Vagueness as Cost Reduction: An Empirical Test

Matt Green (mjgreen@abdn.ac.uk)
Department of Computing Science, Meston Building
Aberdeen, AB24 3UE UK

Kees van Deemter (k.vdeemter@abdn.ac.uk)
Department of Computing Science, Meston Building
Aberdeen, AB24 3UE UK

Abstract

NLG systems that generate text from numerical data must decide between alternative linguistic forms of the given numerical content, such as whether to use a precise or a vague expression. Currently there is little empirical data for these systems to draw on when making these decisions. We performed experiments with human readers in which participants responded to instructions in the form of referring expressions, where we manipulated whether the instruction used a vague or a crisp referring expression, in order to test the hypothesis that vagueness reduces processing costs for the comprehender. Results indicate that people respond more quickly and accurately to vague linguistic expressions than to crisp numerical expressions, but that this benefit also accrues to precise terms that avoid numbers.

Keywords: referring expressions; empirical; vagueness; cost reduction

Introduction

A working definition of vagueness Vagueness can be defined at the level of the word (*tall*) or the referring expression (*Bill is the tall guy over there*). At the word level, we consider that a word is vague if the set of objects that it denotes is not well-defined (Lipman, 2009). For example, *tall* is a vague word because the set of heights that it denotes is not well-defined. Support for this comes from the intuitive observation that there are some heights for which it is not clear whether the height is tall or not-tall. Keefe and Smith give what can be regarded as a widely accepted definition of vagueness: “vague predicates have borderline cases, have fuzzy boundaries, and are susceptible to sorites paradoxes” (Keefe & Smith, 1996). At the level of the referring expression, vagueness has to be defined in context. Referring expressions of quantity can be made at different levels of precision. Suppose, for example, that we and a friend are building a table in a workshop, and we need our friend to pass us one of two trays. One tray has 6 nails on it; the other tray is the one that we need. We could express our request saying (a) “Could you pass me a tray with *n* nails?”, where *n* is the correct number, or (b) “Could you pass me a tray with many nails?” The first description uses an exact number. The second description is less precise: it seems clear that a tray with 100 nails in it could be referred to successfully with (b), and that a tray with only one nail in it could not. However it is not so clear whether (b) could felicitously refer to a tray with, say, 7 nails in it. The uncertainty is whether our friend would consider 7 nails “many” nails. Keefe and Smith (and many others before and after them) would say that the tray with 7 nails is a *borderline case* of a tray with many nails, in this situation. Whether a borderline case exists for an expression of quantity is a kind of litmus test of vagueness, to indicate that the expression does not describe a well-defined set of objects.

The utility of vagueness: Vagueness as cost reduction Standard game-theoretic models of communication (e.g., Crawford & Sobel, 1982) predict that a precise term is always preferable to a vague term when the goals of speaker and listener are in alignment. However, this prediction is at odds with the observation that much of the language that we actually use is vague. This contradiction suggests that vagueness has some utility that is not properly accounted for in such models, leading Lipman to pose the question why vagueness is so common in the language that we use everyday (Lipman, 2009). Van Deemter (2009) suggests the following possible sources of the utility of vagueness in co-operative situations. One is *local vagueness*, where an apparently vague term (like *the tall guy*) is rendered precise in context (like the context of collecting someone from the airport whom one has not met before), which may save the speaker or hearer the effort of making a measurement. Sometimes a precise term (like *a temperature of 39.82 Celcius*) can convey extraneous detail: use of a vague term like *a high temperature* here constitutes *information reduction* and avoids placing an unnecessary processing load on the hearer. Vague terms sometimes convey elements of evaluation (e.g., *expensive*). In such expressions, the evaluative element of the information may constitute utility that a precise expression would lack (Veltman, 2002). A possible benefit for vagueness in cooperative situations is that vague expressions may be easier for the hearer to process than precise expressions. For example, Lipman (2009) writes: “For the listener, information which is too specific may require more effort to analyze”. We shall refer to this characterisation of the utility of vague language as the ‘cost reduction hypothesis’, and it is this notion that we explore in this paper (see also, van Rooij, 2003; Lipman, 2009, for related ‘cost-based’ arguments).

Applied vagueness: the case of Natural Language Generation (NLG) NLG systems that generate text from numerical input must decide between alternative linguistic forms of the same numerical content, including how vaguly to express the information. Such systems include Goldberg, Driedger, & Kittredge (1994); Turner, Sripada, Reiter, & Davy (2006) in the domain of weather forecasting, given data such as temperature and wind speed; Hripcsak, Elhadad, Chen, Zhou, & Morrison (2009); Portet et al. (2009) in the domain of medical decision support, given numerical data such as oxygen saturation, heart rhythm, etc. Currently there is little empirical data to support NLG decisions about vagueness.

Related empirical work on vagueness Peters et al. (2009) carried out a series of studies where participants were required to rate hospitals based on various sources of information about quality of care. The format of the information was manipulated within subjects: either numbers only were presented, or numbers and evaluative categories (e.g., *poor*, *fair*, *good*, *excellent*, with visual boundary lines between the categories). Certainly, terms like *fair* admit the possibility of borderline cases. However, when the visual boundary lines are taken into account, which map the terms to exact ranges, it becomes immediately doubtful whether any borderline cases could be conceived to arise in fact. For example *fair* is mapped to 60 – 70%: a crisp range.

Bisantz, Marsiglio, & Munch (2005) investigated the effects of different representations of probabilistic information. The task was a simulated stock purchase carried out online over three days. Participants made bets on stock that was either profitable or unprofitable, basing their decisions on continuously updated information that they were supplied with about the probability that a stock would be profitable. They manipulated display format and specificity level. An example of their *linguistic expression* display format condition is *fairly unlikely*. This condition is successfully vague in the sense of admitting borderline cases, but Bisantz et al. found no effect of this vagueness.

Mishra, Mishra, & Shiv (2011) manipulated the presentation format of information about quantities. They compared information presented as a range with information presented as a single value at the midpoint of the range (e.g., *between* 0.5 *and* 1.5*g of cocoa* versus 1*g of cocoa*). Information presented as a range of values was considered to be vague information. We note that information presented as an exact range of values does not conform with the definition of vagueness that we use here borrowed from Keefe and Smith (1996), since an exact range does not have the potential to admit borderline cases. We would prefer to call the difference between a range and a single midpoint value a difference of *granularity* or of *specificity*.

Experiments

We used a forced choice paradigm to elicit response times and error rates in a task where participants were issued an instruction in the form of a referring expression to select a target object from among distractors. Objects were squares on screen, containing varying numbers of dots arranged in a pattern that was randomised per-trial. Responses were made via the keyboard, and button-press latency and response accuracy were our dependent variables. Participants were students and staff members at the University of Aberdeen, aged between 18 and 40: there were 25 participants in experiment 1; 20 in experiment 2; and 30 in experiment 3.

Experiment One

Introduction We hypothesised that participants would respond more quickly, and make fewer errors, in the vague conditions than the precise conditions. This is in line with the hypothesis that vagueness reduces costs for the comprehender.

Method This pilot experiment used small numbers of dots (fewer than 10 in each of two squares). Pairs of numbers were $\{2,4\}$, $\{2,6\}$, $\{3,5\}$, $\{5,9\}$, $\{6,8\}$, $\{7,3\}$, $\{7,9\}$, $\{8,4\}$; presented equally often in each left-to-right order. An example stimulus is Figure 1. Per-trial instructions used a referring expression to indicate one of the squares with reference to the number of dots it contained, varying whether the instruction used a crisp or a vague expression of quantity. An example of a vague instruction is *Choose the square with many dots*. An example of a crisp instruction is *Choose the square with seven dots*. We wrote each number out in full (seven) rather than as a numeral (7) because the number words involved did not differ much in length, and therefore their length would not contribute much variance to the response measures. At the start of the experiment, participants were asked to respond quickly while avoiding errors. The apparatus, for all our experiments, was a MacBook Pro laptop with a 13 inch screen. The stimuli were presented using GNU Octave (Eaton, 2002) and the Psychophysics Toolbox extensions (Brainard, 1997).

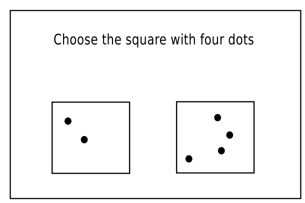

Figure 1: Stimulus, experiment 1

Results Response time data were preprocessed to remove outliers and erroneous responses. One trial (0.06% of the data) was removed because of a programming error that yielded a negative response time. Three trials (0.18% of the remaining data) were removed because response times were greater than our upper limit of $10,000$ ms. No response was less than 250 ms. 40 trials (2.5% of the remaining data) were removed due to incorrect responses. For the analysis of error rates, no trials were discarded. One participant was removed due to not having data in all subconditions. All told, 6.9% of the trials were removed for the analysis of response times. There is evidence to suggest that the numerosity of very small numbers is recognized by a different cognitive mechanism than the numerosity of larger numbers, and that very small numbers are recognized extremely quickly without the need for enumeration (see for example Trick & Pylyshyn, 1994). That research uses the term "subitizable" to indicate that a number is recognized in this way, and we adopt the same nomenclature. Under this scheme, our numbers $1,2,3$, and 4 are subitizable, and our numbers $5,6,7,8$ and 9 are not subitizable. Results were different depending on whether either of the squares contained a subitizable number of dots. The three-way interaction between subitizability, quantity, and vagueness was reliable ($F_{(1,23)} = 6.65, p < .05$). When one square did contain a subitizable number of dots, the task was accomplished faster, and more accurately, when the instruction used a precise quantifier than when it used a vague quantifier (see Figure 2). When both numbers of dots were above the subitizable range, participants were reliably faster to respond to vague quantifiers (e.g., *many dots*) than precise quantifiers (e.g., *nine dots*), but only when the instruction identified the larger of the two numbers.

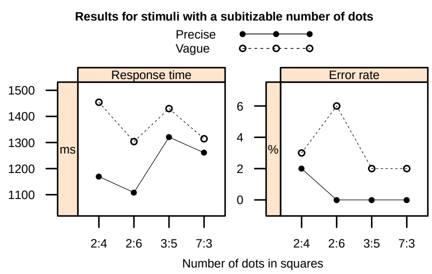

Figure 2: Results, experiment 1

Discussion We were unable to reject the null hypothesis that there was no difference between response times and error rates for the vague and precise instructions. Given the different patterns of results according to whether the stimulus included a subitizable number of dots or not, the null effect seems likely to be due to opposite patterns cancelling each other out. Because stimuli with a subitizable number of dots appear to represent a special case due to the specialised cognitive routines that identify these numerosities, we decided to investigate only non-subitizable numbers in the subsequent experiments.

Experiment Two

Introduction We hypothesised that vague instructions would attract faster response times and fewer errors than the crisp instructions. This is in line with the cost reduction hypothesis.

Method The second experiment used larger numbers of dots and two squares in a similar experiment to the pilot experiment. Pairs of numbers were: {5, 25} {10, 25} {15, 25} {20, 25} {30, 25} {35, 25} {40, 25}, and the same pairs in the other left-to-right-order. All these numbers were above the subitizable range. Per-trial instructions used a referring expression to indicate one of the two squares with reference to the number of dots it contained. An example of a vague instruction is *Choose the square with few dots*. An example of a crisp instruction is *Choose the square with* 45 *dots*. We wrote each number as a numeral (45) rather than in full (forty-five) in this experiment, because the numbers involved were larger than in the previous experiment, and we did not want participants responses to be slowed by longer reading times for the instructions that had longer words for the number. At the start of the experiment, participants were asked to respond quickly while avoiding errors. Experiment two attempted to get a cleaner index of decision time than the pilot experiment, by separating off the time taken to read the instruction from the time taken to choose a square. Participants were required to press a key after reading the instruction, after which a central fixation point was displayed for one second, followed by the stimulus. See Fig (3) for an example instruction and stimulus. RT comprised only the latency between the presentation of the stimuli, and the keypress identifying the decision.

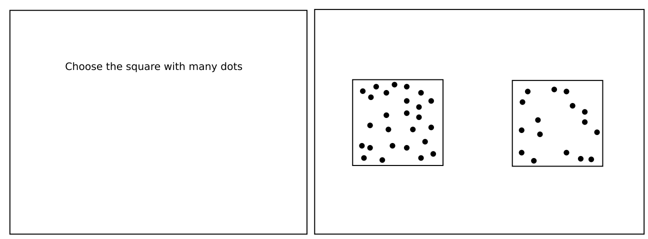

Figure 3: Stimulus, experiment 2. Left panel shows instruction screen, right panel shows stimulus screen

Results We decided not to impose an upper limit on response times in this experiment because the numbers of dots involved were much larger than in experiment one, and we wanted to allow for the possibility that participants might take a long time to count them. No response was less than 250 ms. Response times for trials with erroneous responses were discarded, leading to the loss of 354 trials from 5120, representing 6.9% of the trials. The results indicated that vague quantifiers reliably attracted faster accurate response times and fewer errors than did the crisp quantifiers, for all the combinations of numbers that we used. Figure (4) shows mean response times and error rates, separated out across the levels of numerical distance between the squares in the stimuli. The vagueness advantage in this experiment was large, and very reliable. For response times, the main effect of vagueness was a 316 ms advantage in the vague conditions (mean crisp: 1246 ms; mean vague: 930 ms, $F_{(1,19)} = 12.1, p < .01$ in a repeated measures ANOVA). Mean error rates more than halved in the vague conditions (mean crisp error rate = 9.3%; mean vague error rate = 4.5%; Wald z= 2.3, $p < .05$ in a generalized linear mixed model as recommended by Jaeger, 2008). Vagueness brought diminishing returns as the gap size grew larger, until at the largest gap size, it conferred no advantage relative to crisp equivalents (see Fig. 4).

Discussion We were able to reject our null hypothesis that there would be no difference between responses to the vague and crisp instructions: we found evidence for faster responses, and fewer errors in the vague conditions. We in-

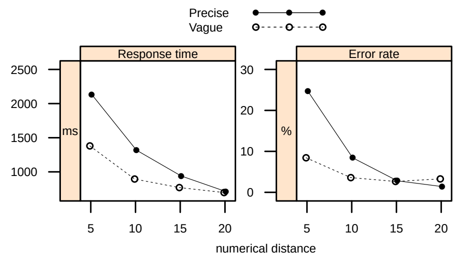

Figure 4: Results, experiment 2

terpret this difference as an advantage for vagueness.

Experiment Three

Introduction Experiments one and two did not have borderline cases in the stimuli. The vague conditions using *few* and *many* had the *potential* for vagueness: but it could be argued that this potential was not realised because there were only two squares and therefore each vague instruction uniquely identified a particular square, and ruled out the other. Therefore it could be argued that the processing advantages that we observed in experiments one and two cannot have been due to vagueness because there were no borderline cases.

Experiments one and two did not isolate vagueness from a possible confound. The vague instructions used the words *few* and *many*. While these are vague, they also refer without using a number. In contrast, the precise instructions used e.g., *five*, *20*. While these are precise, they also refer using a number. Therefore our vagueness advantage could be an advantage for not using a number in the referring expression instead.

Experiment three set out to remedy these shortcomings: we made sure that there were borderline cases of the vague instructions in the stimuli; and we separated the effects of vagueness from the effects of the absence of a number in the instructions.

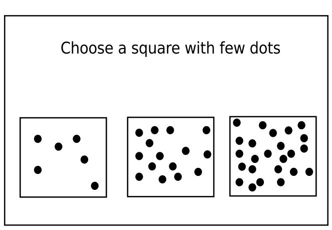

Figure 5: Example stimulus for experiment three

We made sure that there were borderline cases in the stimuli by using three squares in the stimulus (see fig 5); as well as by using indefinite articles in the vague conditions (see Table 1). Using the stimulus $\{16, 25, 34\}$ and the instruction *Choose a square with about 20 dots* as an example, the instruction's scope of application includes the square with 16 dots – this is the closest match for the instruction – but we argue that it also includes the square with 25 dots, and that because it is a poorer match, the square with 25 dots constitutes a borderline case (see fig 6).

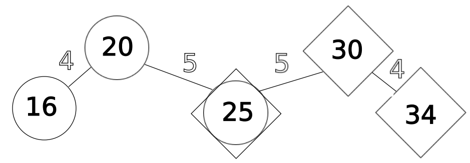

Figure 6: How borderline cases are construed in experiment three. Filled numbers represent numbers of dots; non-filled numbers represent numerical distances

We separated vagueness from non-numerical instructions, by manipulating vagueness and numerical / non-numerical format in a 2 x 2 factorial design, yielding four subconditions as follow, and as tabled in Table (1): number-crisp (*Choose the square with 16 dots*); number-vague (*Choose a square with about 20 dots*); word-crisp (*Choose the square with the fewest dots*); word-vague (*Choose a square with few dots*). The crisp subconditions used definite articles; the vague subconditions used indefinite articles. The crisp subconditions used either a specific number or a superlative quantifier to indicate strictly one object. The vague subconditions used the expressions *about n* and *few* to plausibly indicate more than one object, but with strictly one object as the best match, leaving the other plausible referent as the borderline case that is at the heart of the interpretation of vagueness that we pursue here.

Table 1: Examples of instructions used in Experiment 3. "Choose … square with … "

	Crisp	*Vague*
Number	… the … {16, 34} dots	… a … about {20, 30} dots
Word	… the … the {fewest, most} dots	… a … {few, many} dots

This design allows us to pit our hypothesis (i.e., vagueness reduces processing costs) against the competing hypothesis (i.e., numbers in referring expressions incur processing penalties). The cost reduction hypothesis predicts that the processing advantage that we observed in experiment two will persist both at numerical and at non-numerical levels of the format manipulation (i.e., a main effect of vagueness). In contrast the number penalty hypothesis predicts that processing advantages will manifest in the no-number conditions regardless whether they are crisp or vague (i.e., a main effect of format).

Method We used the following triples representing number of dots in each of 3 squares: {6,15,24}, {16,25,34}, {26,35,44}, {36,45,54}, and the same triples in different left-to-right order. {24,15,6}, {34,25,16}, {44,35,26}, {54,45,36}. The reason for presenting each triple in both left to right orders was to counter-balance facilitation for responding to a small number when it is presented on the left (Dehaene, Bossini, & Giraux, 1993).

Results Responses were preprocessed as follows. No response was treated as erroneous because errors were essentially undefined for the vague conditions. 2 trials were excluded because they had response time (RT) less than 250ms and must therefore have been planned before the stimulus appeared, and 1 trial was excluded because it had RT greater than $50,000$ ms. This excluded less than 0.1% of data points.

Response times were submitted to 2-way within-subjects ANOVA with repeated measures on both vagueness and instruction format. We found a reliable main effect of instruction format (word vs number format) on response times: responses were 1862 ms faster in the word conditions compared with the number conditions (3773 ms vs 1911 ms; $F_{(1,29)} = 13.4, p < .01$). We found no main effect of vagueness on response times, indicating that mean response time in the vague conditions did not differ reliably from mean response time in the crisp conditions. We found no reliable vagueness x format interaction for response times (See Fig 7). This indicates that there was no reliable additional benefit for vagueness either in the no-number conditions or in the number conditions.

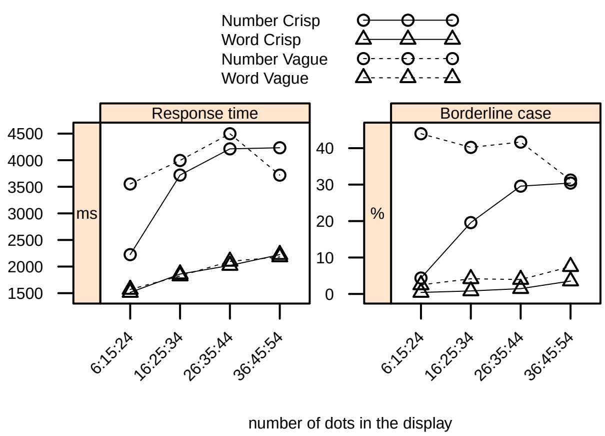

Figure 7: Results for experiment three. Response time in left panel, proportion of trials on which a borderline case was the response in right panel.

The choice of the borderline response (a binary value, chosen vs not-chosen) was analysed with a generalized linear mixed model using a binomial link function (Jaeger, 2008). This measure serves as an analogue of the erroneous response measure in the previous experiments. To choose the borderline response in the crisp conditions is erroneous: in the vague conditions it indicates that the participant did not choose the closest match for the instruction. The vagueness and format manipulations both exerted reliable influences in their own right for choosing the borderline response. Participants were reliably more likely to choose the borderline square for vague instructions - when the borderline response was not strictly an error - than for crisp instructions (21.9% vs $11.3\%, Wald\ z = 12.5, p < .001$). Participants were more likely to choose the borderline response for numerical than for word format (30.1% vs $3.0\%, Wald\ z = -14.8, p < .001$). Numerical format instructions had a scope of application symmetric about the number and included the borderline case, whereas the scope of application of the word format instructions was skewed higher or lower than the declared quantity[1]. The nature of the unreliable vagueness x format interaction was that symmetric number-vague conditions (e.g., *about 20*) attracted the most borderline responses, with symmetric number-precise (e.g., *34*) conditions attracting increasingly more such responses as the size of the number grew; and both the asymmetric word conditions (e.g., *few, fewest*) attracting hardly any such responses.

Discussion Experiment three successfully teased apart vagueness from word vs number format. It also successfully ensured that the vague expressions were still vague when taken in context with the stimulus. It provided an appropriate arena for rejecting the null hypothesis that there is no difference between crisp and vague referring expressions, in terms on response speed and accuracy. It also provided an opportunity for any advantage for verbal format as opposed to numerical format to manifest itself. The findings from experiment three indicate that while it is still true that word-vague instructions are advantageous over number-precise instructions (as found previously in experiment two), when vagueness and format are manipulated factorially (as in experiment three), this advantage is revealed to be an advantage for the word format rather than for vagueness.

General Discussion

Vagueness in a referring expression is a combination of the referring expression's potential for borderline cases; and the specific situation in which the referring expression is used. We found that expressions with the potential for vagueness attracted faster response times than expressions without; but only when the referent set did not allow the possibility of borderline cases. When the referent set did allow borderline cases, precise expressions that used no numbers performed just as well as vague expression that used no numbers; and precise expressions that did use numbers performed just as well as vague expressions that used numbers. Vagueness does exert advantageous effects, but it uses non-numerical format to do the heavy lifting.

Rayna & Brainerd (1989) proposed that reasoners encode representations at different levels of precision, and that reasoning operated at the lowest level that would allow one to accompish the task. The different levels can be described in terms of how many distinctions need to be made between objects in order to accomplish the task. A specific number re-

[1] *few* in the context of $\{16,25,34\}$ has an asymmetric scope of application equal to or lower than 16, but no higher than 16; *about 20* has a scope of application symmetric about 20 that includes both 16 and 25

quires one to draw many distinctions (9 must be distinguished from all cases of not-9). A term like *large* requires fewer distinctions - the objects must be divided into two sets, large and small. A term like *some lives* allows a boolean classification: either # lives = 0; or # lives > 0. Since non-numerical terms usually require fewer distinctions to be made, we suggest that one way in which non-numerical quantifiers might exert a beneficial influence is that they flag up for comprehenders that a lower, less cognitively demanding, level of reasoning might suffice for the task at hand, which would spare them the overhead of marshalling more expensive resources.

Dual process models of decision making (e.g., Sloman, 2007) propose two systems of processing: the quick and affective System 1 and the deliberative and rule-based System 2. This two-systems account can explain our finding of diminishing returns for vagueness in experiment 2. When the gap is small, participants use System 2 (slow, deliberative) for the precise instructions. When the gap is large, they take a shortcut for the precise conditions and merely establish which square is more (or less) numerous, using System 1 (quick, heuristic). In the vague conditions, they use the heuristic system whether the gap is small or large. This account explains the diminishing returns for vagueness that we observed in terms of a change of strategy in the baseline precise conditions as the gap size grew larger.

Conclusions

Although our experiments were limited in focussing on vagueness in descriptive noun phrases only, and although they did show up advantages for certain vague expressions, they do more to cast doubt on the cost reduction hypothesis than to confirm it. NLG systems can condense our findings into a rule that says that for numbers less than 5, a precise term will be easier for the comprehender than a vague term, and for numbers greater than 5, terms (whether vague or crisp) that do not mention numbers will be easier than terms (whether vague or crisp) that do, when the numerical distance from distractors is small, with diminishing returns for vagueness as the distance grows larger.

Acknowledgments

This work was funded in part by an EPSRC Platform Grant awarded to the NLG group at Aberdeen.

References

Bisantz, A., Marsiglio, S., & Munch, J. (2005). Displaying uncertainty: Investigating the effects of display format and specificity. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *47*(4), 777.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*, 433–436.

Crawford, V. P., & Sobel, J. (1982). Strategic information transmission. *Econometrica*, *50*(6), 1431–1451.

Dehaene, S., Bossini, S., & Giraux, P. (1993). The mental representation of parity and number magnitude. *Journal of Experimental Psychology: General*, *122*(3), 371.

Eaton, J. W. (2002). GNU *Octave Manual*. Network Theory Limited.

Goldberg, E., Driedger, N., & Kittredge, R. (1994). Using natural-language processing to produce weather forecasts. *IEEE Expert*, *9*(2), 45–53.

Hripcsak, G., Elhadad, N., Chen, Y., Zhou, L., & Morrison, F. P. (2009). Using Empiric Semantic Correlation to Interpret Temporal Assertions in Clinical Texts. *Journal of the American Medical Informatics Association*, *16*(2), 220-227.

Jaeger, T. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, *59*(4), 434–446.

Keefe, R., & Smith, P. (Eds.). (1996). *Vagueness: a Reader. A Bradford Book.* The MIT Press.

Lipman, B. L. (2009). *Why is Language Vague?* (retrieved 12 April 2011 from `http://people.bu.edu/blipman/Papers/vague5.pdf`)

Mishra, H., Mishra, A., & Shiv, B. (2011). In Praise of Vagueness: Malleability of Vague Information as a Performance-Booster. *Psychological Science*.

Peters, E., Dieckmann, N., Västfjäll, D., Mertz, C., Slovic, P., & Hibbard, J. (2009). Bringing meaning to numbers: The impact of evaluative categories on decisions. *Journal of Experimental Psychology: Applied*, *15*(3), 213.

Portet, F., Reiter, E., Gatt, A., Hunter, J., Sripada, S., Freer, Y., et al. (2009). Automatic generation of textual summaries from neonatal intensive care data. *Artificial Intelligence*, *173*(7-8), 789–816.

Rayna, V. F., & Brainerd, C. J. (1989). Fuzzy-trace theory of framing effects in choice. In *Proceedings of the 30th Annual Meeting of the Psychonomic Society.* Atlanta, GA.

Sloman, S. (2007). Two systems of reasoning. In T. Gilovich, D. Griffin, & D.Kahneman (Eds.), *Heuristics and Biases: The Psychology of Intuitive Judgment.* CUP.

Trick, L., & Pylyshyn, Z. (1994). Why are small and large numbers enumerated differently? A limited-capacity preattentive stage in vision. *Psychological Review*, *101*, 80–102.

Turner, R., Sripada, S., Reiter, E., & Davy, I. (2006). Generating spatio-temporal descriptions in pollen forecasts. In *EACL '06: Proceedings* (pp. 163–166). Stroudsburg, PA: Association for Computational Linguistics.

van Deemter, K. (2009). Utility and Language Generation: The Case of Vagueness. *Journal of Philosophical Logic*, *38*(6), 607–632.

van Rooij, R. (2003). Being polite is a handicap: Towards a game theoretic analysis of polite linguistic behavior. In M. Tenneholz (Ed.), *TARK 9: Theoretical Aspects of Rationality and Knowledge.* Bloomington: Bloomington.

Veltman, F. (2002). *Het verschil tussen 'vaag' en 'niet precies'. (the difference between 'vague' and 'not precise'.).* Inaugural lecture, Vossiuspers, University of Amsterdam.