

Zero and triviality

Andreas Haida
ELSC, HUJI
Edmond J. Safra Campus
Jerusalem, 9190401
andreas.haida@mail.huji.ac.il

Tue Trinh
Leibniz-ZAS
Schützenstr. 18
10117 Berlin
tuetrinh@alum.mit.edu

Abstract This paper takes issue with Bylinina & Nouwen’s (2018) hypothesis that the numeral **zero** has the basic weak meaning of ‘zero or more.’ We argue, on the basis of empirical observation and theoretical consideration, that this hypothesis implies that exhaustification can circumvent L-triviality, but exhaustification cannot circumvent L-triviality. We also provide some experimental results to support our argument.

Keywords: numerals, modification, exhaustification, L-Triviality

1 Theses and antitheses

1.1 Exhaustification

The aim of this squib is to discuss some issues regarding the hypothesis about *zero* put forward by Bylinina & Nouwen (2018). These authors assume that numerals have the weak, ‘at least’ meaning as basic and the strengthened, ‘exactly’ meaning as derived by way of exhaustification, and propose that *zero* be considered no different from other numerals in this respect.¹ This is illustrated by (1a) and (1b).

- (1) a. Zero students smoked \Leftrightarrow
 $| \llbracket \text{students} \rrbracket \cap \llbracket \text{smoked} \rrbracket | \geq 0$
‘the number of students who smoked is zero or greater’
b. EXH [Zero students smoked] \Leftrightarrow
 $| \llbracket \text{students} \rrbracket \cap \llbracket \text{smoked} \rrbracket | \geq 0 \wedge | \llbracket \text{students} \rrbracket \cap \llbracket \text{smoked} \rrbracket | \not\geq 1$
‘the number of students who smoked is zero and not greater’

Before we go on, let us be explicit about the notion of exhaustification which we will adopt for the purpose of this discussion. First, [EXH ϕ] is true iff ϕ is true and

¹ For arguments that numerals have the ‘at least’ meaning as basic see Horn (1972; 1989); von Stechow & Heim (1997); Fox (2007b); Buccola & Haida (2019), among others.

every excludable alternative of ϕ is false. Second, an alternative of ϕ is a sentence derivable from ϕ by replacing certain scalar items in ϕ with their scale mates. And third, ψ is an excludable alternative of ϕ iff ψ is an alternative of ϕ , ψ is stronger than ϕ , and $\phi \wedge \neg\psi$ does not entail any other alternative of ϕ .² These premises, together with the standard view that numerals form a scale (cf. Horn 1989), will derive (1b).³

Coming back now to the sentences in (1), we can see not only that (1a) is entailed by (1b), but also, that (1a) is entailed by every sentence. In other words, (1a) is trivially true: the cardinality of the set of students who smoked, or in fact, of any set, is necessarily 0 or greater than 0. Given this consequence of their analysis, and given the assumption that trivial sentences are “semantically defective,” Bylinina and Nouwen claim that *zero*, unlike other numerals, always requires semantic strengthening by exhaustification.⁴ We will call this claim EXHAUSTIFICATION.

² We say ψ is “stronger” than ϕ iff ψ entails ϕ but not vice versa. That excludable alternatives are stronger than the prejacent, i.e. the sister of EXH, is a standard Neo-Gricean view (cf. Horn 1989; Sauerland 2004) and covers most of the facts about scalar inferences. For arguments for replacing “stronger” with “non-weaker” see e.g. Spector (2006); Schlenker (2012). In this squib we stick to “stronger” as that will simplify the exposition and serves the purposes at hand. Our position, we believe, is consistent with what is said in Bylinina & Nouwen (2018), although these authors were not as explicit about the interpretation of [EXH ϕ] in their paper as we are here (cf. Bylinina & Nouwen 2018: 10). For more discussion on exhaustification, see Magri (2009); Chierchia et al. (2012); Crnič et al. (2015); Buccola & Haida (2019), among others.

In this connection, we would also point out something else which our paper has in common with Bylinina & Nouwen (2018): both works appeal to alternatives, but neither work provides a theory of scales. It is common practice to propose accounts which are premised on specific claims about alternatives but which, at the same time, remain agnostic about how these claims are derived from deeper principles or, in fact, whether such principles exist at all. The validity of such accounts should remain even if it turns out that scales are just “given to us” as primitives of grammar (Gazdar 1979). There is a research program towards an intensional characterization of alternatives (cf. e.g. Matsumoto 1995; Katzir 2007; Fox & Katzir 2011; Katzir 2014; Trinh & Haida 2015; Breheny et al. 2017; Trinh 2018; 2019). However, we believe that the evaluation of proposals which appeal to alternatives but do not take part in this research program, such as Bylinina & Nouwen (2018) and this paper, can be independent of how successful the program is. We thank an anonymous reviewer for raising our awareness of the need to be clear about this issue.

³ Here is how. As the scale mates of *zero* are *one*, *two*, *three* etc., the alternatives of *zero students smoked* will be *one student smoked*, *two students smoked*, *three students smoked*, etc. The conjunction of *zero students smoked* with the negation of any of these alternatives does not entail any other alternative. For example, *zero students smoked* \wedge \neg *two student smoked* does not entail *one student smoked* and does not entail *three students smoked*. Thus, all of these alternatives are excludable, which means that [EXH [*zero students smoked*]] is true iff zero students smoked and not one student smoked and not two students smoked etc., i.e. true iff the number of students who smoked is zero and not greater.

⁴ A semantic representation of *zero students smoked* which would more immediately reflect Bylinina and Nouwen’s analysis is (i).

- (2) EXHAUSTIFICATION
Zero requires semantic strengthening by exhaustification

1.2 Prevention

EXHAUSTIFICATION may be said to receive supporting evidence from the deviance of (3) which, to the best of our knowledge, is a novel observation.

- (3) #At least zero students smoked

Here is how the reasoning may go. It has been noticed that modification of numerals by *at least* prevents semantic strengthening of the numeral (cf. Krifka 1999; Nouwen 2010; Kennedy 2015; Schwarz 2016; Buccola & Haida 2017). For example, (4a) cannot be read as (4b).

- (4) a. At least two students smoked
 b. At least two students smoked and it is not the case that at least three students smoked
 c. (4a) $\not\Rightarrow$ (4b)

Let us state the generalization, which we will call PREVENTION.

- (5) PREVENTION
 Modification of a numeral by *at least* prevents semantic strengthening of the numeral

Although what is crucial for our discussion is that PREVENTION holds, not how it is explained, here is one way this generalization can be accounted for. Suppose that the alternatives of a sentence containing *at least* are derived from it by replacing *at least* with its scale mates *exactly* and *more than* (cf. Kennedy 2015; Buccola & Haida 2017). This means the alternatives of $\phi = \textit{at least two students smoked}$ are $\psi = \textit{exactly two students smoked}$ and $\chi = \textit{more than two students smoked}$. And because $(\phi \wedge \neg\psi) \Rightarrow \chi$ and $(\phi \wedge \neg\chi) \Rightarrow \psi$, neither ψ nor χ is an excludable alternative of ϕ , which means that $[\text{EXH } \phi] \Leftrightarrow \phi$. Thus, modification of a numeral by

$$(i) \quad \exists x(\llbracket \text{zero} \rrbracket(x) = \times \llbracket \text{students} \rrbracket(x) = \times \llbracket \text{smoked} \rrbracket(x) = 1)$$

These authors assume that the linguistic ontology, i.e. \mathcal{D}_e , contains a “bottom element” \perp such that $\llbracket \text{zero} \rrbracket(\perp) = 1$ and $\times \llbracket \alpha \rrbracket(\perp) = 1$ for each expression α of type $\langle e, t \rangle$. We prefer the equivalent but ontologically less controversial representation in (1a), as the discussion below does not hinge on the existence of \perp . Note, also, that in both representations, *zero* ends up being trivially downward and upward entailing in its NP as well as in its VP argument. This means that both representations are compatible with Bylinina and Nouwen’s explanation for *zero*’s inability to license NPIs.

at least makes EXH vacuous, thus preventing the numeral from being semantically strengthened by exhaustification.⁵

Given that modification of a numeral by *at least* prevents semantic strengthening by exhaustification of the numeral, EXHAUSTIFICATION, which says *zero* requires such strengthening, leads to the prediction that modification of *zero* by *at least* will result in deviance. This prediction is confirmed by the deviance of (3). Thus, the deviance of (3) constitutes supporting evidence for EXHAUSTIFICATION.⁶

1.3 Pragmaticism

Now, starting from the supposition that EXHAUSTIFICATION is correct, a question that arises is (6).

- (6) Is semantic strengthening by exhaustification required by *zero* pragmatic or grammatical?

The answer given by Bylinina and Nouwen is that it is pragmatic. To quote from Bylinina & Nouwen (2018: 10): “Unlike other numerals, *zero* invokes exhaustification obligatorily. This is for purely pragmatic reasons.” We will call this claim PRAGMATICISM.

- (7) PRAGMATICISM
Semantic strengthening by exhaustification is required by *zero* for purely pragmatic reasons

Under the perspective of PRAGMATICISM, one would say that the sentence in (8) cannot be parsed as (8a) because (8a) is trivially true, hence uninformative, hence pragmatically odd. The reason (8) is acceptable is because there is another parse, (8b), which is informative and thus pragmatically appropriate.

- (8) Zero students smoked
 a. [Zero students smoked] trivial
 b. [EXH [Zero students smoked]] non-trivial

⁵ If we assume that alternatives of *at least two students smoked* are derived by replacing both *at least* and *two* with their scale mates, so that, for example, *exactly three students smoked* would be an alternative, then we would need a more complicated definition of EXH, specifically one which is based on the notion of innocent exclusion as proposed by Fox (2007a). For a discussion on this point see Schwarz (2016).

⁶ We thank an anonymous reviewer for drawing our attention to the need to make this point clearer.

With respect to the sentence in (9), one would say, in light of PRAGMATICISM, that it is perceived as deviant because both of its parses, (9a) and (9b), are trivial, hence uninformative and pragmatically odd.⁷

- (9) #At least zero students smoked
- | | | |
|----|---------------------------------------|---------|
| a. | [At least zero students smoked] | trivial |
| b. | [EXH [At least zero students smoked]] | trivial |

We believe that PRAGMATICISM can be challenged. If the deviance of (9), repeated in (10a), comes about by it being uninformative, then all sentences expressing the same meaning, and hence are equally uninformative, should be perceived as deviant in the same way. This is not the case, as evidenced by the acceptability of (10b), which is semantically equivalent to (10a).

- (10) a. #At least zero students smoked
 b. Zero or more students smoked

Both (10a) and (10b) are trivial and thus equally uninformative, but only (10a) is perceived as deviant.⁸ This suggests that the contrast between these two sentences is grammatical in nature. The question now is (11).

- (11) Which grammatical criterion distinguishes between (10a) and (10b)?

1.4 L-triviality

To address this question, let us consider (12a) and (12b), which are possible parses of (10a) and (10b), respectively.^{9,10}

⁷ Note that a sentence is acceptable if one of its parses is acceptable, and a sentence is unacceptable if all of its parses are unacceptable.

⁸ Note that the same point is made if we use (8a), i.e. the first parse of (8), instead of (10a). However, this would make the discussion unnecessarily complicated, as (8a) and (8b) are homophonous.

⁹ We should say here that by “parses” we mean Logical Form (LF), which is the structure generated by rules of syntax and interpreted by rules of semantics. LF may differ from Phonetic Form (PF) whose derivation may involve deletion (cf. Chomsky 1995; Fox 2003). In the case of (10b), what has been deleted in the derivation of PF is the words *students* and *smoked* in the first disjunct as well as the words *than* and *zero* in the second disjunct. In this discussion, the relevant structure is LF and not PF. We should point out that an anonymous reviewer raises some concern about whether (12b) is the correct analysis of (10b), i.e. whether **or** is underlyingly propositional. We believe it is standard to assume that it is. For arguments supporting this view, see Hirsch (2017).

¹⁰ We assume that EXH can be appended to any clausal constituent (cf. Fox 2007a; Magri 2009; Chierchia et al. 2012). Note that appending EXH to (12a) will not have any semantic consequence (cf. the discussion on PREVENTION above), and neither does appending EXH to the second disjunct of (12b), or in fact, to that whole disjunction (Fox & Hackl 2006; Chierchia et al. 2012). In other

- (12) a. [At least zero students smoked]
 b. [[EXH [Zero students smoked]] or [more than zero students smoked]]

Note the underlining in (12). Its purpose is to distinguish between the logical expressions, which are underlined, and the non-logical expressions, which are not underlined. For the purpose of this squib, we take logical expressions to be those whose semantic contribution depends only on facts about language, and the non-logical expressions to be those whose semantic contribution depends on both facts about language and facts about the world. To illustrate, suppose we are to know whether (12b) is true in a given possible world w .¹¹ For EXH, *zero*, *or* and *more than*, we would have to know what they mean in English. For *students* and *smoked*, however, we would have to know not only what they mean in English, but also, who the students are in w and among them who smoked in w .

Given the classification of expressions into the logical and the non-logical, we now have a way to describe a crucial difference between (12a) and (12b) regarding the source of their triviality: (12a) is trivial by virtue of its logical vocabulary alone, while (12b) is trivial by virtue of both its logical and non-logical vocabulary. Here is what we mean in more concrete terms. Suppose we replace the word *students* in the second disjunct of (12b) with the word *professors*, the result, which is (13), is not trivial: (13) is true in case no students smoked or some professors smoked, and false in case some students smoked and no professors smoked.

- (13) [[EXH [Zero students smoked]] or [more than zero professors smoked]]

In other words, had the non-logical part of (12b) been different, the sentence would not have been trivial. The same, however, does not hold for (12a): there is no way to make (12a) non-trivial by changing its non-logical part. It remains trivial no matter which noun replaces *students* and which verb replaces *smoked*.

Let us call sentences such as (12a) “L-trivial,” borrowing from Gajewski (2009). Now, L-triviality has been argued to cause deviance (Barwise & Cooper 1981; Fintel 1993; Gajewski 2003; Abrusán 2007; Gajewski 2009). To the extent that that argument is convincing, the fact that (12a) is L-trivial while (12b) is not can be taken to explain the contrast between these two structures, and hence, explain the

words, the only way for EXH to make any difference in (12) is to scope over only the first disjunct of (10b). Presenting semantic calculations to show this would unfortunately take us too far beyond the confines of this squib.

¹¹ Under the standard assumption that knowing the meaning of the sentence means knowing its truth condition, which means knowing what has to be the case for it to be true, which means knowing whether it is true in a given possible world (cf. Wittgenstein 1921). The distinction between facts about language and facts about the world is of course open to debate. When the question is raised as to whether this distinction has any justification, the situation becomes murky (cf. Quine 1951). But this squib is not the place for philosophical contemplation at this level.

contrast between (10a) and (10b).¹² Thus, we have come to a tentative answer to the question in (11), given below.¹³

- (14) The grammatical criterion which distinguishes between (10a) and (10b) is L-triviality

1.5 Circumvention

The argument for (14) relies, crucially, on the assumption that EXH can rescue a sentence from L-triviality. To see this, consider (15a), which has the truth condition in (15b).

- (15) a. $[[\text{Zero students smoked}] \text{ or } [\text{more than zero students smoked}]]$
 b. $|[[\text{students}]] \cap [[\text{smoked}]]| \geq 0 \vee |[[\text{students}]] \cap [[\text{smoked}]]| > 0$

The first disjunct is trival. Moreover, it is L-trivial, as substituting $[[\text{students}]]$ or $[[\text{smoked}]]$ with any other predicate will not change the fact that it is trivial. This means that the whole disjunction is L-trivial. But (15a) is just (12b) without EXH, and since the argument for (14) depends on the claim that (12b) is not L-trivial, this means that that argument depends on the claim that EXH can circumvent L-triviality. Let us call this claim CIRCUMVENTION.

¹² The following contrast might raise doubt about the deviance of *#at least zero* being due to L-triviality, as it is implausible that *Celsius* and *Kelvin* are part of the logical vocabulary.

- (i) a. The temperature is at least zero degrees Celsius
 b. #The temperature is at least zero degrees Kelvin

However, we do not need to assume that *Celsius* and *Kelvin* are logical terms to explain this contrast. What we can say is that the presence of *Celsius* vs. *Kelvin* affects whether *zero* denotes the lower endpoint of the scale or not. Our discussion on *zero* is premised on the understanding that it does denote such a point. Non-standard readings in which it does not, as in (ia), are not relevant.

In this connection, we would also acknowledge a problem which mathematical statements pose for this whole approach. The problem is illustrated by the fact that (ii) is not perceived as deviant.

- (ii) One plus two equals three

Given our notion of “logical expression,” all words in (ii) belong to the logical vocabulary of English. Thus, the sentence is L-trivial, which means it is predicted to be deviant, contrary to intuition. At this point, we have nothing to say about mathematical statements. Our conjecture is that mathematical discourse invokes a different understanding of “facts about the world,” hence can be perceived as being contingent and informative, even though they are either trivially true or trivially false. We thank an anonymous reviewer for drawing our attention to this problem.

¹³ L-triviality is a “grammatical” criterion in the sense that whether a sentence is L-trivial can be determined entirely on its syntax and semantics, without any reference to contextual factors such as discourse participants or common knowledge.

- (16) CIRCUMVENTION
EXH can rescue a sentence from L-triviality

We believe CIRCUMVENTION can be challenged. Consider (17) and its possible parses in (17a) and (17b).

- (17) #Every human has at least zero children
 a. [Every human has at least zero children]
 b. [EXH [Every human has at least zero children]]

This sentence is perceived as deviant in the same way (10a) is. The parse without EXH, (17a), is L-trivial. However, the parse with EXH, (17b), is not. Assuming, again, that *at least* alternates with *exactly* and *more than*, the truth condition of (17b) will be (18a), which is equivalent to (18b), a contingent proposition.¹⁴

- (18) a. (Every human has zero or more children) \wedge \neg (every human has exactly zero children) \wedge \neg (every human has more than zero children)
 b. Some humans have children and some humans have no children

If CIRCUMVENTION is correct, we predict (17) to be non-deviant, because it has a parse which is not L-trivial. Thus, the fact that (17) is perceived as deviant suggests that CIRCUMVENTION is not correct, i.e. that EXH cannot rescue a sentence from L-triviality. This is perhaps not a surprising result. We expect a sentence to be ungrammatical if it contains a constituent which in isolation would itself be ungrammatical. If we take L-triviality to be a grammatical criterion, we expect a sentence with an L-trivial constituent to be deviant regardless of what other linguistic materials it may contain. This means that we cannot append EXH to an L-trivial sentence ϕ to “save” it, even if [EXH ϕ] is not L-trivial.

Another fact which speaks against CIRCUMVENTION is that (19) is perceived as deviant (Barwise & Cooper 1981).

- (19) #There is every student

¹⁴ An anonymous reviewer raises two questions about (17).

- (i) a. What if EXH takes scope below **every**?
 b. What if the alternatives also include those generated by replacing **every** with **some**?

Here are our answers. First, if EXH takes scope below **every**, it would be semantically vacuous, and the parse would be L-trivial. The existence of such a parse, of course, has no bearing on our argument against CIRCUMVENTION, namely that (17) is deviant even though there is a parse for it which is not L-trivial. Our answer to the second question is that it makes no difference whether the **some** alternatives are included in the domain of EXH or not, as these would not be stronger than the preadjacent and hence would not be excludable.

As *every* expresses the subset relation, the truth condition of (19) can be stated as $\llbracket \textit{student} \rrbracket \subseteq \llbracket \textit{there} \rrbracket$, where $\llbracket \textit{there} \rrbracket = \mathcal{D}_e$, i.e. the set of all things that exist. Since every set is a subset of \mathcal{D}_e , the sentence is L-trivial, hence deviant. Now, given the standard assumption that *every* has *some* as its scale mate, a parse of (19) with EXH, (20a), would express the proposition in (20b), which is equivalent to (20c).

- (20) a. [EXH [There is every student]]
 b. (There is every student) \wedge \neg (there is some student)
 c. There is no student

Thus, appending EXH to (19) results in the proposition that there is no student, which is not trivial. If CIRCUMVENTION is correct, (19) should be acceptable and should express this contingent proposition. This is not what is observed.

1.6 The dialectical situation

Let us recap. We start with a claim about the meaning of the numeral *zero*, made by Bylinina & Nouwen (2018), which says that *zero* has the basic weak meaning of ‘zero or more.’ The dialectics then goes as follows. If Bylinina and Nouwen’s claim is correct, then *zero* requires EXH. This requirement is either pragmatic or grammatical. Observation suggests that it is not pragmatic, hence must be grammatical. Saying that the requirement for EXH by *zero* is grammatical, however, implies that EXH can rescue a sentence from L-triviality. But observation suggests that EXH cannot rescue a sentence from L-triviality. Thus, we are lead to the conclusion that Bylinina and Nouwen’s claim is not correct, i.e. that *zero* does not mean ‘zero or more.’ The conclusion is tentative, pending alternative explanations for the facts that Bylinina and Nouwen provided to support their claim.

We would note, ending this section, that all of the data we discussed would be consistent with the view that *zero* has the strong meaning as basic and that EXH cannot circumvent L-triviality. Specifically, (21a) would mean ‘no students smoked,’ (21b) would be trivial but not L-trivial, and both (21c) and (21d) would be L-trivial and therefore deviant.¹⁵

- (21) a. [Zero students smoked]
 b. [~~Zero students smoked~~] or [more than zero students smoked]
 c. #[At least zero students smoked]
 d. #[Every human has at least zero children]

¹⁵ We represent phonological deletion with ~~strikethrough~~. For arguments that **zero** differ from the negative quantifier **no** in a number of formal respects see Bylinina & Nouwen (2018).

However, noting such facts is far from proposing a theory of *zero*. What we hope to have achieved is to raise some issues, present some novel observations, and share some ideas which can stimulate and provide some guiding intuition for further research.

2 Some quantitative data

2.1 *Existential sentences*

To corroborate the intuition that *zero*, unlike other numerals, cannot be modified by the adverb *at least*, we conducted an experiment, hosted on Amazon MTurk, in which participants were asked to rate the **naturalness** of eight English sentences. The sentences were derived from the sentence frames in (22) by replacing the placeholder *n* with the numeral *zero*, for a set of four deviant sentences, and the numeral *two*, for a set of four non-deviant sentences.

- (22)
- a. There are at least *n* students in the seminar room
 - b. The drawer contains at least *n* towels
 - c. The company hired at least *n* employees
 - d. The bartender served at least *n* guests

We asked participants to rate the resulting sentences, presented in pseudo-randomized order, on the following four-point Likert scale: 4 (natural), 3 (relatively natural), 2 (relatively weird), 1 (weird). To illustrate, in advance of the trials, how these values may be associated with English sentences, we gave participants the four sentences in (23) and commented: “You may agree that (23a) is natural, (23b) is weird, while (23c) and (23d) are possibly somewhere in between.”¹⁶

- (23)
- a. Everyone but John came to the party
 - b. Someone but John came to the party
 - c. Everyone who is not John came to the party
 - d. Someone who is not John came to the party

We used an even-numbered scale to force participants to discriminate between sentences which they may be unable to judge as 4 (‘natural’) or 1 (‘weird’), and avoided suggesting that there is agreement on the score of such sentences in order not to stifle participants’ trust in their own judgments. We believe that the use of a four-point scale, excluding a neutral value, does not have an adverse effect on the interpretation of the expected findings of our study: even if there was perceived social or other

¹⁶ The acceptability contrast between (23a) and (23b) is discussed and explained in Fintel (1993).

pressure (not) to classify sentences as ‘(relatively) natural’ or ‘(relatively) weird’, potential skewing would not have a bearing on the expected difference in the score for sentences with *at least zero* and sentences with *at least two*.

We conducted our experiment with 32 Amazon MTurk workers who, after the trials, identified as being native speakers of English. Thus, overall we received 32 scores for each of our eight sentences and hence 128 scores per sentence type. Figure 1 shows that sentences with *at least two* received the highest score 4 (‘natural’) by $\geq 50\%$ of all subjects, while sentences with *at least zero* received the two lowest scores 2 (‘relatively weird’) and 1 (‘weird’) by $\geq 50\%$ of all subjects. The difference in the means of the scores (3.4 v 2.0), depicted in Figure 2, is highly significant ($p < 2.2^{-16}$).

Figure 1: Boxplot of *at least two* & *at least zero*.

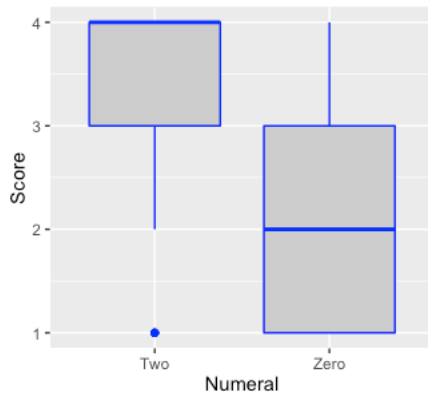
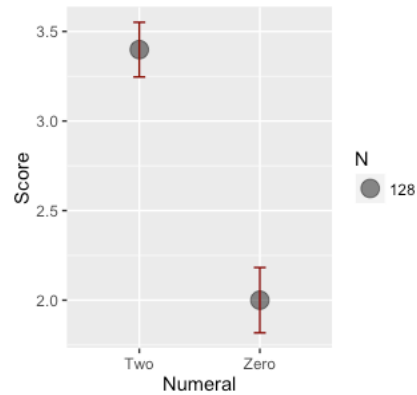


Figure 2: Means of *at least two* & *at least zero*.



This result supports our empirical claim that numerical statements containing *at least zero*, unlike statements containing other superlative modified numerals, are deviant.

2.2 *Universally quantified sentences*

If the deviance induced by the triviality of *zero* is pragmatic, it should be alleviated by exhaustification, and *at least zero* should be acceptable under universal quantification, as such quantification renders exhaustification non-vacuous. Example (17) suggests that the deviance of *at least zero* persists under universal quantification. To corroborate this intuition, we conducted another experiment, again hosted on Amazon MTurk. This experiment targeted the contrast between *#every ... at least zero*

and its counterpart containing *zero or more* as well as the non-zero counterparts of the former two. We used the eight sentences in (24) as stimuli.

- | | | | | |
|------|----|------|---|---------|
| (24) | a. | (i) | #Every human has at least zero children | (TRUE) |
| | | (ii) | #Every human has at least zero biological mothers | (FALSE) |
| | b. | (i) | Every human has zero or more children | TRUE |
| | | (ii) | Every human has zero or more biological mothers | FALSE |
| | c. | (i) | Every human has at least one biological mother | FALSE |
| | | (ii) | Every human has at least two relatives | TRUE |
| | d. | (i) | Every human has one or more biological mothers | FALSE |
| | | (ii) | Every human has two or more relatives | TRUE |

Two of these sentences, viz. the sentences in (24a), are deviant, while the others are non-deviant. Abstracting from the deviance of (24a-i) and (24a-ii),¹⁷ we can characterize the meaning of the sentences in (24) in the following way. All sentences license a distributive inference.¹⁸ For instance, the sentence in (24b-i) licenses the inference that some humans have (exactly) zero children and some humans have one or more children. The distributive inference renders (24b-i), as well as the other sentences (24a) and (24b), non-tautological. The distributive inference of (24b-i) does not, however, render this sentence false in the actual world. In this regard, it differs from e.g. the sentence in (24a-ii): in the actual world, it is false that some humans have (exactly) zero biological mothers. Overall, as indicated, if we take the distributive inference into account four of the eight sentences in (24) are true in the actual world, while the other four are false.¹⁹

We instructed participants to classify the sentences in (24) as either ‘true’, ‘false’, or ‘weird’²⁰ by giving them the following instruction in advance of the trials: “Imagine you are helping an alien who wants to learn English and also learn about humans in general. Your job is to say for several English sentences whether they are true, false, or sound weird.” The scenario we sketched in this instruction served to introduce an addressee without common knowledge. With this, we aimed for subjects to be less drawn to classify non-deviant sentences such as the sentences in (24b-i) or (24c-ii) as weird despite the fact that they express common knowledge with little informational content.²¹ At the same time, the instruction aimed at mak-

¹⁷ That is, we consider these sentences as if language was not sensitive to L-triviality.

¹⁸ As pointed out in Büring (2008), *at least n* licenses the same inferences as *n or more*.

¹⁹ Again, we are abstracting from the deviance of the sentences in (24a).

²⁰ A pilot study using the four-point Likert scale in § 2.1 failed to elicit discriminative judgments for the contrast between (24a-i) and (24b-i). That is, both sentences received equally low scores. We believe that different factors are at play for the two sentences to be judged as (relatively) weird on the four-point scale we used. See the discussion below.

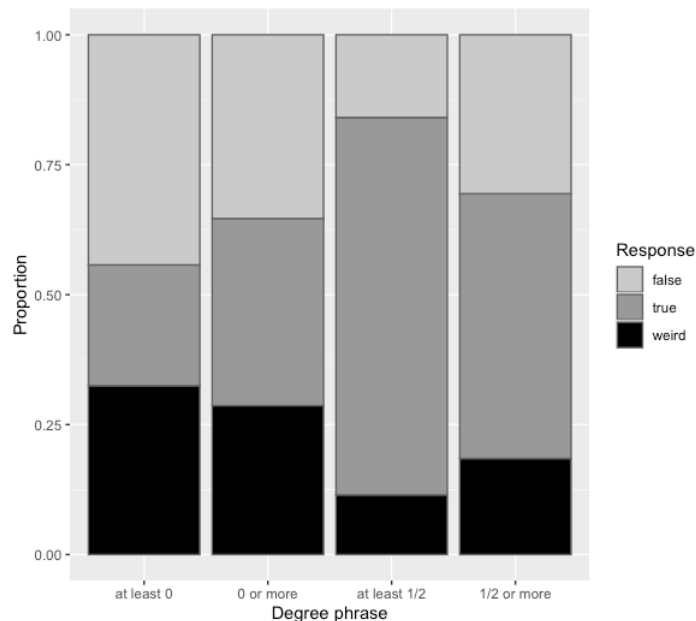
²¹ Recall that (24b-i) has a distributive inference which renders this sentence non-tautological.

ing subjects more inclined to classify the deviant sentences in (24a) as weird by pointing out that the addressee also wants to learn English.

We expect that the proportion of ‘weird’ responses is higher for the sentences in (24a) than for the sentences in (24b), in accordance with our empirical claim that the former, in contrast to the latter, are deviant. Furthermore, we expect that the proportion of ‘weird’ responses is higher for the sentences in (24a) and (24b) than for the sentences in (24c) and (24d). This expectation is based on two factors: (i) the assumption that tautological sentences are more prone to being judged as weird than contingent sentences, and (ii) the experimental finding that there tends to be a substantial subpopulation of subjects that do not to compute distributive inferences (Crnič et al. 2015). Such a subpopulation derives tautological truth conditions for the sentences in (24a) and (24b), while the sentences in (24c) and (24d) are contingent with or without the distributive inference.

We conducted our experiment with 157 Amazon MTurk workers who, after the trials, identified as being native speakers of English. Thus, we received 157 scores for each of our eight sentences (1256 observations in total). Figure 3 shows the proportion of ‘true’, ‘false’, and ‘weird’ responses for the degree phrases in (24a), (24b), (24c), and (24d), respectively.

Figure 3: Proportion of True-False-Weird judgments.



In the subsequent statistical analysis, we consider as dependent variable the binary distinction between ‘weird’ responses and ‘true’ or ‘false’ responses (‘non-weird’ responses). We used binomial logistic regression to analyze the relationship between this binary *response* variable and the following four independent variables: *modifier* with values ‘at least’ and ‘or more’, *numeral* with values ‘zero’ and ‘non-zero’, *at least zero* with values ‘+’ and ‘-’, and *truth value* with values ‘true’ or ‘false’. These variables describe whether the stimulus sentence contains the modifier *at least* or the modifier *or more*, whether the stimulus sentence contains the numeral *zero* or the numeral *one* or *two*, whether the stimulus sentence contains the degree phrase *at least zero* or any other degree phrase, and whether the stimulus sentence is true or false in the actual world, respectively. All four variables turn out to be significant predictors of the *response* variable with *p* values of 0.01449, 0.00273, 0.01984, and 0.00995, respectively. Specifically, the modifier *at least* and the truth value ‘true’ decrease the odds of the ‘weird’ response (reduction of the odds by 43% and 28%, respectively), while the numeral *zero* and the degree phrase *at least zero* increase the odds of the ‘weird’ response (increase of the odds by 78%, and 110%, respectively). The results of the ANOVA between this model and the null model show that, as suggested by the *p* values of the logistic regression, the *numeral* variable is by far the biggest contributor to the fit of the model (reduction of deviance by 44.4 from 1347.7 null deviance, $p = 2.686^{-11}$), followed by the *at least zero* variable (reduction of 6.7, $p = 0.0094$), and *truth value* (reduction of 5.2, $p = 0.0221$), while the *modifier* variable does not contribute significantly to the fit of the model ($p = 0.5$).

These results confirm the expectation that the proportion of ‘weird’ responses is higher for the sentences in (24a) and (24b) than for the sentences in (24c) and (24d), and higher for the sentences in (24a) than for the sentences in (24b). Thus, overall the results are compatible with our empirical claim that the sentences in (24a) are deviant, while those in (24b), (24c), and (24d) are non-deviant.²² Questions remain about the influence of the truth or falsity of the stimulus sentence on the odds of the ‘weird’ response and about the reason for the overall tendency for the modifier *or more* to raise the odds of the ‘weird’ response compared to the modifier *at least*.²³ We hope to address these questions in future research.

²² For further corroboration, we note that an informal Google search on September 14, 2019, shows a clear difference in the production of the expressions under consideration: a search of, e.g., the phrase *zero or more times* gives 237,000 results, while a search for *at least zero times* only gives 453 results.

²³ Recall that, importantly, our results also show that the odds are reversed for *at least zero* relative to other degree phrases.

3 Conclusion

Bylinina & Nouwen (2018) propose the ‘at least’ meaning as basic for *zero* to account for a number of intuitions about this numeral. This squib discusses some issues which arise from this proposal. At the center of the discussion is the observation that *at least zero* leads to deviance while the semantically equivalent *zero or more* does not, as well as the concept of L-triviality and its place in the linguistic system. We present some experimental data to corroborate our observation.

Funding information

This work was funded by the ERC Advanced Grant “Speech Acts in Grammar and Discourse” (SPAGAD), ERC-2017-ADG 787929 and grant 2093/16 of the Israel Science Foundation.

Acknowledgements

The authors wish to thank Brian Buccola, Manfred Krifka, and the audiences at the ZAS Semantics Circle and the XPRAG workshop “The meaning of numerals for helpful discussion.

References

- Abrusán, Martha. 2007. *Contradiction and Grammar: the Case of Weak Islands*: MIT dissertation.
- Barwise, Jon & Robin Cooper. 1981. Generalized quantifiers and natural language. *Linguistics and Philosophy* 4(2). 159–219.
- Breheeny, Richard, Nathan Klinedinst, Jacopo Romoli & Yasutada Sudo. 2017. The symmetry problem: current theories and prospects. *Accepted with revisions for publication in Natural Language Semantics*. .
- Buccola, Brian & Andreas Haida. 2017. Obligatory irrelevance and the computation of ignorance inferences. Unpublished manuscript, v1 of lingbuzz/003600.
- Buccola, Brian & Andreas Haida. 2019. Obligatory irrelevance and the computation of ignorance inferences. *Journal of Semantics* 36. 583–616.
- Büring, Daniel. 2008. The least *at least* can do. In C.B. Chang & H.J. Haynie (eds.), *Proceedings of the 26th West Coast Conference on Formal Linguistics*, 114–120. Somerville: Cascadilla Press.
- Bylinina, Lisa & Rick Nouwen. 2018. On “zero” and semantic plurality. *Glossa: A Journal of General Linguistics* 3(1). 1–23.

- Chierchia, Gennaro, Danny Fox & Benjamin Spector. 2012. The grammatical view of scalar implicatures and the relationship between semantics and pragmatics. In Paul Portner, Claudia Maienborn & Klaus von Stechow (eds.), *Semantics: An International Handbook of Natural Language Meaning*, De Gruyter.
- Chomsky, Noam. 1995. *The Minimalist Program*. MIT Press.
- Crnič, Luka, Emmanuel Chemla & Danny Fox. 2015. Scalar implicatures of embedded disjunction. *Natural Language Semantics* 23. 271–305.
- Fintel, Kai von. 1993. Exeptive constructions. *Natural Language Semantics* 1(2). 123–148.
- Fox, Danny. 2003. On Logical Form. In Randall Hendrick (ed.), *Minimalist Syntax*, 82–123. Blackwell.
- Fox, Danny. 2007a. Free choice disjunction and the theory of scalar implicatures. In Uli Sauerland & Penka Stateva (eds.), *Presupposition and Implicature in Compositional Semantics*, 71–120. Palgrave-Macmillan.
- Fox, Danny. 2007b. Pragmatics in Linguistic Theory. MIT classnotes.
- Fox, Danny & Martin Hackl. 2006. The universal density of measurement. *Linguistics and Philosophy* 29. 537–586.
- Fox, Danny & Roni Katzir. 2011. On the characterization of alternatives. *Natural Language Semantics* 19. 87–107.
- Gajewski, Jon. 2003. L-analyticity in natural language. Unpublished manuscript.
- Gajewski, Jon. 2009. L-triviality and grammar. Unpublished manuscript.
- Gazdar, Gerald. 1979. *Pragmatics: Implicature, Presupposition and Logical Form*. New York: Academic Press.
- Hirsch, Aron. 2017. *An Inflexible Semantics for Cross-categorical Operators*: Massachusetts Institute of Technology dissertation.
- Horn, Laurence. 1972. *On the semantic properties of the logical operators in english*: UCLA dissertation.
- Horn, Laurence R. 1989. *A natural history of negation*. Chicago: University of Chicago Press.
- Katzir, Roni. 2007. Structurally-defined alternatives. *Linguistics and Philosophy* 30. 669–690.
- Katzir, Roni. 2014. On the roles of markedness and contradiction in the use of alternatives. In S Pistoia-Reda (ed.), *Semantics, Pragmatics and the Case of Scalar Implicatures*, 40–71. Palgrave.
- Kennedy, Christopher. 2015. A “de-Fregean” semantics (and neo-Gricean pragmatics) for modified and unmodified numerals. *Semantics and Pragmatics* 8(10). 1–44.
- Krifka, Manfred. 1999. At least some determiners aren’t determiners. In K. Turner (ed.), *The semantics/pragmatics interface from different points of view*, vol. 1 (Current Research in the Semantics/Pragmatics Interface), 257–291. Elsevier

- Science B.V.
- Magri, Giorgio. 2009. A theory of individual-level predicates based on blind mandatory scalar implicatures. *Natural Language Semantics* 17(3). 245–297.
- Matsumoto, Yo. 1995. The conversational condition on Horn scales. *Linguistics and Philosophy* 18. 21–60.
- Nouwen, Rick. 2010. Two kinds of modified numerals. *Semantics and Pragmatics* 3(3). 1–41. <https://doi.org/10.3765/sp.3.3>.
- Quine, Willard Van Orman. 1951. Two dogmas of empiricism. *Philosophical Review* 60. 20–43.
- Sauerland, Uli. 2004. Scalar implicatures in complex sentences. *Linguistics and Philosophy* 27(3). 367–391.
- Schlenker, Philippe. 2012. The semantics/pragmatics interface. In Maria Aloni & Paul Dekker (eds.), *Handbook of semantics*, .
- Schwarz, Bernhard. 2016. Consistency preservation in quantity implicature: The case of *at least*. *Semantics and Pragmatics* 9(1). 1–47. <https://doi.org/10.3765/sp.9.1>.
- Spector, Benjamin. 2006. *Aspects de la pragmatique des operateurs logiques*: Université Paris Diderot Paris 7 dissertation.
- Trinh, Tue. 2018. Keeping it simple. *Natural Language Semantics* 26. 111–124.
- Trinh, Tue. 2019. Exhaustification and contextual restriction. *Frontiers in Communication* 4. 1–7.
- Trinh, Tue & Andreas Haida. 2015. Constraining the derivation of alternatives. *Natural Language Semantics* 23. 249–270.
- von Stechow, Kai & Irene Heim. 1997. Pragmatics in Linguistic Theory. MIT class-notes.
- Wittgenstein, Ludwig. 1921. Logisch-philosophische Abhandlung. *Annalen der Naturphilosophie* 14. 185–262.