Proceedings of the Society for Computation in Linguistics

Volume 4 Article 30

2021

‘Most’ vs ‘More Than Half’: An Alternatives Explanation

Fausto Carcassi
University of Amsterdam, fausto.carcassi@gmail.com

Jakub Szymanik
University of Amsterdam, jakub.szymanik@gmail.com

'Most' vs 'More Than Half': An Alternatives Explanation

Fausto Carcassi **Jakub Szymanik**
Institute for Logic, Language and Computation
Department of Linguistics, University of Amsterdam
Spuistraat 134, 1012 VB Amsterdam
fausto.carcassi@gmail.com jakub.szymanik@gmail.com

Abstract

While 'most' and 'more than half' are generally assumed to be truth-conditionally equivalent, the former is usually interpreted as conveying greater proportions than the latter. Previous work has attempted to explain this difference in terms of pragmatic strengthening or variation in meanings. In this paper, we propose a novel explanation that keeps the truth-conditions equivalence. We support this explanation with a computational model of usage in the Rational Speech Act framework. We find that the difference in typical proportions associated with the two expressions can be explained with previously independently motivated semantic and pragmatic mechanisms.

1 Introduction

According to their standard analysis, sentences 'most As are B' and 'more than half of As are B' (with A and B referring to two sets A and B) are verified by the same As and Bs, i.e. are truth-conditionally equivalent. Specifically, 'most As are B' is analysed as conveying that the size of $A \cap B$ is greater than the size of $A - B$, whereas 'More than half of As are B' is analysed as conveying that the size of $A \cap B$ is greater than half the size of A (Hackl, 2009):[1]

$$[\![\text{most}]\!](A)(B) \iff |A \cap B| > |A - B| \quad (1)$$

$$[\![\text{MTH}]\!](A)(B) \iff |A \cap B| > \frac{1}{2}|A| \quad (2)$$

In contrast to this assumption, the behaviours of 'most' and 'more than half' differ in several ways.

The differences between 'most' and 'more than half' call for an explanation. Early work has focused on the different behaviour of the two expressions with respect to their upper bounds (Ariel, 2003) or their cognitive encoding, which has been argued to lead to different verification procedures (Hackl, 2009). More recent work has focused on the more general fact that 'most' typically conveys proportions higher than 'more than half'. Following Denić and Szymanik (2020), we can categorize the explanations for this difference in two classes.

First, the *pragmatic strengthening hypotheses* claim that while the two expressions are truth-conditionally identical, 'most' is pragmatically strengthened, resulting in a threshold higher than 'more than half'. This strengthening can happen, e.g., through a scalar implicature or through an R-implicature. Solt (2016) argues for the latter option and claims that the reason why 'most' receives a different interpretation to begin with is a (non truth-conditional) difference in the types of scales underlying the two expressions. On the other hand, *lexical meaning hypotheses* attempt to explain these differences in terms of a truth-conditional difference in their logical forms, which can come from, e.g., conventionalization of implicatures or from 'most' being a vague quantifier. More recent work has produced experimental evidence supporting the hypothesis of a semantic difference between the two expressions: Ramotowska et al. (2019) have observed a difference in decision times and behaviour of subjects verifying sentences with the two quantifiers that are consistent with the model in which threshold for "more than half" is 50% and threshold for "most" is higher. Denić and Szymanik (2020) have reported that the thresholds of "most" does not change under the downward monotone environment and remains higher that the threshold for "more than half". This finding also suggest that the difference between the two quantifiers is due to semantics.

In this paper, we propose a novel explanation of the difference between 'most' and 'more than half'. We argue that two independently needed mecha-

[1] In the following, we assume for simplicity that A and B are two finite sets. For a slightly more general definition in terms of measure functions see Solt (2016).

Proceedings of the Society for Computation in Linguistics (SCiL) 2021, pages 334-343.
Held on-line February 14-19, 2021

nisms in the pragmatic interpretation of quantifiers suffice to predict the difference between the two expressions, without assuming a difference in scale structures or truth conditions. The first mechanism is the tendency of the listener to guess points central to a category, in order to minimize the expected distance between their own guess and the speaker's observation. The second mechanism is the structural theory of conceptual alternatives, which lets the alternative set of an utterance depend on the structure of the concept conveyed by the utterance. We show that these mechanisms make the correct predictions with a computational model of pragmatics, the Rational Speech Act model.

2 Solt's account

In order to identify the differences between 'most' and 'more than half', Solt (2016) considers all appearances of the two expressions as quantifiers in the nominal domain in the Corpus of Contemporary American English (COCA) (Davies, 2017). While various differences emerge when comparing the two expressions, in order to compare the typical proportions for which the two expressions are used those appearances were selected which included a specific percentage ($n = 54$ for 'more than half' and $n = 141$ for 'most'). The corpus data shows that (1) 'more than half' is mostly used for percentages in the 50%-65% range, (2) 'most' has a much flatter distribution which covers the whole 50%-100% range.

The difference between the upper bounds and that between the lower bounds of the typical sets of 'most' and 'more than half' receive separate explanations in Solt's account. In order to explain the difference in lower bounds, Solt proposes that the scales that underlie the two expressions are different. According to Solt, 'more than half' uses a *ratio* scale, while 'most' uses an *ordinal* scale.[2] While two points on a ratio scale can be compared in terms of the proportion between them, points on an ordinal scale can only be compared in terms of which one is the greater or lower of the two (Stevens, 1946). Therefore, to say that the scale underlying 'most' is an ordinal scale means that an expression such as 'most A B' only requires us to determine whether the size of $A \cap B$ is greater than the size of $A - B$. On the other hand, 'more than half of As B' requires us to compare the size of $A \cap B$ to a proportion of the size of A, namely $A/2$. Since ratio scales allow arbitrarily precise comparisons between points, Solt's account predicts the lower bound of 'more than half' to be closer to 0.5 than the lower bound of 'most', as observed in the corpus data.

On the other hand, Solt accounts for the difference in the upper bounds of the two expressions with scalar implicatures. Solt points out that 'more than half' has a rich set of alternative utterances, including 'more than two thirds' and 'more than three quarters'. On the other hand, the alternative utterances to 'most' are more sparse, including 'all'. Since the set of alternative utterances is more fine-grained for 'more than half' than for 'most', scalar implicatures constrain the upper bound of the former to be lower than the latter. Solt proposes that the reason why the two expressions have different sets of alternatives is the different types of scales underlying them. In the next section, we introduce an alternative account that explains the difference between 'most' and 'more than half' without assuming a difference in the scales underlying the two expressions.

3 Two mechanisms in the pragmatics of quantifiers

In this section, we present our account in non formal terms. Our account explains the difference in typical set between 'most' and 'more than half', and is based on two phenomena relating to the interpretation of quantifiers. The first is the idea that the listener attempts to minimize the difference between their own guess and the speaker's observation, the second is the fact that different conceptual structures cause different sets of alternatives. We next consider these two mechanisms in turn.

3.1 Distance-minimizing listeners

Some semantic domains, such as nationality, football teams, or personal identity, are not usually structured by relations of similarity. For instance, it is nonsensical to claim that Billy the kid is closer, in terms of his identity, to Jesse James than Doc Holliday.[3] On the other hand, the members of some semantic domains, such as the domain of numbers, colors, or proportions, enter in relations of similarity to each other. For instance, two shades of blue

[2] Technically, Solt argues that 'most' also allows for *semi-ordered* scales. We return to semi-ordered scales below.

[3] There are feature with respect to which two individuals might be more or less close to each other, but this does not concern their identity *as such*.

are closer to each other than either of them is to a shade of red.

In many cases, when communication happens in domains structured by distance, and the listener's task is to construct a representation of the world state given a description produced by the speaker,[4] communicative success is not simply a function of whether the listener's representation is identical or not to the true world state. Rather, success is a(n inverse) function of the similarity between the true state of the world and the listener's guess. In other words, communication is more or less successful depending on whether the listener's guess is more or less dissimilar (respectively) from the true world state.

In this perspective, it is a sensible strategy for a listener to not simply sample from the set of possible world states given their probability after receiving the message, but rather to attempt to minimize the expected distance between their guess and the true world state. For instance, if the speaker utters 'blue', the listener might select a shade of blue that is located around the center of the blue category, because a point near the center of the category will have a lower expected distance to the true world state than a point that is around the margin of the category. Previous literature supports this idea that listeners should tend to guess the center of a category when communicative success depends on the similarity between true state and listener's guess, e.g. Jäger et al. (2011) showed that the optimal strategy for such so-called *simmax* signaling games involves a listener that guesses the central point in the category.

Consistently with previous literature,[5] we will assume in the rest of this paper that communication with quantifiers happens on a semantic domain structured by a distance, namely the scales of proportions and numbers.[6] Moreover, we claim that in communication with quantifiers, communicative success is of the graded type presented above. For instance, if 1/2 of the As are B, the communication is more successful if the listener guesses $|A \cap B|/|A| = 0.6$ than if the speaker guesses 0.9. This implies that a rational listener does not guess a proportion simply by sampling from the posterior over proportions conditional on the received signal, but rather they attempt to minimize the expected distance between their guess and the true state of the world.

The tendency for the listener to guess a state that minimizes the expected distance to the speaker's observation, when in a scalar semantic domain, is not only a result about rational agents, but also aligns with the way we use quantifiers in practice. For instance, imagine receiving the signal 'between 50 and 100', and creating a representation of the world state. Even within the part of the scale of integers covered by the expression—e.g. numbers between 50 and 100—the guess does not happen uniformly. Rather, we intuitively tend to guess an integer that is around the center of the category, i.e. around 75. In other words, we are less likely to select a number close to the category boundaries, such as 99. As we discuss in more detail below, the situation is subtler when multiple possible utterances are involved.

3.2 The structural account of alternatives

As discussed above, Solt's explanation for the different comparison sets of 'most' and 'more than half', which produces in her account the different upper bounds, is based on the difference between the scale types. We propose an alternative explanation for the difference in the sets of alternative utterances which does not rely on scale structure. In particular, we rely on the *structural account of alternatives* (Katzir, 2007; Fox and Katzir, 2011; Trinh and Haida, 2015) to explain why 'most' and 'more than half' have different sets of alternative utterances.

The structural theory of alternatives starts with the idea of a structural alternative. ψ is a structural alternative to ϕ ($\psi \lesssim \phi$) iff ψ is structurally at most as complex as ϕ, i.e. ψ can be obtained from ϕ through a "finite series of deletions, contractions, and replacements of constituents of ϕ" with constituents of the same category taken from the lexicon (Katzir, 2007). The core idea is to define the set $A_{str}(\phi)$ of utterances alternative to ϕ as follows:

$$A_{str}(\phi) = \{\psi | \psi \lesssim \phi\} \tag{3}$$

In words, the set of utterances that enter in the calculation of implicatures for ϕ is the set of utterances that are structurally at most as complex as ϕ.

[4] This communicative setup is called *descriptive* by Franke (2014), who opposes it to *referential* communicative games. In the following, we limit ourselves to discussions of descriptive communication.

[5] See chapter 2 of Carcassi (2020) for an overview.

[6] In the following, we will focus on the scale of proportions, but what we say can be easily generalized to the scale of numbers.

While the original criterion for alternatives in Katzir (2007) is a syntactic one, there is emerging theoretical and experimental evidence that it is best characterized as acting not on the syntactic structure, but rather on the *conceptual* structure of utterances (Chemla, 2007; Buccola et al., 2018). The structural account of conceptual alternatives is still being developed, however in the following we limit ourselves to using the basic idea in equation 3, as it is all we need for the present purposes.

In the following, we make two crucial assumptions about the way alternatives are generated for the expressions under consideration. First, not every expression of the form 'a b' is considered, where a is a cardinal number and b an ordinal number such that $a \leq b$. If every a and b were considered, the set of alternatives to 'one half' would be the set of rational numbers in the unit interval. Various factors plausibly restrict the set of considered numbers. First, the listener can generally assume that the speaker has noisy measurement of the true proportion, and therefore only produces utterances implying at most a certain level of granularity. Moreover, the communicative aims generally do not require transmission of precise proportions. The idea that the notion of alternative is graded and depends on the complexity of the concept, developed in Buccola et al. (2018), could also account for why 'five sixth' seems to compete with simpler fractions such as 'two thirds', while the opposite is not true; the difference would depend on the different conceptual complexity of different fractions. We do not develop this idea further in the present paper.[7]

The second assumption we make is that 'most' and 'more than half' are conceptually structured as proposed by Hackl (2009), i.e. as in equations 1 and 2 above. The main consequence of this assumption is that 'two thirds' and structurally equivalent expressions are alternatives to 'half' according to the criterion in equation 3.

Under the two assumptions just discussed, the criterion defined in equation 3 has the correct consequences for the cases at hand. Namely, A_{str}('most') contains 'all' and does not contain 'more than three quarters'. On the other hand, A_{str}('(one) half') contains e.g. 'three quarters'.

In this section, we have presented two mechanisms that play a role in the way quantifiers are interpreted. These two mechanisms have already been discussed in the literature in other contexts. While our approach relies on a difference in conceptual structures, the account remains a pragmatic account insofar as conceptual structure affects the results only indirectly by causing a difference in alternative sets, rather than directly as in Solt (2016). The main contribution of this paper is therefore to show how these two mechanisms can, without recourse to the scale types discussed by Solt, explain the difference between 'most' and 'more than half' with respect to the proportions that the two expressions typically convey.

In the next section, we model the mechanisms we discussed. We consider a pragmatic speaker who picks 'most' or 'more than half' not simply as a function of their extension on the scale of proportions, but instead also implicitly selecting the set of alternatives that will allow a pragmatic listener to choose a proportion that is as close as possible to the speaker's observation. Since a pragmatic listener guesses points closer to 0.5 for a rich alternative set such as the one induced by 'more than half', the speaker selects 'more than half' for such proportions. On the other hand, since the listener will guess alternatives close to 0.75 for 'most', the speaker produces 'more than half' for such proportions. In order to formalize this intuition, in the next section we present the RSA modelling framework for pragmatic language use.

4 An RSA model of the two mechanisms

4.1 Basic RSA model

The RSA framework is meant to model the process of recursive mindreading that lies behind the pragmatic interpretation or production of utterances (Goodman and Stuhlmüller, 2013; Franke, 2014; Frank et al., 2017). RSA models usually start with a pragmatic listener who interprets utterances based on the simulated behaviour of a pragmatic speaker. The pragmatic speaker in turn given an observation tends to choose the most useful utterance for a literal listener who interprets it based solely on its literal meaning. We will first explain the simplest type of RSA model, and then a modification that will be useful to model numerals.

The simplest RSA model starts with a set of utterances u and a set of possible states s. The meaning of each utterance can be encoded as the set of those states that verify the utterance. The pragmatic listener L_1 receives an utterance u and calculates a

[7]Thanks to Milica Denić for recommending the literature on conceptual alternatives.

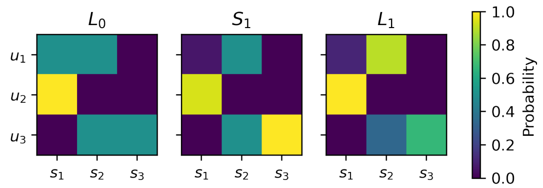

Figure 1: Simple RSA model with three possible utterances u (y-axis) and three states s (x-axis). L_1 calculates a scalar implicature for utterances u_1 and u_2 ($\alpha = 4$). The left, central, and right plots correspond to L_0, S_1, and L_1 respectively. Note that the color indicates the probability of guessing a state given a signal for L_0 and L_1, and the probability of producing a signal given a state for S_1.

posterior over states by Bayesian update, combining their prior over states with the probability that the pragmatic speaker S_1 would have produced the utterance given each state:

$$p_{L_1}(s|u) \propto p_{L_1}(s) p_{S_1}(u|s) \tag{4}$$

The pragmatic speaker in turn observes a state and produces an utterance that aims at optimizing the utility $U(u|s)$ for a literal listener L_0 given the state, while minimizing the utterance cost:

$$p_{S_1}(u|s) \propto \exp(\alpha U(u|s) - c(u)) \tag{5}$$

The utility $U(u|s)$ is the negative surprisal of the state given the utterance, so that the speaker favours utterances that make the state less surprising for the literal listener:

$$U(u|s) = \log(p_{L_0}(s|u)) \tag{6}$$

Finally, the probability that literal listener L_0 attributes to each state given an utterance is simply 0 if the utterance is not verified by the state, and proportional to the prior for the state otherwise:

$$p_{L_0}(s|u) \propto \begin{cases} p(s) & \text{if } s \text{ verifies } u \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

Figure 1 shows L_0, S_1, and L_1 in this simple RSA model. The crucial phenomenon that can be observed in figure 1 is that L_1 calculates a scalar implicature: although utterance u_1 is, in its literal sense, compatible with both s_1 and s_2, S_1 tends to produce u_1 mostly for s_2, because when s_1 is observed S_1 tends to use the more useful signal u_1. Therefore, when hearing u_1 L_1 is more likely to guess s_2.

4.2 Distance based listeners

In the simple RSA models above, the success of communication is binary, solely a function of whether the listener's guess coincides with the speaker's observed state. This is plausible in cases where the set of states has no internal structure. However, as discussed above in the case where a notion of distance is well-defined on the set of states, the listener might not be simply trying to guess the speaker's observation, but rather might strive to minimize the (expected) distance between the state they select and the speaker's observation.[8]

In order to model the effects of a well-defined distance D on the set of states, we modify the listener L_1 so that instead of selecting a state by sampling from their posterior distribution given the signal, they try to minimize the expected distance between their selection s and the true state. Therefore, we define the choice probability for listeners as follows:[9]

$$p_{CL}(s|u) \propto \exp\left(-\rho \sum_{i \in \text{states}} p_{L_1}(i|u) D(i, s)\right) \tag{8}$$

where ρ is the parameter of a softmax function which determines how strongly the listener tends to minimize the expected distance and p_{L_1} is defined as above in equation 4. The listener described in equation 8 tends therefore to minimize the expected linear distance function. Figure 2 shows the effects of this modification of the model for 20 states, when $D(s_n, s_m) = |n - m|$. The right plot shows that in this modified RSA model, L_1 tends to guess points that are located centrally in the category, after the category has been restricted by scalar implicature.

4.3 Varying sets of alternatives

The modification to the basic RSA model above is an implementation of the first mechanism discussed in section 3. The second mechanism con-

[8] Cf previous work where the effects of a distance structure affects the speaker but not the listener, such as Franke (2014). It is worth noticing that our model does not have more degrees of freedom than Franke's model: while we introduce one more parameter than the basic RSA model to regulate the listener's tendency to minimize expected distance, Franke introduces one parameter to regulate the amount of pragmatic slack. We do not investigate in this work the differences between the two approaches.

[9] We apply this modification only to L_1, assuming that the attempt to minimize distance is something above and beyond the literal reading of the signals. We leave to future work an investigation of the effects of modifying both listeners.

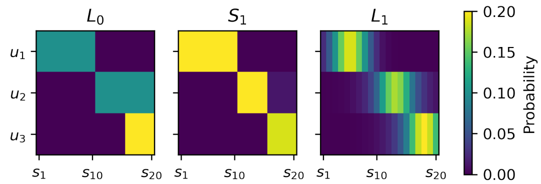

Figure 2: RSA model with a distance-minimizing L_1. The model displayed in the plot use a language with three utterances and 20 states. The listener L_1 do not simply guess the signal observed by the speaker by sampling their posterior, but rather attempt to minimize the expected distance between their guess and the speaker's observation ($\alpha = 4, \rho = 2$). See figure 1 for more detail.

L_0	Receives a signal and guesses a state given the literal meaning of the signal.
S_1	Given a state, tends to select a signal with a high utility for the literal listener L_0.
L_1	Tends to minimize the distance between own guess and S_1's observation, given the behaviour of S_1. Alternative utterance set depends on S_2's signal.
S_2	Like S_1, but utility for L_1. Implicitly selects a set of alternatives for L_1.

Figure 3: Structure of modified RSA model. The set of alternative utterances considered by L_1 is not fixed, but rather depends on the received utterance. Moreover, L_0 and L_1 do not simply guess a state based on their posterior probability given the received signal, but rather tend to guess a state that is expected to be close to the speaker's observation.

cerns the way that the comparison set depends on the speaker's utterance.

In the basic RSA framework, the set of possible utterances considered by the pragmatic speaker and the pragmatic listener that the speaker models are identical. However, according to the structural account of alternatives discussed above the set of utterances considered by the listener depends on the actual utterance that is picked by the speaker. For instance, if the speaker utters '101', the listener will consider all alternative utterances that are at most at a similar level of granularity as 101, such as 91 and 100. However, if the speaker utters '100', the listener in the model considers an alternatives set containing e.g. only 90 and 100, but not 101.

In order to model this in the RSA model, we introduce a speaker S_2. S_2, much like S_1, tends to select the signal that minimizes the listener's surprise for the real state given the signal. However, the set of alternative utterances considered by L_1 is not independent of the signal received by L_1. Instead, the set of alternative utterances considered by L_1 (and therefore by the lower levels S_1 and L_0) depends on the actual utterance of S_2, as described in the section on the structural account of alternatives above. The model below is therefore a production model.

This picture of alternatives is in many respects a simplification. For instance, it is likely that from the point of the listener there is uncertainty as to the set of alternatives that ought to be considered in the context. More complex discussions of issues related to granularity and alternatives can be found in the literature, see e.g. Bastiaanse (2011) for numerals. However, these more complex models are not needed to explain the issue at hand, and therefore we leave investigation of the subtleties to future work.

In sum, the only requirements for this model to apply are (1) that the listener is trying to minimise the distance between their guess and the speaker's observation, and (2) that some terms have an alternatives set that is more granular than other ones. In particular, it applies even if the two expressions that induce different granularities are synonymous.

In this section, we have formalized the two mechanisms discussed in section 3 within the RSA framework. The resulting model is summarized in natural language in figure 3. In the resulting model, structurally different expressions induce the pragmatic listener to consider different sets of alternative utterances. Moreover, the listener does not simply guess uniformly from the enriched part of the parameter space, but rather tends to guess points that are central in the pragmatically enriched category. Therefore, even intensionally equivalent expressions will be used differently, as long as they are structurally different. In the next section, we show how this model applies to the specific case of the contrast between 'most' and 'more than half'.

5 An alternatives account of 'most' vs 'most than half'

5.1 An RSA model of the contrast

In the following, we will model communication with quantifiers by applying the RSA model described above to the following simple referential communication task, modelled after Pezzelle et al.

Utterance	Extension
All ($\forall$)	$\{1\}$
Most	$(1/2, 1]$
None ($\neg\exists$)	$\{0\}$
Some ($\exists$)	$(0, 1]$
MT a half ($> 1/2$)	$(1/2, 1]$
MT one third ($> 1/3$)	$(1/3, 1]$
MT two thirds ($> 2/3$)	$(2/3, 1]$
LT a half ($< 1/2$)	$[0, 1/2)$
LT one third ($< 1/3$)	$[0, 1/3)$
LT two thirds ($< 2/3$)	$[0, 2/3)$

Table 1: Meaning of each signal in the model. MT='more than', LT='less than'.

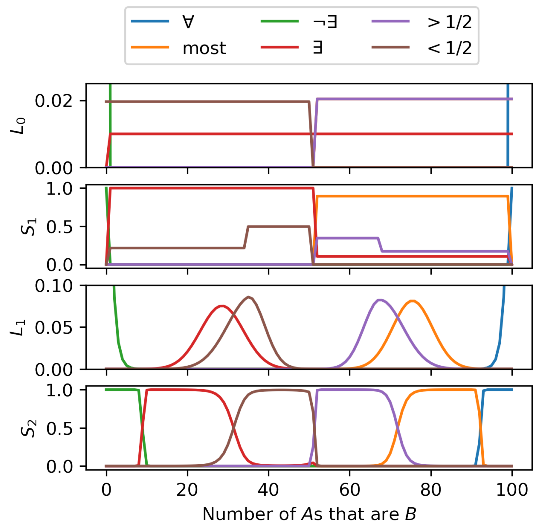

Figure 4: The plots shows the results with $|A| = 100$, $\alpha = 3$ and $\rho = 1$. Each plot shows the behaviour of a different agent in the RSA model.

(2018) where communication was set up similarly in a production task. A speaker observes two sets, A and B, and attempts to communicate to a listener which proportion of A is also in B in the way modelled by the modified RSA model introduced above. As possible signals, we have chosen the lexically simple Aristotelian quantifiers and some minimal set of alternatives for 'more than half'. The literal meaning of each quantifier in the model corresponds to a portion of the scale of proportions (see table 1). This set of utterances is closed under substitution of 'more' by 'less', and by (semantically meaningful) substitutions of one, two, and three (both their cardinal and ordinal versions) with each other.

As in the modified RSA model presented above, the alternatives considered by the pragmatic listener depend on the speaker's utterance. For instance, if the speaker uttered 'some' the listener would consider a set of alternatives containing 'all' but not 'more than two thirds', while if the speaker uttered 'more than one third' both 'all' and 'more than two thirds' would be possible options for the listener. In the present case, the utterances above can be divided in two groups, the first containing 'all', 'most', 'none', and 'some', and the second containing the remaining utterances. Each utterance in the first group contains all other utterances in that group as alternatives, and none of the utterances in the second group. Each of the utterances in the second group contains all utterances in its set of alternatives.[10]

In order to isolate the effects of the account of alternatives discussed above from the consequences of utterance cost, we assume that signals have no cost. Moreover, to keep the results as simple as possible S_2 can only produce 'all', 'most', 'none', 'some', 'more than a half' and 'less than a half'. In order for the speaker to be able to calculate a distribution over utterances given any state, there has to be at least one utterance to refer to each state in each set of alternative utterances.

The results of the model are shown in figure 4. L_0 guesses uniformly within the categories expressed by each signal considered by S_2. L_0 treats 'most' and 'more than half' identically, guessing uniformly among the states between 51 and 100. Finally, S_0 selects the maximum for 'every' and the minimum for 'none'.

With S_1, the set of alternatives for each signal matters (second plot from top in figure 4). More specifically, while the lower bound for both 'most' and 'more than half' are similar for S_1, their upper bounds are different as a consequence of the different ways that the respective set of alternatives cover the scale. 'More than half' implicates less than two thirds, and therefore tends to not be used for proportions higher than two thirds, while most only implicates 'not all'. Note that while the six signals are plotted together in figure 4, the distribution for each signal is computed independently with a possibly different set of alternatives utterances. Therefore, S_1 does not suffice to explain the difference between 'most' and 'more than half'.

[10] While previous work has argued that utterances with different monotonicity profiles do not appear in the same set of alternatives (e.g. Horn, 1989), Katzir (2007) has argued the structural theory of alternatives can lift this restriction.

L_1 tends to pick the central point in the categories as produced by S_1 (third plot from top in figure 4). Therefore, L_1 tends to guess points closer to the middle of the scale for 'more than half' than for 'most', because the former is produced by S_1 for a range of proportions closer to the scale's midpoint. Finally, the pragmatic speaker S_2 tends to pick 'more than half' for signals closer to the midpoint of the scale than 'most' (bottom plot in figure 4).

The results in figure 4, while qualitatively correct, are quantitatively surprising in two respects. First, the upper bound of 'most' is lower in the results shown in figure 4 than in the data presented by Solt. This is a consequence of the unusually low bound for 'all', which goes down to 80%. The exact value of this bound changes depending on the exact set of utterances available to the speaker. For instance, the availability of 'almost all' would push the lower bound for 'all' higher up on the scale. The second difference is that the upper bound of 'more than half' goes higher than in the data presented by Solt. The positions of the involved bounds are sensitive to the parameter values. For instance, figure 5 shows a parameter setting that makes predictions closes to Solt's data. Moreover, the upper bound of 'more than half' is also sensitive to the set of alternatives. For instance, including proportions such as 'three fourths' would push the upper bound of 'more than half' closer to 0.5.

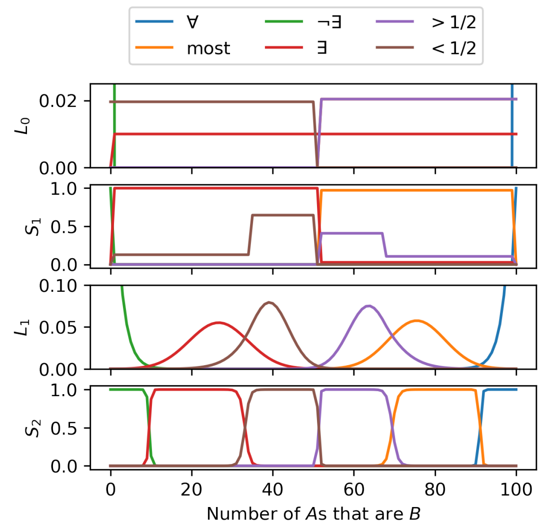

Figure 5: The plots shows the results with $|A| = 100$, $\alpha = 5$ and $\rho = 0.5$. This parameter setting produces results closer to Solt's observed typical sets for 'most' and 'more than half'.

5.2 Ignoring the mechanisms

In order to see what role each of the two mechanisms play in the predictions of the model above, it is instructive to observe the consequences of ignoring each of the two mechanisms.

When both mechanisms are ignored, the result is a simple RSA model as described in 4.1. In particular, we consider an RSA model with different costs for the signals. The results are shown in figure 6. When both mechanisms are ignored and production costs are implemented, speaker S_2 does not introduce any substantial innovation over the listener L_1, and therefore we stop the computation at the level of L_1. The effect of cost in this setting is simply to make S_1's production probability for 'more than half' uniformly lower than the one of 'most' for any given state. However, the difference in cost cannot be exploited by L_1 to draw an inference about the state observed by S_1. In sum, a difference in cost alone without the mechanisms discussed above cannot be exploited by a pragmatic speaker to convey different information with truth-conditionally equivalent signals. For similar reasons, when only the first mechanism, namely the structural account of alternatives, is ignored, all utterances compete with each other, and therefore S_2 does not introduce interesting results. Again, since the symmetry is not broken by the different set of alternatives, 'most' and 'more than half' end up conveying identical information to S_1.

When only the second mechanism—the distance-minimizing listener—is ignored, we still obtain the crucial result that 'more than half' is generally used to convey proportions closer to the scale's midpoint than 'half'. However, the results differ from section 4 in two crucial ways. First, the speaker S_2 ends up producing each signal with uniform probability within the pragmatically enriched category, as shown in figure 8. For instance, 'more than half' is produced with uniform probability for proportions above 1/2 and below 1/3. Second, the model without distance-minimizing L_2 predicts that the speaker would use 'all' and 'none' exclusively for the maximum and minimum of the scale respectively. These two consequences contradict both Solt's data and the data in Pezzelle et al. (2018). In the data, the production probabilities for 'more than half' resembles a Gaussian distribution rather than a uniform distribution, 'all' and 'none' are

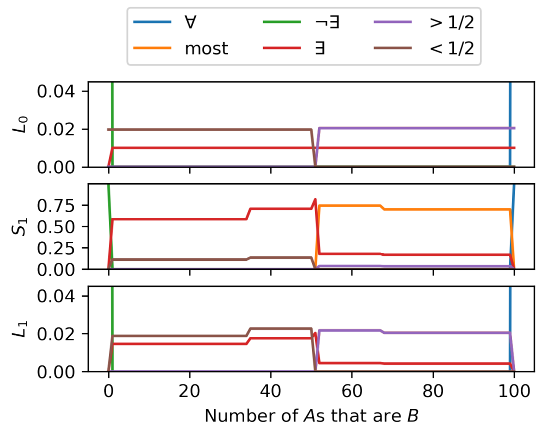

Figure 6: Results without both mechanisms. We stop the computation at the level of pragmatic listener L_1. S_1 is less likely to produce ‘more than half’ for any given state, because of its higher cost. However, L_1 does not derive any difference between the information conveyed by ‘most’ and ‘more than half’, so that the two lines perfectly overlap for L_1. In this plot, $\alpha = 2$. While the speaker in this model can produce all signals, for ease of comparison with the previous plots we only plot the 6 signals that the speaker could produce in the previous models. We model the cost of each utterance simply as the number of words in the utterance: ‘all’, ‘most’, ‘none’, and ‘some’ get cost 1, while all other signals get cost 4.

used for signals close to the scale’s extremes rather than exclusively to the extremes.

In this section, we have shown that the two mechanisms discussed in section 3 are not only independently motivated, but are also both needed to make sense of the difference in typical sets between ‘most’ and ‘more than half’.

6 Conclusions

‘Most’ and ‘more than half’, while traditionally assumed to be truth-conditionally equivalent, are typically associated with different proportions. In the most developed explanation of this difference, Solt (2016) introduces a difference between the structures of the scales used by the two expressions. In contrast, in this paper we proposed a novel account of the difference that is based on independently motivated mechanisms and does not rely on different scale structures. Moreover, we analysed the predictions of the account by implementing it in a popular computational model of pragmatic reasoning, the RSA model. Further and especially experimental work is needed to compare the empirical accuracy of the different emerging accounts

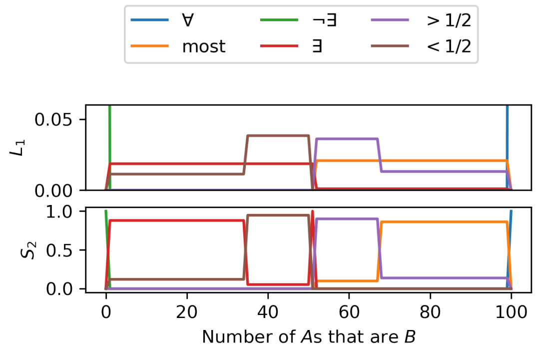

Figure 7: No distance minimizing.

Figure 8: Results without the second mechanism: when the distance-minimizing listener is substituted with the pragmatic listener of equation 4, speaker S_2 has a uniform probability of producing ‘more than half’ across its whole pragmatically enriched domain. Only L_1 and S_2 are plotted as L_0 and S_1 are essentially the same as in figure 4.

of the differences between ‘most’ and ‘more than half’.

The work we presented can be extended in various possible directions. First, a similar model could be used to account for the usage of modifier numerals, since a similar contrast to the one discussed here can be found e.g. between ‘more than 100’ and ‘more than 101’, where the typical guessed number for the former utterance is higher than the for the latter. Second, the model above can be used to fit experimental production data with quantifiers. A third possible development would look at whether the model predicts that the quantifiers’ thresholds stay the same even in downward entailing contexts, as suggested by the experimental data in Denić and Szymanik (2020). Lastly, in the models presented in this paper we only considered a small set of alternative proportions for ‘half’, namely ‘one third’ and ‘two thirds’. However, it would be valuable to study the predictions of the model when more alternative utterances, containing more complex proportions, are included. We leave all these exciting possible developments to future work.

Acknowledgments

The authors have received funding from the European Research Council under the European Union’s Seventh Framework Programme (FP/2007–2013)/ERC Grant Agreement n. STG

716230 CoSaQ.

References

Mira Ariel. 2003. Does most mean ‘more than half’? *Annual Meeting of the Berkeley Linguistics Society*, 29(1):17–30.

Harald Bastiaanse. 2011. The Rationality of Round Interpretation. In *Vagueness in Communication*, Lecture Notes in Computer Science, pages 37–50, Berlin, Heidelberg. Springer.

Brian Buccola, Manuel Križ, and Emmanuel Chemla. 2018. Conceptual alternatives: Competition in language and beyond.

Fausto Carcassi. 2020. *The Cultural Evolution of Scalar Categorization*. PhD Thesis, University of Edinburgh, Edinburgh.

Emmanuel Chemla. 2007. French both: a gap in the theory of antipresupposition.

Mark Davies. 2017. Corpus of Contemporary American English (COCA).

Milica Denić and Jakub Szymanik. 2020. Differences in thresholds between 'most' and 'more than half': Semantics or pragmatics? In *Sinn Und Bedeutung 25*, London. OSF.

Danny Fox and Roni Katzir. 2011. On the characterization of alternatives. *Natural Language Semantics*, 19(1):87–107.

Michael C. Frank, Andrés Gómez Emilsson, Benjamin Peloquin, Noah D. Goodman, and Christopher Potts. 2017. Rational speech act models of pragmatic reasoning in reference games.

Michael Franke. 2014. Typical use of quantifiers: A probabilistic speaker model. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 36(36).

Noah D. Goodman and Andreas Stuhlmüller. 2013. Knowledge and Implicature: Modeling Language Understanding as Social Cognition. *Topics in Cognitive Science*, 5(1):173–184.

Martin Hackl. 2009. On the grammar and processing of proportional quantifiers: Most versus more than half. *Natural Language Semantics*, 17(1):63–98.

Laurence R. Horn. 1989. *A Natural History of Negation*. University of Chicago Press, Chicago.

Gerhard Jäger, Lars P. Metzger, and Frank Riedel. 2011. Voronoi languages: Equilibria in cheap-talk games with high-dimensional types and few signals. *Games and Economic Behavior*, 73(2):517–537.

Roni Katzir. 2007. Structurally-defined alternatives. *Linguistics and Philosophy*, 30(6):669–690.

Sandro Pezzelle, Raffaella Bernardi, and Manuela Piazza. 2018. Probing the mental representation of quantifiers. *Cognition*, 181:117–126.

Sonia Ramotowska, Shane Steinert-Threlkeld, Leendert van Maanen, and Jakub Szymanik. 2019. Most, but not more than half, is proportion-dependent and sensitive to individual differences. In *Proceedings of Sinn Und Bedeutung 24*, pages 165–182, Osnabrueck.

Stephanie Solt. 2016. On measurement and quantification: The case of most and more than half. *Language*, 92(1):65–100.

S. S. Stevens. 1946. On the Theory of Scales of Measurement. *Science*, 103(2684):677–680.

Tue Trinh and Andreas Haida. 2015. Constraining the derivation of alternatives. *Natural Language Semantics*, 23(4):249–270.