A Numeral Oddity

Luca Gasparri, IJN Paris | December 17, 2018

(penultimate draft, final version published in the *Journal of Semantics*)

Abstract

Natural language appears to allow the ascription of properties of numeral symbols to the denotation of number referring phrases. The paper describes the phenomenon and presents two alternative explanations for why it obtains. One combining an intuitive semantics for number referring phrases and a predicate-shifting mechanism, the other assigning number referring phrases a structured denotation consisting of two parts: a mathematical object (the number) and a contextually determined numeral symbol. Some preliminary observations in favor of the second analysis are offered.

1. Introduction

There is considerable debate on the semantics and pragmatics of number words and constructions. The aim of this short contribution is to draw attention on a familiar, yet surprising pattern found in everyday number talk, and use it to consider an unorthodox hypothesis about the semantics of number referring phrases. The plan is as follows. Section 2 will outline two rival approaches to the denotation of bare numerals and describe a numeral phrase unanimously believed to be number referring (i.e., to evaluate to a number). Section 3 will describe a pool of felicitous sentences which appear to ascribe properties of numeral symbols to the denotation of the number referring phrase introduced in Section 2. Section 4 will consider explaining the phenomenon by way of two alternative hypotheses. One on which number referring phrases evaluate to numbers and the odd pattern is licensed by predicate-shifting mechanisms. Another on which number referring phrases have a structured denotation consisting of two parts: a mathematical object (the number) and a contextually determined numeral symbol. Section 5 will make some preliminary observations in favor of the second analysis. Section 6 will conclude.

2. Reference to Numbers

The current landscape in semantics and philosophy of language distinguishes what we can call "referentialist" and "non-referentialist" theories of the denotation of bare numerals (cf., a.o., Landman 2003; 2004; Hofweber 2005; 2016; Geurts 2006; Ionin and Matushansky 2006; Moltmann 2013; Kennedy 2015; Rothstein 2017). Consider (1a-c).

1) a. Fifteen is divisible.
   b. Fifteen children went to the movies.
   c. The fifteen children went to the movies.

For referentialists, who recast the classical Fregean analysis, the basic meaning of bare numerals is the one found in (1a), where "fifteen" is a singular term denoting the number `15`.[1] The readings found in (1b) and (1c) are derived from this basic meaning. For non-referentialists, the opposite is the case. Depending on the specific brand of non-referentialism involved, bare numerals are believed to encode as their basic meaning either the reading found in (1b), where the numeral "fifteen" operates as a quantificational determiner, or the reading found in (1c), where the numeral "fifteen" operates as a cardinality predicate. On these views, bare numerals have a basic predicative or quantificational meaning, and the number-referring reading found in (1a) is derived from this basic meaning through appropriate shifts.

Notwithstanding the debate on the basic denotation of bare numerals, there is thus consensus that bare numerals and complex numeral phrases behave as number referring (i.e., evaluate to numbers) when positioned in appropriate environments. They might do so "directly", as referentialists would say, because the environment where the numeral phrase occurs allows the numeral to contribute its basic denotation (a number). Or they might do so "indirectly", as non-referentialists would say, because the environment where the numeral phrase occurs forces a shift of the basic denotation of the numeral (e.g., from a cardinality property to a number). For example, there is consensus that appositive phrases combining a definite determiner, the 'number' sortal and a number word evaluate to numbers in environments like (2), where φ expresses a number property such as being divisible, being odd, or being prime (cf. Moltmann 2017; Snyder 2017).

2) [*definite D*, 'number' *sortal*, *number word*] is φ.

1 Notational caveat: to avoid ambiguities, numbers will be notated in code font, whereas numeral symbols will be notated within single quotes. Thus, `15`, `600`, and `1224` will designate the respective mathematical objects (the numbers), whereas '15', '600', and '1224' will designate the respective numeral symbols.

Hence, there is consensus that the numeral phrase "the number fifteen" in (3) is number referring and evaluates to the number 15.

3) The number fifteen is odd.
   ⟦The number fifteen⟧ = 15

3. The Oddity

We all agree that numbers and numeral symbols are different (kinds of) entities. Numbers are abstract mathematical objects and numeral symbols are tools to designate numbers within a numeral system. For example, the numeral symbols '15' and '1111' designate 15 in decimal and binary, respectively. Likewise, '11' can be interpreted as the binary symbol for 3 or as the decimal symbol for 11. Now consider (4).[2]

4) ✓ The number fifteen is odd and has two digits.

(4) opens with a variant of the appositive construction found in (3), which we expect to evaluate to 15. The sentence then ascribes two properties to the denotation of "the number fifteen": the property of not being divisible by 2 and the property of having two digits. But this is strange. While the property of not being divisible by 2 can be predicated of the number 15, "having two digits" is not a possible property of 15; it's a property of the numeral symbol '15'. On account of this, (4) should be perceived as infelicitous (a category mistake) and guilty of ascribing an object a property that can be instantiated only by the expression or symbol that names it. Witness (5a-b).

5) a. # The city of Boston is densely populated and has six letters.
   b. # My flatmate Andrew is a great cook and has two syllables.

Yet, unlike (5a-b), (4) sounds perfectly felicitous. How is this possible?

2 The judgments reported in the paper have been reviewed at various stages of the preparation of the manuscript with two native informants. Preliminary tests with other informants indicate that the contrasts I will consider are not restricted to English, and can be replicated in other languages (e.g., French, German, Italian).

One reaction would be to dismiss the question and argue that the reason why (4) succeeds is simply that we are confused about the distinction between `15` and '15'. From a linguistic standpoint, there is nothing interesting to say about (4) besides that it shows we are bad at keeping track of the distinction between numbers and numerals. It is true that (4) is *perceived* as felicitous, unlike (5a-b). However, (4) is *objectively* just as bad as (5a-b). We perceive a contrast between (4) and (5a-b) only because we cannot help conflating numeral symbols with the objects they are designed to pick out within the numeral system we are most familiar with. But deep down there is no actual contrast in felicity between (4) and (5a-b). Hence, nothing we should worry about or bother trying to explain.[3]

The objection raises a fairly natural concern. But I think it should be resisted, for two reasons. The first is that the dichotomy between "real" and "surface" felicity pressed by the objection introduces an unnatural divide between a class of "objectively grounded" judgments of acceptability or felicity we should try to account for versus a class of "unreliable" judgments that do not constitute linguistic explananda. Assuming that the business of linguistic theory is to account for the distribution of speakers' intuitions about natural language sentences (no matter how consistent with our prior assumptions such intuitions might be), perceived felicity just *is* felicity. Hence, the fact that listeners do perceive a contrast between (4) and (5a-b) should be enough to make us wonder why that happens and try to answer the question with the instruments of linguistic theory.

The second reason is that one can accept that the source of the felicity of (4) is the conflation of numbers and numeral symbols, *and* at the same time ask if, and how, this conflation is reflected in the linguistic properties of the expressions we use to talk about numbers. Notice, in this regard, two things. First, the odd pattern is accepted in sentences ascribing numbers (e.g., `1000003`) the properties of numeral symbols (e.g., '1,000,003'), but is consistently rejected in sentences ascribing numbers (e.g., `1000003`) the properties of numeral words (e.g., 'one million and three'), irrespective of whether the numeral in the number referring phrase is expressed in digits or characters. See (6a-d) and, for further illustration, notice how "the number twelve" can pick out the symbol '12' in (7b) but cannot replace the word 'twelve' in (7a).[4]

6) a. ✓ The number one million and three is odd, prime, and starts with a '1'.
   b. ?? The number one million and three is odd, prime, and starts with an 'o'.

3 Thanks to an anonymous reviewer for pressing me on this.

4 Thanks to Jeremy Kuhn for (7a-b).

c. ✓ How many digits does the number 1,000,003 have?

d. ?? How many consonants does the number 1,000,003 have?

7) a. ?? The salesperson asked Mary how many apples she wanted. Mary said the number twelve.

b. ✓ Noticing that John was deaf, the salesperson used a piece of paper to ask him how many apples he wanted. John wrote the number twelve.[5]

Second, the contrast between (4) and (5a-b) seems to be unaffected by numerical literacy and literacy in linguistic theory. Based on the judgments I've gathered, (4) sounds felicitous not only to the ears of speakers unaccustomed to the distinction between numbers and numeral symbols, but also to listeners educated in mathematics and linguistic theory, and in any case well aware of the distinction between numbers and numeral symbols. The same goes for (8a-c).

8) a. ✓ The number one divides all numbers and has a straight shape.

b. ✓ The number two is even, prime, and resembles a swan's neck.

c. ✓ The number three is an approximation of π and contains as many lines as it represents.

Even upon noticing their technical inconsistency, listeners keep perceiving (4) and (8a-c) as viable sentences. The judgments persist in face of the attempt to explicitly recall that numbers (1, 2, 3) are not (concatenations of) digits ('1', '2', '3'). Which gives another reason to bet on the hypothesis that

5 A reviewer points out that variants like "✓ The number 33 is difficult to pronounce" or "✓ The number 133 is often misspelled" appear acceptable. Assuming that "being difficult to pronounce" and "being often misspelled" can only be properties of numeral words (e.g., 'thirty-three'), this would seem to indicate that the oddity does extend to numeral words. I don't think, however, that this invalidates the observation. For one thing, the contrasts in (6a-d) and (7a-b) remain robust. Second, I would argue that the felicity of "The number 33 is difficult to pronounce" and "The number 133 is hard to spell" is due, rather than to any special property of the numeral phrases they contain, to the more general fact that the phrases "being difficult to pronounce" and "being often misspelled", albeit selecting primarily linguistic objects (words), can take non-linguistic objects as subject arguments. So, it is true that "✓ The number 133 has four divisors and is hard to spell" is better than "?? The number 133 has four divisors and many consonants", but so is, e.g., "✓ Pittsburgh has a continental climate and is often misspelled" versus "?? Pittsburgh has a continental climate and many consonants". Thus, the felicity of variants like "The number 33 is difficult to pronounce" is not necessarily revealing of any surprising feature specific to number referring phrases.

the contrast between (4) and (5a-b) may teach us something about the way in which the distinction between numbers and numeral symbols is reflected in natural language.

So let us try to explain why (4) is felicitous despite the odd conjunction of predicates it contains. I see two main options to do that. The first option is to stick to the assumption that "the number fifteen" evaluates to 15 and argue that the sentence is felicitous because the predicate "has two digits" shifts in meaning. The second option is to construe the success of the predication as having to do with the semantics of the numeral phrase opening the sentence and, hence, as symptomatic of the need to reconsider the assumption that the "the number fifteen" evaluates to 15 *simpliciter*. Let us explore these two options in more detail.

4. Two Explanations

As mentioned, the first strategy would be to explain the success of (4) by combining the assumption that "the number fifteen" denotes 15 *simpliciter* with the appropriate shift in the predicate "has two digits". Consider (9).

9) [*Context: listeners have to guess what cities were visited by Rebecca last summer*]
✓ Every city Rebecca visited last summer is in India and has six letters.

How come that (9) is felicitous even if cities cannot possibly be literally predicated of the property of "having six letters"? Presumably because, partly upon the pressure from the goals at stake in the context (guessing a series of city names), the predicate "having six letters" is applied to "every city" after being implicitly shifted into the more complex predicate "being designated by a six-letter name". Since, one could argue, this seems everything we need to figure out why (9) work so well, we can extend the same explanation to (4). On this line of thinking, the felicity of (4) is perfectly consistent with the assumption that "the number fifteen" in (4) evaluates to 15 *simpliciter*. The felicity of the example is explained by positing that the interpretive machinery shifts the literal meaning of the numeral-selecting predicate "having two digits" into "being associated by convention to a double-digit numeral" (or something in that vicinity), and ascribes 15 the so shifted predicate. Building on this analysis, we would seem able to capture the felicity of (4) while not having to embark on any unorthodox speculation about the denotation of "the number fifteen".

Suppose instead we react to the oddity by conjecturing that the felicity of (4) casts doubt on the idea that the "the number fifteen" in (4) ranges semantically over an entity of plain and simple type NUMBER. This conjecture can be backed up by observing that the odd conjunction of predicates featuring in (4) has a close analogous in co-predication examples like (10).

10) ✓ The book on Pam's desk weighs five pounds and was originally written in Russian.

Sentences like (10) have received considerable attention in recent years (e.g., Asher 2011; Luo 2012; Gotham 2017). Most available analyses of co-predication cases claim that sentences such as (10) are acceptable because "the book" evaluates to a structured semantic matrix whose specifications or facets combine with the basic meaning of the ensuing predicates. So the co-predication of "weighs five pounds" and "was originally written in Russian" to "the book" is felicitous because the "the book" in (10) evaluates to a PHYSICAL-OBJECT • INFORMATIONAL-CONTENT object which simultaneously designates a physical entity with a certain weight and a collection of sentences in a natural language.

The parallel suggests another simple way of explaining the felicity of (4): positing that "the number fifteen" in (4) has a structured denotation describable as a NUMBER • NUMERAL-SYMBOL matrix. Basically, we should reconsider the assumption that number referring phrases evaluate to numbers *simpliciter* and allow them to range over structured semantic objects consisting of two components: a mathematical object (e.g., 15) and a contextually determined numeral symbol (e.g., '15') which names the number in the numeral system assumed by the speakers at the production of the sentence (in ordinary English, the default is the Arabic decimal system). For illustration, see (11a-b) and (12a-b).

11) [*Context: a conversation among ordinary speakers of English*]
a. ✓ The number fifteen has two digits. (*true*)
b. ✓ The number fifteen has four digits (*false*)
⟦The number fifteen⟧ = 15 • '15'

12) [*Context: a conversation among binary enthusiasts*]
a. ✓ The number fifteen has two digits. (*false*)

b. ✓ The number fifteen has four digits. (*true*)

⟦The number fifteen⟧ = 15 • '1111'

5. Points for the Unorthodox Moral

Although both analyses sound workable, the one positing a NUMBER • NUMERAL-SYMBOL denotation is probably going to inspire particular caution. While the predicate-shifting explanation is semantically conservative, in the sense that it produces an explanation for the oddity without asking us to renegotiate our assumptions about the denotation of number referring phrases, the appeal to NUMBER • NUMERAL-SYMBOL objects asks us to depart from our initial view of what number referring phrases denote. The costs are comparatively higher, and so is the amount of linguistic evidence we would ideally want to gather before concluding that the option deserves consideration. With this in mind, let me provide a few exploratory remarks in favor of structured-denotation analysis.

To start, note that besides providing a straightforward explanation for the success of (4), the hypothesis generalizes smoothly to neighboring cases. For example, by allowing number referring phrases to contribute semantically a NUMERAL-SYMBOL component, it immediately foresees equal felicity for (13a-b) contra (14a-b), where "the city of Boston" can be safely assumed not to contribute a CITY-NAME component.

13) a. ✓ The number two hundred and thirty is larger than the number fifteen.
b. ✓ The number two hundred and thirty is longer than '15'.

14) a. ✓ The city of Boston is larger than Somerville.
b. # The city of Boston is shorter than 'Somerville'.

It also correlates nicely with evidence that numeral-symbol properties can be recruited to function as pointers to numbers in sentences like (15a-b), whereas the mechanism does not extend to (15c), which uttered out of the blue sounds very odd.

15) a. ✓ In 1979, the US had a two-digit inflation.
b. ✓ Jane earned a six-figure salary last year.
c. ?? Jane has a two-syllable flatmate.

Finally, by allowing number referring expressions to range over objects with a NUMERAL-SYMBOL component, the hypothesis seems to capture the contrast in (16).

16) a. [*Context: the speaker is correcting a child attempting to write* '3' *but producing instead an inscription shaped as* 'Ɛ']
✓ That's not the number three.
b. [*Context: the speaker is correcting a child attempting to write* 'Denver' *but producing instead an inscription shaped as* 'Dənvər']
That's not the city of Denver.

Coming now to potential points of advantage of the structured-denotation analysis over the one appealing to a shift in meaning of the numeral-selecting predicate, recall the parallel between (4) and (9) encouraged by the predicate-shifting explanation.

4) ✓ The number fifteen is odd and has two digits.
9) [*Context: listeners have to guess what cities were visited by Rebecca last summer*]
✓ Every city Rebecca visited last summer is in India and has six letters.

Now, it is true that (9) sounds felicitous in the particular context specified for the sentence, and that whenever the aim of the communicative exchange is the identification of an object via a consideration of the properties of the expression that names it, the parallel holds.[6] See (17a-b).

17) a. [*Context: listeners have to guess a series of numbers*]

6 For the sake of completeness, it should be mentioned that contexts à la (9) also improve the infelicitous examples in (6a-d). We have seen that sentences ascribing numbers the properties of numeral words tend to be rejected. Interestingly, in situations where listeners have to guess numbers based on properties of numeral words, context seems to repair the pattern: e.g., "✓ Tell me what numbers between ten and twenty have three syllables". But again, the repair is hardly revealing of anything specific to the semantics or the pragmatics of number referring phrases, since it's available across the board (e.g., "[*Context: talking to a child at the zoo*] ✓ Tell me which of these animals has five letters"), and does not invalidate the initial contrast: outside specific contexts of this sort, only the predication of properties of numeral symbols seems allowed.

✓ Tell me what numbers between five and fifteen have two digits.

b. [*Context: listeners have to guess a series of cities*]

✓ Tell me what cities in Massachusetts have eight letters.

However, as soon as the goal at stake stops being the identification of an object based on the properties of the word or symbol that names it, the parallel breaks down. See (18a-c) versus (19a-c).

18) [*Context: the speaker is talking about the properties of a number*]

a. ✓ The number one thousand is a power of ten and contains three '0's.

b. ✓ The number fifteen is a divisor of thirty and has two digits.

c. ✓ How many digits are there in the number 632?

19) [*Context: the speaker is talking about the properties of a city*]

a. # The city of Boston is rich and contains two 'o's.

b. # The city of Kyoto covers 320 square miles and has two syllables.

c. # How many syllables are there in the city of Shanghai?

The contrast casts doubt on the idea that the felicity of (18a-c) (and (4)) is ultimately a matter of predicative shifts, for the following reason. Supposing that the crux of felicity lies in the shift in meaning of the appropriate predicate, the fan of the predicate-shifting explanation needs to explain why the shift is available in (18a-c) but unavailable in (19a-c), despite the fact that (19a-c) are produced in a context and built on a structural blueprint which resemble so closely the ones found in (18a-c). By contrast, the structured-denotation hypothesis already comes with the resources required to produce an explanation for the pattern. Namely, there is no mechanism of predicative shift in (18a-c), nor, retrospectively, in (4): the phenomenon really has to do with what the numeral phrases involved denote. What makes (18a-c) acceptable, as opposed to (19a-c), is the fact that the numeral phrases featuring in (18a-c) denote a NUMBER • NUMERAL-SYMBOL matrix (15 • '15') whose facets can selectively combine with the non-shifted meaning of the ensuing predicates. The contrasts in (20a-d) and (21a-d) also give support to the idea that "the number fifteen" evaluates to a semantic object featuring a NUMBER and a NUMERAL-SYMBOL component, as they show that both readings are available to

anaphora and that ellipsis is licensed. By contrast, it is unclear how the success of these examples might be derived by appealing to the machinery of shifts.

20) a. ✓ The number fifteen$_i$ is a simple integer. Its$_i$ divisors are one, three, five, and fifteen.
b. ✓ The city of Boston$_i$ is well-known. Its$_i$ history attracts many tourists.
c. ✓ The number fifteen$_i$ is a simple integer. But few children can write it$_i$ well.
d. ?? The city of Boston$_i$ is well-known. But few children can write it$_i$ well.

21) a. ✓ John wrote 'bazooka' and Mary the number twenty-five.
b. ?? John wrote 'bazooka' and Mary the city of Boston.
c. ✓ 'Aramara' reads the same backward as forward, and the number forty-four does so too.
d. ?? 'Honolulu' starts with a three-letter syllable, and the city of Boston does so too.

6. Going Forward

I have provided some preliminary observations in favor of the view that number referring phrases have a structured dot-type semantics combining a mathematical object (the number) and a contextually determined numeral symbol. Of course, the evidence considered in this short contribution is far from conclusive. But I think it gives us an initial case for the NUMBER • NUMERAL-SYMBOL analysis, one worth exploring in more depth. While the structured-denotation hypothesis produces a straightforward explanation for the intuitions we have reviewed, it is not immediately clear how such data might be accounted for based on the notion that, e.g., "the number fifteen" in (4) evaluates to `15` *simpliciter*. One possibility could be to assume the simple semantics, leave shifts out of the picture, and claim that the phenomenon obtains because number talk inevitably generates a strong salience effect for numeral symbols (unlike city talk, which allows reference to the properties of city names only in contexts where the name of the city matters to the question under discussion, à la (9)), and that this is the source of the oddity. But the hypothesis appears ill-suited to capture simple cases like (11a), and faces an uphill battle if measured against examples like (20c), which suggest that we are dealing with something deeper than salience. See also (22a-b), where a variant of the conjunction of predicates found in (4) is licensed even if `55` is introduced descriptively, (22c), which reproduces the pattern with a quantified expression, and (22d), which indicates that the phenomenon can be found in the denotation of cardinality.

22) a. ✓ The tenth Fibonacci number is odd and reads the same backward as forward.

b. ✓ The product of eleven and five is odd and reads the same backward as forward.

c. ✓ Every natural number has a successor and at least one digit.

d. ✓ The number of protons in an atom of caesium is odd and resembles a double 's'.

Should we be able to provide further evidence that the pervasiveness of the oddity is best explained under the complex semantics foreseen by the structured-denotation analysis, our current assumptions about the semantics of number referring phrases may deserve some fine-tuning.[7]

References

Asher, Nicholas (2011), *Lexical Meaning in Context: A Web of Words*. Cambridge University Press. Cambridge.

Geurts, Bart (2006), Take 'five'. In Svetlana Vogeleer and Liliane Tasmowski (eds.), *Non-Definiteness and Plurality*. Benjamins. Amsterdam. 311-329.

Gotham, Matthew (2017), Composing criteria of individuation in copredication. *Journal of Semantics* 34: 333-371.

Hofweber, Thomas (2005), Number determiners, numbers, and arithmetic. *Philosophical Review* 114: 179-225.

Hofweber, Thomas (2016), *Ontology and the Ambitions of Metaphysics*. Oxford University Press. Oxford.

Ionin, Tania, & Ora Matushansky (2006), The composition of complex cardinals. *Journal of Semantics* 23: 315-360.

Kennedy, Christopher (2015), A "de-Fregean" semantics (and neo-Gricean pragmatics) for modified and unmodified numerals. *Semantics and Pragmatics* 8, Article 10: 1-44.

Landman, Fred (2003), Predicate-argument mismatches and the adjectival theory of indefinites. In Martine Coene and Yves d'Hulst (eds.), *From NP to DP Volume 1: The Syntax and Semantics of Noun Phrases.* John Benjamins. Amsterdam. 211-237.

Landman, Fred (2004), *Indefinites and the Type of Sets*. Blackwell. Oxford.

Luo, Zhaohui (2012), Formal semantics in modern type theories with coercive subtyping. *Linguistics and Philosophy* 35: 491-513.

Moltmann, Friederike (2013), Reference to numbers in natural language. *Philosophical Studies* 162: 499-534.

Moltmann, Friederike (2017), Number words as number names. *Linguistics and Philosophy* 40: 331-345.

Rothstein, Susan (2017), *Semantics for Counting and Measuring*. Cambridge University Press. Cambridge.

Snyder, Eric (2017), Numbers and cardinalities: What's really wrong with the Easy Argument for numbers? *Linguistics and Philosophy* 40: 373-400.

7 I am grateful to Salvador Mascarenhas for discussion in preparation of the paper, and to Brian Buccola, Benjamin Spector, two anonymous reviewers, and the editor in chief, Rick Nouwen, for comments on the penultimate version of the manuscript. All errors are mine. This research was supported by grants ANR-17-EURE-0017 FrontCog and ANR-10-IDEX-0001-02 PSL*.