics) *a book*, *a car* and *a tree* refer to an entity in the world that meets the description of these indefinite NPs given by the context, i.e., on a round table that the speaker is looking for, with four cylinders that John is searching for, and in the Vondelpark that the parents of the speaking are looking for, respectively.

(13) a. I am looking for *a book* on a round table.
     b. John is searching for *a car* with four cylinders.
     c. My parents are looking for *a tree* in the Vondelpark.

In none of the examples above does the indefinite NP refer obligatory, as it might be the case that no entity existed such that it was a book or a car or a tree, meeting the description given by the context. However, referring is obligatory when the same NPs appear in the following sentences.

---

[37] More than 2000 utterances containing *shenme* were attested in the subcorpus *KouYu*; we randomly selected and analyzed 1000 of these for practical reasons. Out of these 1000 utterances, 66 contained *wei shenme* ('for a reason'), which partially overlaps with the target morphologically, but is syntactically and semantically different. The total number of utterances included in Table 2 is therefore 934.

(14) a.   I read *a book* yesterday.

    b.   John bought *a car* last year.

    c.   My parents planted *a tree* in 2010.

Utterances (14a) to (14b) necessarily presuppose the existence of at least one entity that meets its contextual description, i.e., read by the speaker yesterday, bought by John last year and planted by the parents of the speaker in 2010. Therefore, the indefinite NPs' *a book*, *a car* and *a tree* must obligatory refer. The examples given in (14) are all veridical expressions. In other kinds of veridical contexts, such as in perfectives (see (15a)), complement clauses of a factive verb (see (15b)), NPs must also refer.[38]

(15) a.   I have read *a book* since the last time I visited my parents.

    b.   I know that John bought *a car* last year.

On the contrary, the obligation to refer disappears when NPs appear in the following contexts: under the scope of negation in (16a) and (16b), in conditional clauses in (16c), in complement clauses of an intensional verb in (16d) and scoped over by a modal adverb in (16e).

(16) a.   I did not read *a book* yesterday.

    b.   Nobody read *a book* yesterday.

    c.   If John bought *a car* last year then he does not have to do it this year.

    d.   I guess that John bought *a car* last year.

    e.   Perhaps my parents planted *a tree* in 2010.

The contexts illustrated above are all examples of non-veridical contexts. As introduced in Section 2, non-veridical contexts are contexts that cannot entail the truth of an embedded proposition. This is why NPs uttered in such conditions do not necessarily presuppose the existence of a certain entity that meets the description provided by the context, as such explaining why non-veridical contexts do not require obligatory referring.

Given the generalization that veridical contexts involve obligatory reference whereas non-veridical contexts do not, the conclusion is that only indefinites that are able to refer

---

[38] As pointed out by a reviewer, perfectives in English are not necessarily referential contexts. In case of an intentional verb, for instance, perfectives may also be nonreferential: *I have looked for a solution* does not necessarily entail the existence of a solution.

may survive in contexts that presuppose existential import. Consequently, indefinites and/or quantifiers that are not able to refer cannot survive in contexts that presuppose existential import and may therefore appear in non-veridical contexts (Giannakidou 2002). On the basis of the distribution of the Mandarin indefinite restricted to non-veridical contexts only, we analyse *shenme* as an existential quantifier that lacks referentiality in its lexical semantics (see Li 1992 and Lin 1998 for a similar but not identical approach). Hence, it is *shenme*'s referential deficiency that restricts this indefinite to non-veridical contexts only that do not force it to refer only, explaining why *shenme* is a superweak NPI.

5. Acquiring shenme as a superweak NPI

Our analysis that *shenme* is a superweak NPI due to its referential deficiency explains why Mandarin speakers only use this indefinite in non-veridical contexts; but it also raises a learnability problem. S*henme*'s absence in veridical contexts such as those shown in (10) does not necessarily indicate its referential deficiency in the target grammar. It then appears that children would not be able to acquire that *shenme* can only appear in non-veridical contexts. We would thus expect children to overuse *shenme* in veridical contexts. However, without being confronted with any negative evidence, i.e., information about what is impossible and ungrammatical in a target grammar (Pinker 1995, among others),[39] it is impossible for children to unlearn the overgeneralized use of the NPI. As to understand how Mandarin children can acquire *shenme*'s non-referentiality based on positive evidence only, i.e., information about what is possible and grammatical in a target language (e.g., Pinker 1995), we follow Van der Wal (1996) in hypothesizing a conservative widening learning strategy in children's acquisition of the Mandarin NPI.

According to the conservative widening learning hypothesis, the acquisitional process of the NPI is analysed as having different developmental stages. In the first stage, children are assumed to establish the strictest possible analysis of the NPI based on limited input data available at the beginning of acquisition. While confronted with more input data falsifying the strict initial analysis children are assumed to weaken down the strict analysis and establish a reanalysis of the NPI. Such a weakening-down process

---

[39] For a discussion on the absence of negative evidence, the reader is referred to Brown & Hanlon (1970), Braine (1971), Baker & MacCarthy (1981), Pinker (1984, 1995, 2013), Gropen et al. (1991), Marcus et al. (1992), Marcus (1993), Cowie (1997). For a discussion on the unreliability of negative evidence, the reader is referred to Marcus (1993). For the fact that language learners do not seem to benefit from negative evidence if there is any, we refer the reader to McNeill (1966), Stromswold (1994), MacWhinney et al. (2002), among others.

according to language input continues until a reanalysis is achieved that explains all input data.

In order to provide empirical evidence for this learning strategy, we searched the CHILDES database (MacWhinney 2009) to investigate Mandarin children's acquisition of the superweak NPI. A total of 734 CHAT files of subcorpora *Beijing 2* (Tardif 1993, 1996), *Zhou 1* and *Zhou 2* (Zhou 2004) were analysed, covering spontaneous speech data of more than 40 monolingual Mandarin children aged between 1 and 5 years old. The procedure of our corpus research is as follows. First we divided all children into 4 different groups depending on their age at the time of recording: Group 1 (1 to 2 years old), Group 2 (2 to 3 years old), Group 3 (3 to 4 years old) and Group 4 (4 to 5 years old). After that we collected all utterances containing the target NPI *shenme* per age group. All the utterances of *shenme* were then categorized depending on their semantic property. Raw results are presented in the table below.

Semantic contexts	Group 1	Group 2	Group 3	Group 4
Anti-morphic	0	1	5	15
DE but not anti-additive	0	1	1	3
Interrogative sentences	0	53	170	335
Other non-veridical but not DE	0	0	0	30
Unclear	0	0	1	0
TOTAL	**0**	**55**	**177**	**386**

Table 3: Distribution of *shenme* in child Mandarin

Looking at how the target NPI was distributed, we only found a significant difference between children older than the age of 4 and those below 4 years old (*p*=.000, df=6) (see Table 4). Moreover, we found that the contributor to this significant effect is the emergence of non-veridical contexts that are neither interrogative nor DE, forming a new type of licensing contexts for the target NPI. Whereas Mandarin children below the age of 4 are only able to use the target NPI in interrogative sentences (more than 99% of the times), their older counterparts are also capable of employing other kinds of non-veridical contexts that are not DE to license *shenme* (at approximately 8% of the times).

Semantic contexts	Early child Mandarin (Age < 4 years old)	Late child Mandarin (Age > 4 years old)
Interrogative sentences	99.13%	86.79%
Anti-morphic contexts	2.6%	3.89%
DE but not anti-additive contexts	0.87%	0.26%
Other non-veridical but not DE contexts	0	7.77%

Table 4: Distribution of *shenme* in early and late child Mandarin

We take the developmental pattern illustrated above to represent an analysing and a reanalysing process of Mandarin children in the acquisition of the NPI. Since we do not assume any inborn linguistic knowledge of *shenme* being lexically non-referential, we started by looking at language input in order to understand Mandarin children's initial step to acquire the target NPI. In child-directed Mandarin in the investigated subcorpora of CHILDES, we found that *shenme* appears in an interrogative sentence at a frequency of more than 97%. Given *shenme*'s overwhelming occurrences under the scope of an interrogative operator in the input, we hypothesize that Mandarin children start out with a narrow assumption of the target NPI being a WH-quantifier. However, this initial analysis by Mandarin children can be falsified by input evidence showing *shenme* in a non-interrogative sentence. In child-directed Mandarin, we observed that at approximately 3% of the times the NPI is used in a non-veridical context that is not interrogative. *Shenme*'s appearance in such non-veridical contexts is sufficiently infrequent. Nevertheless, it still poses a problem for children's strong analysis of *shenme* as a WH-quantifier, due to its inability to explain why *shenme* is also allowed to appear in non-interrogative non-veridical contexts. In order to explain this, children need to establish a less narrow reanalysis that is compatible with all input data. Given the fact that only non-referential existential quantifiers are subject to a restricted distribution of non-veridical contexts only, including both interrogative and non-interrogative sentences (cf. Giannakidou 2002), we hypothesize that *shenme* is reanalysed as exactly that; a non-referential existential quantifier that cannot give rise to any existential import (after Lin 1998), similar to WH-quantifiers, but is allowed to appear in non-interrogative non-veridical contexts as well. Based on the significant difference observed between the Mandarin children below the age of 4 and their older counterparts (Table 4), we further hypothesize that the reanalysis of *shenme* – being referentially deficient – is established shortly after the age of 4.

The analysis and the reanalysis sketched above explain the developmental pattern attested in this corpus research. The fact that *shenme* only appears in interrogative sentences in early child Mandarin is understood by children's strict assumption of this

indefinite as a WH-quantifier, at ages younger than 4: the analysis. The broader distribution of *shenme* in a variety of non-veridical contexts including interrogative sentences in late child Mandarin is accounted for by the weaker reanalysis of *shenme* being referentially deficient after the age of 4: the reanalysis. Moreover, the analysing and the reanalysing processes provide evidence for the conservative widening learning strategy in Mandarin children's acquisition of the superweak NPI. First, in both the analysing and the reanalysing process, Mandarin children make use of positive evidence only. Second, the acquisitional pathway of *shenme* exhibits a clear widening development, as the NPI is distributed in a broader set of contexts in late child Mandarin than in earlier stages. Finally, Mandarin children – regardless of their age – do not overuse *shenme* in veridical contexts since we did not attest any overgeneralization errors of the NPI. We therefore conclude that Mandarin children acquire *shenme* as a non-referential superweak NPI via the conservative widening learning strategy.

6. Summary

In his 1995-paper, Zwarts proposed that non-veridicality is the logico-semantic property that licenses the weakest type of NPIs. Although the existing body of literature has hardly reported any polarity items that are systematically licensed in all non-veridical contexts, the current paper presents a prototypical superweak NPI that indeed exhibits such a distribution, providing crucial empirical evidence for Zwarts' proposal of almost twenty years ago.

We start by introducing NPIs and non-veridical environments. By examining the distribution of the Mandarin indefinite *shenme* in the Chinese literature, we conclude that *shenme* is a superweak NPI, allowed to appear in non-veridical contexts only. Our data collected in the PKU-CCL Corpora confirm this as well, since virtually all contexts containing the target NPI were non-veridical. We then present an explanation for why *shenme* has become an NPI of the weakest type, systematically banned from veridical contexts only. Following Giannakidou (2002), *shenme* is analysed as a lexically deficient indefinite that cannot refer on its own and therefore may only appear in non-veridical contexts that do not presuppose any existential import. The acquisition of the Mandarin NPI is examined by means of an intensive search in the CHILDES database. The child data show that the acquisition of *shenme* exhibits an analysing and a reanalysing process. After an initial narrow assumption of *shenme* being a WH-quantifier, which is a specific type of non-referential quantifier, Mandarin children reanalyse the target NPI more generally as an existential quantifier that lacks referentiality.

Our treatment of the Mandarin indefinite that states that it lacks referentiality explains why *shenme* is a superweak NPI restricted to non-veridical contexts only. However, it raises several further questions as well. The first question concerns a series of quantifiers in Mandarin Chinese that are analysed as WH-terms according to traditional Chinese grammar: *shei* ('a person'), *weishenme* ('for a reason'), *nali* ('a place'), etc. Similar to *shenme*, these so-called WH-terms can also appear in non-interrogative but still non-veridical contexts (Huang 1982, Lin 2011). If the distribution of all these quantifiers is indeed also restricted to non-veridical contexts, we can generalize our NPI-analysis of *shenme* to these quantifiers by adopting Zwarts' notion of non-veridicality. Moreover, the acquisition of this series of superweak NPIs in Mandarin Chinese will also be driven by the notion of non-veridicality – assuming a similar distributional pattern as that of *shenme* in the language input.

Secondly, our proposal motivates a typological investigation of NPIs of the weakest type. Mandarin Chinese is a WH-in-situ language; interrogative sentences are therefore not syntactically marked by WH-movement as in Dutch or English, for instance. Ni (2005) and Zhou (2010), among others, list several criteria to distinguish an interrogative sentence from its non-interrogative counterparts in Mandarin Chinese. Apart from two prosodic requirements,[40] an important grammatical property of interrogative sentences in Mandarin Chinese is the presence of an overt Q-marker in the sentence-final position. However, as Ni (2005) points out, this Q-marker may also be covertly present. This leads to two possible resolutions of a sentence containing *shenme* but without an overt Q-marker. Speakers either assign this sentence as interrogative by assuming a covertly present Q-marker; or they assign this sentence as non-interrogative by their analysis of *shenme* as a superweak NPI. Since Chinese is presumably not unique in this sense we may expect other WH-in-situ languages that do not require an overtly present Q-marker to exhibit superweak NPIs for the same reason of non-referentiality, similar to *shenme*; but we leave this for further exploration.

A third topic for further exploration is related to the current methodology. We executed a corpus study. But because corpus research restricts our observation to children's production only, which does not necessarily indicate what children can or will produce, another approach of interest for further research is to confirm the widening learning pathway attested here in an experimental setting. By manipulating *shenme's* appearance in different types of non-veridical contexts, e.g., interrogative sentences, under the scope of negation, modal contexts, we executed a sentence repetition task with

---

[40] The two prosodic requirements for interrogative sentences in Mandarin Chinese are a rising intonation at sentence level and a main sentential stress on a WH-quantifier in a traditional sense.

monolingual children aged between approximately 3 and 5 years old (authors in prep.). Preliminary results appear to confirm our corpus findings that the acquisition of this Mandarin NPI exhibits an initial interrogative assumption. A detailed description of our data and a discussion of our experimental results are part of our further research.

7. References

Baker, C. L., & McCarthy, J. J. 1981. *The logical problem of language acquisition*, Cambridge, MA: MIT Press.

Braine, M. D. 1971. 'On two types of models of the internalization of grammars'. *The Ontogenesis of Grammar*: 153-186.

Brown, R., & Hanlon, C. 1970. 'Derivational complexity and order of acquisition in child speech'. *Cognition and the Development of Language. New York: Wiley, 8*

Cheng, L. L-S. 1994. 'Wh-words as polarity items'. *Chinese Languages and Linguistics,* 2: 615-640.

Cheng, L. L-S. 1995. 'On *dou*-quantification'. *Journal of East Asian Linguistics,* 4:197-234.

Cheng, L. L-S & Huang, J. C-T. 1996. 'Two Types of Donkey Sentences'. *Natural Language Semantics*, 4: 121-163.

Cowie, F. 1997. 'The logical problem of language acquisition'. *Synthese, 111*(1): 17-51.

Ding Sh-Sh. 1961. Xiandai Hanyu Yufa JiangHua (*Speeches on modern Chinese grammar*). Beijing: Commercial Press.

Gropen, J., Pinker, S., Hollander, M., & Goldberg, R. (1991). 'Affectedness and direct objects: The role of lexical semantics in the acquisition of verb argument structure'. *Cognition, 41*(1): 153-195.

Giannakidou, A. 1998. *The landscape of polarity items*. Amsterdam: John Benjamins

Giannakidou, A. 1999. 'Affective Dependencies'. *Linguistics and Philosophy,* 22: 367-421.

Giannakidou, A. 2002. 'Licensing and sensitivity in polarity items: from downward entailment to (non)veridicality'. In M. Andronis, E. Debenport, A. Pycha & K. Yoshimura (eds.) *CLS* 38: 29-54.

Giannakidou, A. 2010. 'Negative and positive polarity items: licensing, compositionality and variation.' In C. Maienborn, K. von Heusinger and P. Portner (eds.) *Semantics: An International Handbook of Natural Language Meaning*. Berlin: Mouton de Gruyter.

Hoeksema, J. 1999. 'Aantekeningen bij ooit, deel 2: De opkomst van niet-polair ooit'. *TABU: Bulletin Voor Taalwetenschap, 29*: 147-172.

Hoeksema, J. 2010. 'Dutch ENIG: from nonveridicality to downward entailment'. *Natural*

Language & Linguistic Theory, *28*(4): 837-859.

Hoeksema, J. 2012. 'On the natural history of negative polarity items'. *Linguistic Analysis,* 38 (1–2): 3–33.

Huang, J. C-T. 1982. *Logical Relations in Chinese and the Theory of Grammar.* PhD Dissertation, MIT.

Ladusaw, W. 1979. *Polarity Sensitivity as Inherent Scope Relations*. Unpublished PhD Dissertation, University of Texas. Distr. by the Indiana University Linguistics Club.

Ladusaw, W. 1996. 'Negation and Polarity Items'. In Lappin, S. (ed.) *The Handbook of Contemporary Semantic Theory*: 321-341. Oxford and Malden: Blackwell Publishers.

Li, J-X. 1924. Xinzhu Guoyu Wenfa (*The Grammar of Mandarin Chinese: New Edition*). Beijing: Commercial Press.

Li, A. Y-H. 1992. 'Indefinite Wh in Mandarin Chinese'. *Journal of East Asian Linguistics* 1: 125–155.

Li, Y-M. 2002. Yufa Yanjiulu (*Record on Grammar Research*). Beijing: Commercial Press.

Lin, J-W. 1996. *Polarity licensing and wh-phrase quantification in Chinese*. PhD Dissertation, University of Massachusetts, Amherst.

Lin, J-W. 1998. 'On existential polarity wh-phrases in Chinese'. *Journal of East Asian Linguistics* 7: 219-255.

Lin, J. 2011. *Mandarin shenme as a negative polarity item: Distribution, interpretation and acquisition.* (Unpublished rMA thesis), University of Amsterdam.

Lü, SH-X. 1982. Zhongguo Wenfa Yaolue (*Essentials of Chinese Grammar*). Beijing: Commercial Press.

MacWhinney, B. 2009. 'The childes project'. *Tools for Analyzing talk–electronic Edition, 2*

MacWhinney, B., Givón, T., & Malle, B. 2002. 'The gradual emergence of language'. *The Evolution of Language Out of Pre-Language*: 231-263.

Marcus, G. F. 1993. 'Negative evidence in language acquisition'. *Cognition, 46*(1): 53-85.

Marcus, G. F., Pinker, S., Ullman, M., Hollander, M., Rosen, T. J., Xu, F., & Clahsen, H. 1992. Overregularization in language acquisition'. *Monographs of the Society for Research in Child Development*, i-178.

McNeill, D. 1966. 'Developmental psycholinguistics'. *The Genesis of Language*: 15-84.

Ni, L. 2005. Xiandai Hanyu Yiwen Daici de Jiben Yuyi Fenxi (*Basic semantic study on wh-words*). *The Northern Forum,* 192: 55-57.

Pinker, S. 1984. *Language learnability and language development*. Cambridge, MA: Harvard University Press.

Pinker, S. 1995. 'Language acquisition'. *Language: An Invitation to Cognitive Science, 1*: 135-182.

Pinker, S. 2013. *Learnability and cognition: The acquisition of argument structure*. MIT press.

Stromswold, K. 1994. *What a mute child tells us about language*. Unpublished Manuscript, Rutgers University.

Tardif, T. 1993. *Adult-to-child speech and language acquisition in Mandarin Chinese*. Unpublished PhD Dissertation, Yale University.

Tardif, T. 1996. 'Nouns are not always learned before verbs: Evidence from Mandarin speakers' early vocabularies'. *Developmental Psychology, 32*(3): 492.

Van der Wal, S. 1996. *Negative Polarity Items & Negation: Tandem Acquisition*. PhD Dissertation, University of Groningen.

Van der Wouden, T. 1994. 'Polarity and 'illogical negation'. In *Dynamics, Polarity and Quantification.* 17-45. Stanford, California: CSLI Publications.

Van der Wouden, T. 1997. *Negative Contexts: Collocation, Polarity and Multiple Negation*. London: Routledge.

Xie, Z. 2007. 'Nonveridicality and existential polarity wh-phrases in Mandarin'. In *Proceedings from the Annual Meeting of the Chicago Linguistic Society*. Vol. 43 (2): 121-135. Chicago Linguistic Society.

Zhao, Y-R. 1979. Hanyu Kouyu Yufa (*A Grammar of Spoken Chinese*). Beijing: Commercial Press.

Zhou, J. 2004. *The Zhou 1 corpus*, the CHILDES database.

Zhou, J. 2004. *The Zhou 2 corpus*, the CHILDES database.

Zhou, P. 2010. *Interface conditions in child language: A view from Mandarin Chinese*. PhD Dissertation, Macquarie University.

Zwarts, F. 1981. 'Negatief polaire uitdrukkingen I'. *Glot, 4*(1): 35-132.

Zwarts, F. 1993. 'Three types of polarity'. Ms. Also appeared as Zwarts 1998 in F. Hamm and E. Hinrichs (eds.) *Plural quantification.* Dordrecht: Kluwer.

Zwarts, F. 1995. 'Nonveridical contexts'. *Linguistic Analysis 25*, 3-4: 286-312.

8. **Appendix 1: Shenme appearing in different linguistic contexts given in Table 1**

(1) under the scope of a negative marker

Ta   mei you   *shenme* hua   xiang   he ni   shuo   .

s/he NEG have   *shenme* word   want   with you say

'S/he does not want to say anything to you.'

(2) in a restrictive clause of a universal quantifier

Ta   *shenme* hua   dou xiang   he ni   shuo   .

s/he *shenme* word   all   want   with you say

'S/he wants to tell you everything.'

(3) in a conditional clause

Ruguo ta   you   *shenme* hua   xiang   he ni   shuo,

if   s/he have   *shenme* word   want   with you say

ta   jiu hui gei ni   dadianhua   .

s/he then will to   you call

'If s/he has something to tell you, s/he will then call you.'

(4) in a *BEFORE*-clause

Zai   ta   xiangdao *shenme* fangfa   zhiqian, yinggai

at   s/he think of *shenme* solution before   should

he   ni   haohao   shangliang   yixia

with   you well   discuss   while

'Before s/he thinks of any solution, s/he should have a good talk with you.'

(7) in a matrix interrogative sentence

Ta   he ni   shuo   guo *shenme* (ne) ?

s/he with you say   PRF *shenme* Q-marker

'What did s/he tell you?'

(8) in an embedded interrogative sentence

Wo   xiang   zhidao ta   he ni   shuo   guo *shenme* .

I   want   know   s/he with you say   PRF *shenme*

'I want to know what s/he told you'

(9) imperfectives: futural aspect

Wo   mingtian   qu shichang   gei wo ma   mai

I   tomorrow   go market   for I   mother buy

dian   *shenme* chi de   .

CL   *shenme* eat PAR

'I will go to the market tomorrow to buy something to eat for my mother.'

(10)   imperfectives: habitual aspect

Ta   zongshi baoyuan *shenme* buhao   *shenme* budui   .
s/he   always   complain*shenmei* bad   *shenme* wrong
'S/he always complains that something is wrong.'

(11)   imperfectives: progessive aspect

Ta   xianzai zhengzai kan   *shenme* dianshiju   ne .
s/he   now   PRG   watch   *shenme* television program   PAR
'S/he is now watching a television program.'

(12)   in an imperative

 Kuai   qu   mai dian *shenme* zhixue   de yao !
quickly   go   buy   CL   *shenme* haemostatic   PAR medicine
'Quickly go buy some haemostatic medicine!'

(13)   in a modal context

Ta   haoxiang zai   kan   *shenme* dianshi jiemu .
s/he   probably at   watch   *shenme* television program
'S/he is probably watching some television program.'

(14)   in a matrix polar question

Ni   hai xiang   mai *shenme* yao   ma ?
you   still want   buy *shenme* medicineQ-marker
'Do you still want to buy some/any medicine?'

(15)   in an embedded polar question

Wo   tebie   xiang   zhidao ta   shifou   gei wo mai le
I   very   want   know   s/he whether   for   I   buy PRF
shenme   liwu .
shenme   present
'I really want to know whether s/he has bought me a present.'

(16)   in a complement clause of an intensional verb

Wo   cai   ta yijing   gei ni   mai le   *shenme* liwu .
I   guess   s/he already for you buy PRF *shenme* present
'I guess that s/he has already bought you a present.'

(17)   under the scope of inference *le*

Shenme   lingjian er   huai   le   .
shenme   spare part   broken   INF
'Some part appears to be broken.'

9. Appendix 2: Abbreviations

CL	classifier
COP	copula
DE	downward entailing
DEM	demonstrative
FUT	future tense marker
GEN	genitive marker
HAB	habitual marker
INF	inference marker
MOD	modifier
NEG	negation
NP	noun phrase
NPI	Negative Polarity Item
PAR	particle
PRF	perfective marker
PRG	progressive marker
Q-marker	question marker