Up to n: pragmatic inference about an optimal lower bound

Ciyang Qing
Stanford University

Abstract

Blok (2015) observes that the directional modified numeral *up to n* can trigger a *proxmity inference*: after hearing *up to 100 people will come to my wedding*, the listener can infer that the speaker thinks the number of guests will be close to 100. In this paper, I draw a parallel between *up to n* and vague gradable adjectives such as *tall*, and argue that the proximity inference is due to the vagueness of *up to n* in *speaker ignorance* contexts. Complicating the picture, whether *up to n* is vague depends on the speaker's contextual level of uncertainty. In *authoritative permission* contexts, e.g., *you are allowed to borrow up to 100 books*, *up to n* is not vague and receives a *full range interpretation*: borrowing 0–100 books is allowed. To explain the opposite inference patterns of *up to n* in different contexts, I propose a new semantics of *up to n*, which has an unspecified lower bound similar to vague gradable adjectives, and leverage previous probabilistic models of gradable adjectives to pragmatically infer the distribution of the lower bound (Lassiter & Goodman, 2013, 2015; Qing & Franke, 2014a, 2014b). Crucially, the interaction between *informativity* and *applicability* in different contexts predicts the opposite inference patterns of *up to n*.

The modified numerals *up to n* and *at most n* form an interesting pair. Intuitively they seem to have the same semantic content. However, as Blok (2015) points out, *up to n* but not *at most n* triggers a *proximity inference*. For example, (1a) but not (1b) implies that the number of guests will be under but close to 100. Blok leaves this contrast as an open issue.

(1) a. Up to 100 people will attend my wedding. $\rightsquigarrow$ the number is close to 100
  b. At most 100 people will attend my wedding. $\rightsquigarrow$ no such implication

On the other hand, the contrast in (1) disappears when the modified numerals are under the scope of permission modals and the speaker is assumed to be *authoritative*: both (2a) and (2b) grant the listener permission to borrow from 0 to 100 books. This is the opposite of the proximity inference and I will call it the *full-range* inference.[1]

(2) a. You are allowed to borrow up to 100 books. $\rightsquigarrow$ any number in $[0, 100]$ allowed
  b. You are allowed to borrow at most 100 books. $\rightsquigarrow$ any number in $[0, 100]$ allowed

Therefore, *up to n* is sensitive to the linguistic environment. It triggers a proximity inference in unembedded sentences but a full-range inference under permission modals.

Puzzle Why does *up to n* trigger opposite inference patterns in unembedded sentences and under permission modals?

Previous work Blok (2015) uses a *contrary-to-expectation* diagnostic (among other things) to argue that *up to n* sets a semantic lower bound that excludes 0 (3).

(3) a. ? I expect to see at most ten people, but maybe no one will show up.
  b. I expect to see up to ten people, but maybe no one will show up.

According to Blok, (3a) sounds less coherent because the possibility of seeing no one is already included in the expectation of seeing at most ten people. Hence there is no contrast between the two clauses to license *but*. On the other hand, the *but*-clause in (3b) is felicitous, which means that the possibility of seeing no one is contrary to the expectation described in the main clause. Therefore Blok concludes that *up to n* sets a semantic lower bound that excludes 0.

New data & generalization Blok's proposal is informative, but she only considers the possibility of 0 in the *but*-clause. Naturally-occurring examples suggest that the *but*-clause can also contain non-zero numbers (4).

(4) a. Vernell expected up to 10 vendors but only six materialized.
  b. Allison had expected up to 1,000 extras, but only about 60 people were in costume on the set.

This suggests that *up to n* can also set a semantic lower bound that excludes non-zero numbers. Nevertheless, as the number in the *but*-clause approaches the upper bound n, the entire sentence sounds less and less coherent and eventually becomes totally infelicitous (5).

[1] An alternative interpretation of (2a) and (2b) is that the speaker does not know the number of books the listener is allowed to borrow and is giving an estimation. In this *speaker-insecure* interpretation, (2a) but not (2b) triggers a proximity inference. This interpretation corresponds to a wide-scope reading of *up to n* and the analysis of unembedded sentences in this paper can be similarly applied.

(5) I expect to see up to 100 people, but maybe only/no more than 10 / 20 / ... / 50 / (?)60 / ?80 / ??90 / #95 will show up.

Similarity to gradable adjectives To motivate an analysis of the graded exclusion of numbers in (5), I note that vague gradable adjectives such as *tall* have a similar pattern (6).

(6) I expect John to be tall, but maybe he is only/no taller than 5'5" /... / 5'10" / (?)6' / ?6'2" / ??6'4" / #7' ... (tall).

In a degree-based semantics for gradable adjectives (e.g., Kennedy & McNally, 2005; Kennedy, 2007), the positive form of a gradable adjective A introduces a contextually determined *standard of comparison* θ, and *x is A* is true iff x's degree of A-ness exceeds θ. For vague adjectives, although the precise value of θ is unknown, people typically have a sense of the probability distribution of θ in the context. For instance, suppose we know that the standard of *tall* for a US adult male is likely to be between 6' and 6'2". Then we can be fairly certain that small degrees of height such as 5'10" are lower than the standard θ and hence not part of the expectation. This is why the *but*-clause in (6) is coherent for small degrees. We can also be fairly certain that a large degree such as 6'4" is greater than the standard, and hence it is part of the expectation, making the *but*-clause incoherent. But there are also borderline cases in between, i.e., degrees that are unclear whether they exceed the standard. In other words, the graded exclusion of a degree d is explained by the probability that $d < \theta$.

The parallel between (5) and (6) suggests that there are borderline cases in the interpretation of *up to n*, which is a characteristic property of vagueness (Kennedy, 2007).

A new semantics of *up to n* Based on its similarity to gradable adjectives, I propose that *up to n* imposes a contextually determined semantic lower bound θ. Following previous work (Coppock & Brochhagen, 2013; Blok, 2015), I use the inquisitive semantics framework (Ciardelli, Groenendijk, & Roelofsen, 2012) for implementation (7).

(7) a. $[\![\text{up to n}]\!] = \{\lambda M_{\langle d,t\rangle}.\, \max_d(M) = k \mid k \in [\theta, n]\}$, where θ is a contextual lower bound ($0 \leq \theta < n$).[2] [3]

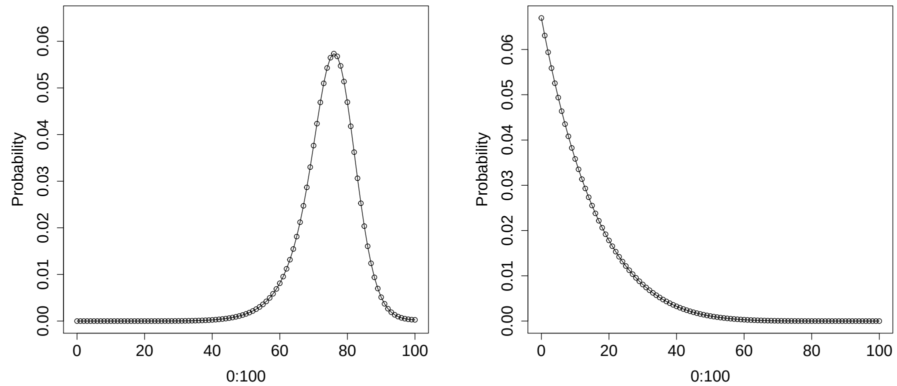
b. $[\![\text{Up to 100 people will attend my wedding}]\!] = \{p_\theta, p_{\theta+1}, \ldots, p_{100}\}$, where $0 \leq \theta < n$ and p_i is the proposition that exactly i people will attend my wedding.

c. $[\![\text{You are allowed to borrow up to 100 books}]\!] = \{\Diamond\{p_\theta, p_{\theta+1}, \ldots, p_{100}\}\}$, where $0 \leq \theta < n$ and p_i is the proposition that you borrow exactly i books.

An unembedded sentence such as (1a) denotes a set of propositions (7b) and the informative content is the union of all the propositions, i.e., the proposition that the number of guest is within the range $[\theta, n]$. In the authoritative reading of (2a), the permission modal takes scope above the set of propositions (7c).

[2] Blok (2015) argues that the upper bound of *up to n* is pragmatically implicated rather than semantically entailed. Here I put the max operator in the denotation of *up to n* to gloss over the pragmatic exhaustification procedure, but nothing hinges on the nature of the upper bound.

[3] Blok (2015) proposes that the lower bound is a contextual starting point of the scale, which seems very similar. However, her motivation is to capture different granularity of the contextual quantity scale, and the lower bound is always the smallest non-zero degree (e.g., if eggs are bought in half dozens then the lower bound would always be 6). The current semantics is more flexible than Blok's in that the lower bound can be any number less than n. This difference is crucial to capturing the proximity inference.



(a) unembedded sentence (7b)   (b) authoritative permission (7c)

Figure 1: Probability distributions of the lower bound θ in two linguistic environments

Pragmatic inference about the optimal lower bound What remains to be answered is how the unspecified semantic lower bound θ is contextually determined. I will provide a pragmatic mechanism independently motivated in the study of gradable adjectives (Lassiter & Goodman, 2013, 2015; Qing & Franke, 2014a, 2014b). The core idea is that the distribution of θ is due to an interaction between *informativity* and *applicability*, where informativity is measured by the strength of the informative content and applicability is the probability that a sentence is assertible. The optimal lower bound θ should strike a balance between the two.

When *up to n* is unembbeded (7b), a higher θ corresponds to a narrower range $[\theta, n]$, which makes the sentence more informative. Meanwhile, (7b) is assertible iff the speaker believes that the actual number of guests is within the range $[\theta, n]$. Given that the speaker does not know the exact number of guests, a wider range is more likely to contain all the numbers of guests that the speaker considers possible and hence a lower θ would make the sentence more applicable. Assuming that the speaker is reasonably informed but has residual uncertainty, the optimal lower bound should be close to n but not too close. This captures the proximity inference in unembedded sentences.

In the authoritative reading (7c), when the permission modal scopes above the set of propositions, it gives rise to a *free-choice* inference that each proposition in the set is allowed (8) (e.g., Kamp, 1973, 1978; Zimmermann, 2000; Kratzer & Shimoyama, 2002).

(8) $\Diamond\{p_\theta, p_{\theta+1}, \ldots, p_{100}\} \rightsquigarrow \Diamond p_\theta \wedge \Diamond p_{\theta+1} \wedge \ldots \wedge \Diamond p_{100}$

Assuming that the result of the free-choice inference is the informative content of (7c) that feeds into the pragmatic mechanism, a lower θ corresponds to more conjuncts in (8), which makes the sentence more informative. Meanwhile, the sentence is assertible iff the speaker has authority, which is assumed to be the case in the authoritative reading. Therefore there is no preference in terms of applicability. Hence the optimal θ is 0, which captures the full range inference under permission modals.

The above analysis accounts for the different inference patterns of *up to n* in unembedded sentences and under permission modals and can be extended into a probabilistic model that makes testable quantitative predictions (Figure 1 illustrates the different distributions of θ in the two environments). It suggests a close relation between *up to n* and gradable adjectives and the generality of the informativity-applicability trade-off in pragmatics.

References

Blok, D. (2015). The semantics and pragmatics of directional numeral modifiers. In *Semantics and linguistic theory (SALT)* (Vol. 25). LSA and CLC Publications.

Ciardelli, I., Groenendijk, J., & Roelofsen, F. (2012). Inquisitive semantics. *NASSLLI 2012 lecture notes*.

Coppock, E., & Brochhagen, T. (2013). Raising and resolving issues with scalar modifiers. *Semantics and Pragmatics*, *6*(3), 1–57. doi: 10.3765/sp.6.3

Kamp, H. (1973). Free choice permission. *Proceedings of the Aristotelian Society*, *74*, 57-74.

Kamp, H. (1978). Semantics versus pragmatics. In F. Guenthner & S. J. Schmidt (Eds.), *Formal semantics and pragmatics for natural languages* (pp. 255–287). Dordrecht: Reidel.

Kennedy, C. (2007). Vagueness and grammar: The semantics of relative and absolute gradable adjectives. *Linguistics and Philosophy*, *30*, 1–45.

Kennedy, C., & McNally, L. (2005). Scale structure, degree modification, and the semantics of gradable predicates. *Language*, *81*(2), 345–381.

Kratzer, A., & Shimoyama, J. (2002). Indeterminate pronouns: The view from japanese. In Y. Otsu (Ed.), *Proceedings of the third tokyo conference on psycholinguistics* (pp. 1–25). Tokyo, Japan: Hituzi Syobo Publishing Company.

Lassiter, D., & Goodman, N. D. (2013). Context, scale structure, and statistics in the interpretation of positive-form adjectives. In *Proceedings of SALT 23.*

Lassiter, D., & Goodman, N. D. (2015). Adjectival vagueness in a Bayesian model of interpretation. *Synthese*.

Qing, C., & Franke, M. (2014a). Gradable adjectives, vagueness, and optimal language use: A speaker-oriented model. In J. Grieser, T. Snider, S. D'Antonio, & M. Wiegand (Eds.), *Proceedings of SALT* (pp. 23–41). elanguage.net.

Qing, C., & Franke, M. (2014b). Meaning and use of gradable adjectives: Formal modeling meets empirical data. In P. Bello, M. Guarini, M. McShane, & B. Scassellati (Eds.), *Proceedings of cogsci* (p. 1204-1209). Austin, TX: Cognitive Science Society.

Zimmermann, T. E. (2000). Free choice disjunction and epistemic possibility. *Natural Language Semantics*, *8*, 255-290.