

Minding Our Manners: Investigating The Nature Of Gricean Manner Implicature Using Psychological Priming Techniques

Joe Cowan

Department of Theoretical and Applied Linguistics
Darwin College, University of Cambridge

This dissertation is submitted for the degree of

Doctor of Philosophy

December 2024

This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the preface and specified in the text. It is not substantially the same as any work that has already been submitted, or is being concurrently submitted, for any degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the preface and specified in the text. It does not exceed the prescribed word limit for the relevant Degree Committee.

Minding Our Manners: Investigating the nature of Gricean manner implicature using psychological priming techniques.

Joe Cowan

Summary

That we often mean more than we *say* is a hallmark of human communication. As per the Gricean-inspired perspective of communication, the gap between *what is said* and *what is implicated* is bridged using a set of conversational maxims that can be exploited to trigger *conversational implicatures*. This thesis investigates the nature of Gricean *Manner Implicature* (MI) and its relationship with other types of conversational implicature, specifically *ad-hoc and adjectival scalar Quantity Implicature* (QI) and *I-Implicature*.

While QI, which arises from violations of the Maxim of Quantity, has enjoyed broad theoretical and experimental investigation, MI—triggered by violations of the Maxim of Manner (e.g., “Be perspicuous”)—remains comparatively underexplored. Moreover, there is no consensus between popular models of conversational implicature derivation concerning the extent to which shared or partially shared mechanisms derive MI, QI and I-Implicature. Similarly, there is a lack of clarity concerning the nature of different implicature subtypes, their qualities, and the extent to which they constitute unique types of conversational implicature.

This thesis contributes to research in experimental and theoretical pragmatics by extending Bott and Chemla’s (2016) priming paradigm to MI, which contributes to the scant body of research concerning Manner-based enrichments. The thesis presents a series of experiments with adult native English speakers that used a novel sentence-to-picture matching task to prime the derivation of MI using MI primes, two types of QI primes and I-Implicature primes, to explore the extent to which priming occurs within and between different types of conversational implicature.

In Experiment 1, a novel methodology for eliciting MI and I-Implicature is developed, establishing baseline rates for each and providing evidence of the psychological validity of marked-based contrasts. Experiment 2 explores whether other MI can prime MI and whether the effect is influenced by implicit or explicit instruction. Experiments 3–6 extend the priming paradigm to test cross-phenomenon effects, investigating whether I-Implicature, ad

hoc QI, adjectival scalar QI, and a neutral condition can prime MI. The findings suggest that both MI and QI and I-Implicature can prime MI but suggest that the nature of the within-category and between-category priming effects are fundamentally different. These findings are discussed with respect to theoretical models of implicature derivation.

Acknowledgements

First and foremost, I'd like to thank Napoleon Katsos, a supervisor of this thesis, for his enduring support. As an undergraduate, I asked Richard Breheny if I should apply to the University of Cambridge to continue my studies. At the time, I was working alongside Richard on a re-run of a study by Lewis Bott and Emmanuel Chemla (referenced countless times in this thesis). Richard said that if I got in, I'd be working with 'Napoleon Katsos', and I remember being struck by the uniqueness of his name. Then, I did not imagine that I'd now be finishing up a PhD under Napoleon's guidance, having been supervised by him for an MPhil, too. Neither did I imagine that I'd still be using some of the experimental stimuli I'd made five years earlier, nor that I'd still be furiously typing the reference 'Bott and Chemla (2016)'.

Napoleon has supervised me with patience, empathy, and friendliness and has helped me make the most of my years in Cambridge. I am fortunate to be co-supervised by Jenny Gibson, too, who has offered an invaluable perspective and influenced me to explore academic arenas beyond those of Experimental Pragmatics.

Alongside both, the wider Experimental Pragmatics community has given me insightful, thought-provoking feedback each step of the way and has provided me with immense help in refining the ideas and methodologies of this thesis. In addition, I owe thanks to my peers in TAL, who have helped me in a number of ways, from assisting with my relentless statistics questions to providing companionship and direction. Moreover, I appreciate the thousands of participants from across the world who have participated at various stages in this research. Although I do not know who they are, I could only have completed this thesis with their input.

While I am lucky enough to have had a great deal of academic support, Darwin College has provided me with a much-needed social outlet that has undoubtedly played a role in the fruition of this thesis. I want to thank my friends at Darwin – and beyond – for making my time spent completing this PhD unforgettable, irreplaceable and doubtlessly much more bearable than it would have been had I gone it alone. Outside of Cambridge, I have the continued support of family and friends, far and wide.

Finally, I'd like to acknowledge everything the PhD process has taught me. Researching and writing this thesis has changed my approach to both learning and working, and I am eternally

grateful for the opportunity I have been given. Although the hours spent engaged, challenged and curious are equally matched with those spent wanting to pull at tufts of my hair, the ups and downs of the journey have shaped my perspective in ways that I could never have anticipated.

Parts of this thesis have been written in Venice and other parts in Istanbul, while I typed up the bulk in Darwin's study centre, overlooking the Mill. However, by sheer coincidence, I polished things off where everything began, in the library of UCL's Chandler House.

Contents

1.	The Elusive Manner Implicature	1-14
2.	Background.....	2-17
2.1.	Grice’s Account of Implicature	2-17
2.2.	Neo-Gricean Accounts of Implicature	2-24
2.2.1.	Horn’s Account of Conversational Implicature.	2-24
2.2.2.	Levinson’s Account of Conversational Implicature.....	2-28
2.2.3.	The (neo-)Gricean Accounts: comparison and conflict	2-32
2.3.	The Gricean-Inspired Framework.....	2-32
2.3.1.	QI in the Gricean-inspired perspective	2-34
2.3.2.	Post-Gricean Accounts of QI	2-41
2.3.3.	MI in the Gricean-inspired perspective.....	2-43
2.3.4.	Post-Gricean Accounts of MI.....	2-57
2.3.5.	I-Implicature in the Gricean-Inspired Perspective.....	2-58
2.3.6.	The Similarities and Differences Between QI and MI in the Gricean-Inspired Perspective.	2-67
2.3.1.	The Similarities and Differences Between QI, MI and I-Implicature in the Gricean-Inspired Perspective.	2-70
2.4.	Conclusion	2-72
2.5.	Probabilistic Accounts of Conversational Implicature	2-72
2.5.1.	The Question-Under-Discussion (QUD)	2-74
2.5.2.	Conclusion	2-76
3.	Experimental Background	3-77
3.1.	Experimental Investigation of QI.....	3-77
3.2.	Experimental Investigation of MI.....	3-78
3.3.	Conclusion	3-83
3.4.	The Current Study.....	3-84
3.4.1.	Priming.....	3-85
3.4.2.	Conclusion	3-89
3.4.3.	Methodological Background.....	3-90
3.4.4.	Conclusion	3-97

4.	Priming MI with MI.....	4-98
4.1.	Overview of Experiments	4-98
4.2.	Pre-Experimental Item Screening	4-99
4.2.1.	Images	4-100
4.2.2.	Item Screening Task 1	4-102
4.2.3.	Item Screening Task 2	4-105
4.3.	Experiment 1 - MI and I-Implicature Baselines	4-107
4.3.1.	Experiment 1a – MI Baseline	4-108
4.3.2.	Experiment 1b – I-Implicature Baseline	4-118
4.4.	Experiment 1: Combined Analysis	4-124
4.4.1.	Data analysis	4-124
4.4.2.	Main Effect of Condition	4-125
4.4.3.	Item Analysis	4-125
4.5.	Experiment 1: Combined Discussion.....	4-129
4.5.1.	Manner Implicature.....	4-129
4.5.2.	I-Implicature	4-133
4.5.3.	Overall Conclusion	4-135
4.6.	Experiment 2a – Manner-to-Manner Priming.....	4-135
4.6.1.	Participants.....	4-136
4.6.2.	Design	4-136
4.6.3.	Materials	4-137
4.6.4.	Procedure	4-139
4.6.5.	Research Hypotheses & Predictions	4-139
4.6.6.	Coding.....	4-140
4.6.7.	Results & Analysis	4-140
4.6.8.	Discussion	4-144
4.7.	Experiment 2b – Manner-to-Manner priming (with training block)	4-149
4.7.1.	Motivations	4-149
4.7.2.	Participants.....	4-150
4.7.3.	Design	4-150
4.7.4.	Materials	4-150
4.7.5.	Results & Analysis.....	4-154
4.7.6.	Discussion.....	4-156

4.8.	Experiment 2c – Manner-to-Manner priming (Training Block & Guided Primes).	4-158
4.8.1.	Motivations	4-158
4.8.2.	Participants.....	4-159
4.8.3.	Design	4-159
4.8.4.	Materials	4-159
4.8.5.	Procedure	4-161
4.8.6.	Predictions.....	4-162
4.8.7.	Coding.....	4-162
4.8.8.	Results & Analysis.....	4-162
4.8.9.	Discussion	4-165
4.8.10.	Experiment 2c - Post-Hoc Item Analysis.....	4-171
4.8.11.	Effect of item typicality	4-173
4.9.	Experiment 2 – Overall Conclusions	4-176
5.	Priming MI with I-Implicature.....	5-177
5.1.	Experiment 3 – I-Implicature-to-MI Priming	5-177
5.1.1.	Motivations	5-177
5.1.2.	Participants.....	5-180
5.1.3.	Design	5-180
5.1.4.	Materials	5-181
5.1.5.	Procedure	5-182
5.1.6.	Predictions.....	5-183
5.1.7.	Coding.....	5-183
5.1.8.	Results & Analysis.....	5-184
5.1.9.	Discussion	5-186
5.1.10.	Conclusion	5-195
6.	Priming MI with QI	6-196
6.1.	Motivations	6-197
6.1.1.	The nature of QI.....	6-197
6.1.2.	Investigating a shared QI & MI derivational mechanism	6-202
6.1.3.	Predictions of the competing accounts	6-203
6.2.	Experiment 4 – Ad-hoc QI-to-MI Priming	6-204
6.2.1.	Participants.....	6-204

6.2.2.	Design	6-204
6.2.3.	Materials	6-205
6.2.4.	Procedure	6-207
6.2.5.	Predictions.....	6-208
6.2.6.	Coding.....	6-208
6.2.7.	Results & Analysis.....	6-208
6.2.8.	Discussion.....	6-211
6.3.	Experiment 5: Adjectival Scalar QI-to-MI Priming.....	6-212
6.3.1.	Participants.....	6-212
6.3.2.	Design	6-213
6.3.3.	Materials	6-213
6.3.4.	Procedure	6-216
6.3.5.	Predictions.....	6-217
6.3.6.	Coding.....	6-217
6.3.7.	Results & Analysis.....	6-218
6.3.8.	Discussion Experiments 4 & 5.....	6-221
6.4.	Experiment 6: Scalar Ambiguity.....	6-228
6.4.1.	Motivations	6-228
6.4.2.	Participants.....	6-229
6.4.3.	Design	6-229
6.4.4.	Materials	6-230
6.4.5.	Procedure	6-232
6.4.6.	Predictions.....	6-233
6.4.7.	Coding.....	6-233
6.4.8.	Results & Analysis.....	6-233
6.4.9.	Discussion.....	6-236
7.	Conclusions.....	7-237
7.1.	Summary of key findings.....	7-237
7.1.1.	Key conclusions & Unanswered Questions.....	7-239
7.2.	Implications for the nature of the priming effect(s).....	7-240
7.2.1.	Locus of the priming effect.....	7-242
7.3.	Implications of the structure of the ‘mechanism’	7-244
7.3.1.	Conclusions about the nature of mechanism	7-246

7.4.	Ramifications theories of conversational implicature.....	7-247
7.4.1.	The Gricean-inspired perspective	7-247
7.4.2.	Other accounts of implicature derivation.....	7-252
7.4.3.	Adjusting the Gricean-inspired perspective?	7-253
7.5.	Research Outlook.....	7-256
7.5.1.	Limitations	7-256
7.5.2.	Contributions.....	7-258
8.	References.....	8-261
9.	Appendices.....	9-279
9.1.	The item pairs used in the MI and I-Implicature critical and prime trials	9-280
9.2.	Trial Blocks (see 14 for context).....	9-285
9.3.	MI Critical And Prime Trial Captions.....	9-286
9.4.	I-Implicature Critical And Prime Trial Captions	9-287
9.5.	Experiment 4: Ad-hoc Captions.....	9-288
9.6.	Experiment 4: Ad-Hoc Items	9-289
9.7.	Experiment 5: Adjectival Scalar Captions	9-291
9.8.	Experiment 5: Adjectival Scalar Items	9-292
9.9.	Experiment 6: Scope Ambiguity Captions.....	9-295
9.10.	Experiment 6: Scope Ambiguity Items	9-296
9.11.	Filler Trials.....	9-300
9.12.	Consent Form (All Experiments).....	9-301
9.13.	Demographic Questionnaire	9-301

Table of Figures

Figure 2-1: A comparison of the different principles used by the Gricean-inspired perspective....	2-33
Figure 2-2: A summary of the equivalent concepts in the Gricean-inspired perspective.....	2-34
Figure 2-3: MI and QI implicature derivation in the Gricean-inspired perspective.	2-68
Figure 2-4: Table of qualities shared or not shared between QI and MI as per the Gricean-inspired perspective.	2-70
Figure 2-5 Table of qualities shared or not shared between QI,MI and I-Implicature as per the Gricean-inspired perspective.....	2-72
Figure 3-1: An example of the experimental stimuli used by Wilson & Katsos (2016) across conditions with both possible answers.....	3-81
Figure 3-2: The visual stimuli used in Wilson (2017) for the item associated with the caption: ‘The girl cut the apple’	3-83
Figure 3-3: The ‘strong’ and ‘weak’ trials and the ‘Better Picture?’ target trial used in Bott & Chemla (2016)	3-91
Figure 3-4: The priming-via-the-alternative prime trial in the ‘some QI’ condition from Rees & Bott (2018).....	3-94
Figure 3-5: An illustration of the salience and combination model taken from Rees & Bott (2018)	3-95
Figure 4-1: MI critical trial for the item ‘t-shirt’	4-100
Figure 4-2: marked and unmarked image pairs for different MI critical trials.	4-102
Figure 4-3: trial screen for the item ‘hat’ in Item Screening Task 1	4-104
Figure 4-4: trial screen for the item ‘bathrobe’ in Item Screening Task 2	4-106
Figure 4-5 MI critical trial for the item ‘t-shirt’.....	4-110
Figure 4-6: ‘Yes’ filler trial for the item ‘fork’.	4-113
Figure 4-7: The mean rate of MI derivation in Experiment 1a by item ordered smallest to largest.	4-118
Figure 4-8: The I-Implicature critical trial for the item ‘vest’	4-120
Figure 4-9: Filler trial for the item ‘fork’	4-121
Figure 4-10: The by-item rates of unmarked image selection in Experiment 1b.....	4-123
Figure 4-11: Combined rates of respective implicature derivation by item for both Experiment 1a and 1b.....	4-126
Figure 4-12: the Marked-Unmarked-False filler trial for the item ‘lamp’	4-134
Figure 4-13: the trial progression of Experiment 2a	4-138
Figure 4-14: The mean rates of MI derivation by item in Experiment 1a and 2a ordered by the baseline	4-143
Figure 4-15: bar chart depicting the rate of MI derivation by ‘primed’ status in Experiment 2b with error bars representing the standard error of the by-item mean rate of MI derivation.....	4-145
Figure 4-16: MI critical trial for the item ‘tuxedo’	4-151
Figure 4-17: the critical trial for the item ‘chair’ in the training block.....	4-152

Figure 4-18: bar chart depicting the rate of MI derivation by 'primed' status in Experiment 2b with error bars representing the standard error of the by-item mean rate of MI derivation.....	4-155
Figure 4-19: Critical trial for the item 'tuxedo' in the MI condition.	4-160
Figure 4-20: 'Guided prime' critical trial for the item 'polo shirt' in Experiment 2c.	4-160
Figure 4-21: bar chart depicting the rate of MI derivation by 'primed' status, with error bars representing the mean rates of by-item MI derivation in each condition for Experiment 2c. ..	4-164
Figure 4-22: The mean rate of MI derivation across Experiment 1a and Experiments 2a, 2b, 2c and 3, and 5 with error bars representing the standard error of the by-item mean rate of MI derivation.	4-166
Figure 4-23: Mean rate of MI derivation by the number of correct primes across Experiments 2a, 2b, 2c, with error bars representing the standard error of the by-item mean rate of MI derivation.	4-169
Figure 4-24: the results of the polynomial regression: <code>model_poly <- lm(Magnitude of Priming Effect in Exp 2c ~ poly(Typicality of Unmarked Image, 2) * poly(Typicality of Marked Image).</code>	4-172
Figure 4-25: contour plot based on the predictions of Figure 5-24's polynomial regression. Here, the magnitude of the priming effect is predicted across continuous combinations of typicality scores for the marked and unmarked images.	4-173
Figure 4-26: marked and unmarked images for the item 'skirt' and data from Item Screening Task 2 and Experiments 1a and 2c.	4-174
Figure 4-27: marked and unmarked images for the item 'cowboy boot' and data from Item Screening Task 2 and Experiments 1a and 2c.	4-175
Figure 5-1: MI critical trial for the item 'tuxedo'	5-181
Figure 5-2: I-Implicature prime trial for the item 'vest'	5-182
Figure 5-3: The mean rate of MI derivation across Experiment 1a and Experiments 2a, 2b, 2c and 3, and 5 with error bars representing the standard error of the by-item mean rate of MI derivation.	5-184
Figure 6-1: MI critical trial for the item 'tuxedo'	6-205
Figure 6-2: prime trial for the item 'circle'	6-207
Figure 6-3: The mean rate of MI derivation across Experiment 1a and Experiments 2a, 3, and 4 with error bars representing the standard error of the by-item mean rate of MI derivation. . .	6-209
Figure 6-4: MI critical trial for the item 'tuxedo'	6-214
Figure 6-5: Sentence and images for the item 'ajar' take from Pankratz & van Tiel (2021)	6-215
Figure 6-6: list of Items, Captions and Horn Scales for the items used in Experiment 5	6-215
Figure 6-7: Prime trial for the item 'content' in Experiment 5.	6-216
Figure 6-8: The mean rate of MI derivation across Experiment 1a and Experiments 2a, 3, 4 and 5 with error bars representing the standard error of the by-item mean rate of MI derivation.	6-218
Figure 6-9: MI critical trial for the item 'tuxedo'	6-230
Figure 6-10: scope ambiguity prime trial for the item 'blue circle' in Experiment 6	6-232

1. The Elusive Manner Implicature

Any well-versed communicator will know that the messages we seek to convey are often far richer, far more complex, and far more nuanced than the string of words with which we choose to convey them; to paraphrase, linguistic meaning is more than the sum of its parts. The utterance ‘Are you heading off?’ can mean ‘Wait for me!’, ‘You’ve embarrassed yourself, please leave’, or ‘Where are you going?’ depending on when, where, and by whom it is spoken. The ability to communicate pragmatically, i.e., to mean things that we don’t explicitly say, is a hallmark of typical human-to-human interaction and pervades our social exchanges; comedians use irony to elicit laughs from their audiences, subtext powers the caustic slights thrown at swanky cocktail parties, and hidden meaning is perhaps the most powerful tool in the arsenal of the romancer.

While most palpable when used for purposes of rhetoric, humour, or euphemism, the presence of pragmatic inference in everyday conversation goes often undetected. Even simple exchanges, like the following, require a listener to hear beyond the literal words of the speaker and make inferences about what the speaker intends to communicate:

A: ‘Would you like coffee and cake?’

B: ‘I would like coffee please’

Example 1-1

The intuitive interpretation of B’s reply is that B would like coffee and *not* cake. Although intuitive, the arrival at this inference relies on a complex series of mutually-taken cognitive steps. Firstly, to interpret B’s reply as exhaustive and avoid a follow-up question (e.g., ‘And what about the cake?’) A must make the assumption that B is attempting to be as informative as possible when speaking, i.e., willing to supply all the information required for the purposes of harmonious exchange. On the flip side, B must also assume that A is assuming B’s own

informativeness, otherwise B must either risk their reply being interpreted as under-informative by A or instead say, 'I would like coffee but not cake, please'. Finally, A must assume that, in light of the assumption that B is informative and knows what they want, that if B had wanted cake they would have said so, and that therefore B must *not* want any cake, rather than being ambivalent about whether or not they receive any. Moreover, for the exchange to occur effectively, both parties must assume that the other is telling the truth, e.g., that B does indeed want coffee and that A is able to provide it.

These mutually assumed cognitive strategies provide a vital lubricative function in conversation, ensuring that we are neither burdened by the cumbersome process of articulating absolutely everything that we wish to communicate nor burdened with the doubt that what we hear may be untrue, inaccurate or incomplete – at least according to the dominant model of conversational inference proposed by Grice (1975). While quotidian pragmatic inference may fly under the radar, its necessity in ensuring cheap and easy interaction becomes abundantly clear when they it is scrutinised. In the above example, A's assumption of B's informativeness is the basis of the inference that B would just like coffee, which allows B to avoid having to oversupply information so as not to confuse A.

Informativity-based inferences are referred to as *quantity implicature* (Grice, 1975).

Another assumption we make of our fellow interlocutors is that they will express themselves as perspicuously as possible, avoiding unnecessary ambiguity, prolixity, or obscurity (Grice, 1975). This assumption concerns not what is said, but rather *how* it is said, i.e. the utterance's form. In general conversation, abidance to this principle goes relatively unnoticed – after all, speaking 'normally' is unlikely to attract particular attention. However, when apparent commitment to the mutually assumed rule of clarity of expression falters, this gives rise to conversational inference (Grice, 1975). Take the following exchange from a Wall Street Journal interview with Aretha Franklin, the legendary singer, in which the legacy of newer popstars is discussed:

Interviewer: When I say the name Adele, what comes to mind?

Aretha: Young singer, good singer.

Interviewer: Alicia Keys?

Aretha: Young performer, good writer, producer [of music].

(Wall Street Journal, 2014)

Example 1-2

The interview garnered attention for Aretha Franklin's slighting of Alicia Keys (Bond, 2018), the vehicle of which appears to be a form-based conversational inference. Here, there appears to be a contrast between Aretha's use of the word 'singer' in describing Adele and the prolixity of the description of Alicia Keys as a performer, writer and a singer. The expanded definition of 'singer', in the context of pop stardom, can be understood to encapsulate the duties of performing, songwriting, and producing and is, theoretically, applicable to Alicia Keys, too. However, to justify Aretha Franklin's apparent opting out of a commitment to perspicuousness, listeners make conversational inferences that imbue this opting out with a sense of pragmatic meaning. Interpreting the utterance pragmatically allows the listener to continue to assume that Aretha is committed to the principles of rational communication, which is necessary in order to prevent a communication breakdown between the interlocutors. Given that Aretha Franklin could have described Alicia Keys as a 'singer', but instead used a prolix description thereof, we can infer that Aretha Franklin seeks to communicate that Alicia Keys is not a singer *in the stereotypical sense*, for example, is perhaps not as talented a vocalist as a typical popstar (or, at least, Adele). As per Grice's account, this type of conversational inference is referred to as *manner implicature* and is the focal point of this thesis.

The thesis presents a series of experiments that seek to investigate the nature of manner implicature. As described, quantity implicature and manner implicature are interrelated phenomena, according to theoretical perspectives inspired by Grice (1975). Nevertheless, they have received unequal theoretical and experimental investigation and manner implicature remains relatively underexplored. Moreover, as demonstrated in Experiment 1a – MI Baseline, manner implicature's nature is far more elusive and heterogeneous than that of quantity implicature. In Experiment 1a of the thesis, I establish an experimental baseline for manner implicature to investigate how manner implicature can be elicited in experimental settings. Experiment 1a allows me to better understand the frequency with which manner implicature is derived generally and the degree to which this frequency varies between different types of ad-hoc manner implicature. Experiments 2 – 5 investigate whether manner implicature is subject to psychological priming effects. Firstly, Experiment 2a, 2b and 2c test

whether rates of manner implicature derivation can be increased by priming manner implicature and investigate the factors that modulate an increase in manner implicature derivation. Finally, Experiments 3 – 5 test whether manner implicature can be primed by quantity implicature and another type of implicature, *I-Implicature*, to investigate the extent to which the phenomena are interrelated.

To understand why manner implicature warrants the type of experimental investigation presented in this thesis, an extensive theoretical groundwork must be laid. In this introductory chapter, I will present the dominant theoretical accounts of conversational implicature: the Gricean account, Neo-Gricean accounts, the Relevance Theory account, a.o. I will then describe the theoretical assumptions of both Pragmatics and Experimental Pragmatics under which the undertakings of this thesis operate. Finally, I will direct attention to MI, specifically, and explore the ways in which MI's theoretical distance from other types of conversational implicature, namely, its elusivity and heterogeneity, justify the experimental investigation presented in this thesis. I will then present the subsequent theoretical questions the thesis seeks to address

2. Background

2.1. Grice's Account of Implicature

Considered the forefather of Pragmatics, Grice's seminal work *Logic and Conversation* (1975) seeks to address the described gap that exists between the words that we articulate and the messages that we wish to communicate. In response to an era of philosophical thought committed to the principles of Ideal Language Philosophy – the notion that we should seek to understand language unambiguously by appeal to the tools of formal logic – the ideas presented in *Logic and Conversation* seek to instead account for the existence of the types of contextually bound implied meaning that cannot be accounted for by propositional logic (Noveck, 2018). For Grice, the distinction, and sometimes conflict, between *what is said*, the literal meaning of an utterance, and *what is implicated*, what is understood by the utterance in a particular conversational context, is the launching point of his framework of conversational meaning.

Referred to by Grice as the *Co-operative Principle*, Grice's thesis is centered around the notion that when interlocutors engage in conversation, they tacitly commit to the mutually assumed principle of cooperative and rational communication. In practice, the commitment to abiding by the Co-operative Principle is reliant on the interlocutors' adherence to a set of follow-on subordinate constituents that give rise to fluent, efficient communication. In certain contexts, these maxims can be exploited in order to elicit and derive *conversational implicature*, a process in which an utterance is pragmatically enriched. Being a product of the mutually saddled Co-operative principle, conversational implicature is therefore bidirectional and is predicated on both the speaker's intention and the listener's willingness to derive the implicature.

As exemplified in the introduction to this chapter, the specific assumptions that interlocutors make about the meaning of one another's utterances are manifold and this heterogeneity is accounted for in Grice's framework by a diverse set of conversational maxims, the maxims of *quantity, quality, relevance, and manner*, which themselves comprise of a set of subordinate maxims too, as presented below:

i. The Co-operative Principle:

i. Be co-operative.

ii. Conversational Maxims:

Quality: *Be truthful*

- i. Don't say what is false.
- ii. Don't say what lacks evidence.

Quantity:

- i. Don't say less than is required.
- ii. Don't say more than is required.

Relation: *Be relevant*

Manner: *Be perspicuous*

- i. Avoid obscurity.
- ii. Avoid ambiguity.
- iii. Be brief.
- iv. Be orderly.

(Grice, 1975)

Example 2-1

While adherence to these maxims ensures efficient conversation, interlocutors often fall short of such conversational ideals. Interlocutors may opt out of a particular maxim by refusing to be cooperative, for instance, by responding ‘No comment’ to a question. Or an interlocutor may quietly violate a maxim in an effort to obscure information they may not want to share, e.g., a company may seek to dress up a job advertisement by substituting the job title ‘Cleaner’ for that of ‘Waste Management and Disposal Technician’ and, in doing so, violate the maxim of Manner. For Grice, the ostentatious disregarding of conversational maxims that gives rise to conversational implicature is called either *flouting* or *violating* a maxim. While Grice’s delineation of the two is somewhat vague, the flouting of a maxim refers to the blatant, ostensive disregarding of a maxim, such that it is abundantly clear to the listener that the speaker is doing so. Often, the type of conversational implicatures generated by the flouting of maxims aligns with the layman’s understanding of sarcasm or irony. Take the following exchanges:

A: Do you like chocolate?

B: No, I think it is absolutely disgusting.

(Spoken while B eats a chocolate bar.)

Example 2-2

A: What do you think about Cuba?

B: I love mojitos!

(Spoken while watching a documentary about the Cuban missile crisis)

Example 2-3

In Example 2-2, Speaker B flouts the maxim of Quality due to the obvious contradiction between their professed disgust at chocolate and the conversational context. In Example 2-3, we can consider Speaker B to have violated the maxim of Relation, assuming that A's question itself implies '+> what do you think about Cuba from a geopolitical standpoint?'¹ and that B understands this. Instead, B answers the question as though understanding it more generally and therefore provides an irrelevant answer to A's question for the purposes of ironic effect. Crucially, in both, the addition of the conversational implicature (e.g., +> I do like chocolate) to the utterance does not rectify the violation of the maxim, but rather produces a conflict between the utterance and the implication.

Comparatively, the violation of a maxim occurs when a maxim is simply contravened with no attempt on the part of the speaker to make this particularly obvious. As mentioned, one can violate a maxim for the purposes of deception. In these instances, the violation of the maxim is either covert and goes unnoticed or, when noticed, is interpreted as deliberately deceptive by the listener. However, in order to consider a speaker deceptive, a listener must suspend their belief in the mutually assumed Co-operative Principle. Given the fundamental role the Co-operative principle plays in Gricean communication, its wanton suspension is therefore undesirable and its maintenance requires a benefit of the doubt approach from listeners. Here, listeners derive conversational implicatures that, when combined with the articulated utterance, rectify this violation – such instances shall be the focus of this thesis.

Take the following examples and their respective implicatures:

A: Do you like the Gallagher brothers?

B: I like Liam

¹ The symbol '+>' is borrowed from Levinson (2000) and denotes that whatever follows is an implicature. For instance, '*some of the cars are green*' implies '+> not all of the cars are green'.

+> I do not like Noel.

Example 2-4

A: Are you going skiing this year?

B: My stipend doesn't stretch that far.

+> I cannot go skiing because I cannot afford it.

Example 2-5

In Example 2-4 we have an example of a quantity implicature, whereby B has supplied an informationally inadequate response to A, providing their opinion on just one, not both, of the Gallagher brothers. Conversely, in Example 2-5 we have an example of a *relational implicature*, whereby Speaker B's response, superficially, fails to address the question asked by Speaker A.

In accordance with Grice's framework of conversational implicature, the steps taken to achieve reach the implicatures in the examples using Example 2-4 as an illustration, are as follows:

- i. The listener expects the speaker to share their opinion on the Gallagher brothers.
- ii. The listener recognises that the speaker could have said more than they did, i.e., shared their opinion on both Liam and Noel.
- iii. Assuming the speaker is rational, the listener realises that the speaker has chosen not to impart the omitted information (rather than, say, has forgotten to, or has been unable to).
- iv. To reconcile conflict between an adherence to the Co-operative Principle and the violation of the maxim of Quantity, the listener assumes that the speaker has chosen not to explicate the omitted information for a particular reason.

- v. The listener, using contextual information, makes an assumption that renders the supposition tenable. For instance, the listener might assume that the speaker seeks to implicate ‘+> I like Noel, but I do not like Liam’.

Or this process can be more succinctly summarised as follows:

‘Speaker B said ϕ , and could as easily have said ψ , which is stronger. Because Speaker A is cooperative, she makes the strongest true statement possible, so ψ can’t be true.’

Whereby ψ refers to the informationally stronger alternative, in this case: ‘I like Liam and Noel’.

(Russell, 2006).

Example 2-6

As described, the violation of the Maxim of Manner gives rise to manner implicature, the derivation of which occurs in a comparative fashion to that of quantity implicature, but operates along lines of convention, rather than strength of information. Grice (1975) cites the presence of implied meaning in the latter of the following sentences as an illustration of *manner implicature*:

- i. ‘Miss X sang ‘Home sweet home.’
+> Miss X produced a series of sounds that corresponded closely with the score of ‘Home Sweet Home.’

Example 2-7

Here, the utterance gives rise to the following implicature ‘+> Miss X’s performance of ‘Home Sweet Home’ was bad.’ Revisiting the strikingly similar demonstration of manner implicature illustrated in Example 1-2, the derivation thereof can be described as such:

- i. Given Aretha’s prior use of the word ‘singer’, the interviewer expects Aretha Franklin to comment on Alicia Key’s abilities as a ‘singer’.

- ii. Any listener recognises that Aretha could have, or should have, used more perspicuous, concise language to describe Alicia Keys i.e., also referred to her as a ‘singer’.
- iii. Assuming Aretha is rational, any listener realises that Aretha has chosen not to use the more perspicuous, synonymous alternative ‘singer’.
- iv. To reconcile conflict between Aretha’s adherence to the Co-operative Principle and the violation of the maxim of Manner, the listener assumes that Aretha has chosen not to use the more perspicuous alternative for a reason.
- v. The listener assumes that if Aretha considered Alicia Keys a ‘singer’, she would have used the word ‘singer’ and therefore seeks to indicate a difference between Alicia Keys and singers to whom the label ‘singer’ is commonly applied. Clearly, contextual input is necessary in determining the nature of this difference and, in Example 1-2, the most salient interpretation is that Alicia Keys is not a talented vocalist.

Example 2-8

As illustrated by Example 2-4 and Example 2-7, there is a subtle difference in the way the Gricean programme accounts of the derivation of quantity implicature (henceforth QI) and manner implicature (henceforth MI). Crucially, MI’s derivation is not driven by the informational strength of an utterance, unlike QI, but rather the perspicuousness of its expression. In simpler words, MI concerns itself with not *what* is said, but rather *how* something is said.

It must be noted that the derivational process explicated in the examples is extrapolated from Grice’s (1975) seminal description of conversational implicature and the process by which a listener derives it. Grice’s original thesis is not explicit in the delineation of the particular ‘steps’ to this cognitive process and Example 2-6 and Example 2-8 represent an interpretation of the consensus of Grice’s thesis, inspired by the type of derivation processes laid bare by Horn (1984) and Levinson (2000).

2.2. Neo-Gricean Accounts of Implicature

Before discussing the nature of MI and its use in more depth, I will introduce the influential Neo-Gricean accounts of conversational implicature proposed by Levinson (1983, 2000) and Horn (1984, 1991, 2007) that provide reconceptualised frameworks of MI and QI, heavily inspired by Grice's seminal account.

2.2.1. Horn's Account of Conversational Implicature.

Horn's 'Division of Pragmatic Labor' (1984, 1989) posits the commitment of participants in a conversation to two counteracting principles: the Q-principle and R-principle, presenting a slimmed-down repackaging of Grice's more exhaustive set of conversational maxims. The principles are as follows (Horn, 1984, p. 13):

Q-principle = Make your contribution sufficient; Say as much as you can (given the R-principle)

R-Principle = Make your contribution necessary; Say no more than you must (given the Q-principle).

Example 2-9

Horn's Q-principle governs how informative a speaker should be, encapsulating the first of Grice's Quantity submaxim, 'don't say less than is required' and two of the subordinate maxims of Manner, 'avoid obscurity' and 'avoid ambiguity'. Implicatures arising from the Q-principle (or *Q-Inference*) are lower-bounding and are generally triggered by 'Horn scales' – lexical scales, e.g., < *some, most, all* >, in which any lower (leftwards) constituent is unilaterally entailed by those higher (in 'negative' Horn scales, this order is reversed), e.g., 'some' is entailed by 'most' and 'all', but not vice-versa (Horn, 1984, 2005, 2009). Q-inference serves to restrict access to meanings on the upper end of a scale, e.g., in using the word 'some' as opposed to 'all', a speaker is said to contravene the Q-principle, as the speaker could have used the informationally stronger scale member 'all' instead yet chose not to. As a result, the upper end of the scale ('all') is restricted and the implicature '+> some but not all' is derived. Put simply, as per the Q-principle, saying an informationally weak

utterance ‘x’ dictates that a speaker means ‘exactly x’ or ‘no more than x’ (Wilson & Katsos, 2023; Huang, 2009).

Conversely, Horn’s countervailing R-Principle is an upper-bounding principle that is most commonly exploited to restrict access to lower-bounded meanings. The R-Principle encapsulates Grice’s maxim of Relation, the second maxim of Quantity ‘Don’t say more than is required’ and two of the subordinate maxims of Manner: ‘be brief’ and ‘be orderly’ (Wilson & Katsos, 2023; Huang, 2009). In contrast to the generic Q-Inference that ‘a speaker means no more than they have said’, *R-Inferences* arise when a listener infers that by ‘x’ a speaker means ‘more than x’. There is a difference between the stratification of conversational implicature in Horn’s countervailing Q- and R-Principles and Grice’s original proposition. In the latter, QI and MI are stratified by whether the locus of an implicature’s trigger is content, i.e., QI’s reliance on the quantity of information, or form, i.e., MI’s reliance on a departure from what Grice calls ‘reasonable dispatch’.

In contrast, Horn’s proposal rehashes this division and QI and MI can both be triggered by both form and content, so differ in nature to Grice’s original conception. Horn’s rehashing of Grice’s framework is inspired by Zipf’s (1949) description of a communicative economy that takes a cost-benefit approach to the counteracting principles of least effort, benefitting the speaker, and anti-ambiguity, or specificity, benefitting the listener. Consequently, for Horn, over-informativity, prolixity and ambiguity are considered to violate the R-Principle in the same way, as they are non-beneficial to the speaker and give rise to unnecessary contributions, which are burdensome for the listener. In contrast, Q-based violations benefit the speaker by conserving effort, but give rise to contributions that are insufficient for the listener.

Naturally, the antinomic nature of Horn’s Q- and R- principles results in a conflict between the constraints of each principle; the resolution of this conflict is presented by Horn (1984) as the *division of pragmatic labour*, with the general pattern thereof summarised as follows:

‘The use of a marked (relatively complex and/or prolix) expression when a corresponding unmarked (simpler, less ‘effortful’) alternate expression is available tends to be interpreted as conveying a marked message (one which the unmarked alternative would not or could not have conveyed’ (Horn, 1984:22)

Example 2-10

This pattern of reasoning can be explicated more specifically, as demonstrated below, paraphrasing Horn by applying the reasoning to Example 1-2's MI:

- i. The speaker used the marked expression 'Young performer, good writer, producer*' or E*, containing extra information (i.e., prolixity) when a corresponding unmarked and essentially coextensive expression 'singer', or E is available.
- ii. Either (i) the 'extra' material was irrelevant and unnecessary, or (ii) was necessary and 'singer' could not have been appropriately used.
- iii. As (i) is in conflict with the R-Principle, this interpretation is to be rejected.
- iv. Therefore, if either (i) or (ii) must be the case, and (i) is to be rejected, a listener must assume (ii), that 'singer' could not have been appropriately used.
- v. The unmarked *alternative* E, i.e., 'singer', tends to be associated with an unmarked situation ('s'), which represents the stereotypical or salient features or elements of the extension of the set E/E* (i.e., the stereotypical meanings or referents encompassed by the terms singer and '[..] performer [...] writer [...] producer*', e.g., a [talented] vocalist. (R-based inference)
- vi. In contrast, the marked alternative, E*, is associated with the complement of 's' with respect to the original extension of E/E*, e.g., the less stereotypical, or salient elements, like performing, authoring songs, and producing music. (Q-based inference)²

Example 2-11

² Here, the contravening of Horn's R-Principle is resolved via the 'division of pragmatic labour', whereby the countervailing principles, Q- and R-, associate marked meanings with marked utterances (via the Q-Principle) and unmarked meanings to unmarked utterances (via the R-Principle). Horn's explanation for

Here, Step (vi) appears somewhat tautological as the resolution of the Q- and R-based conflict is superficially identical to the articulated utterance. However, it must be noted that ‘Young performer, good writer, producer’ is not explicitly restricted from communicating ‘+> a talented singer’ unless subject to the meaning-based constraints imposed by Horn’s countervailing principles. However, when these constraints are applied, the association of the marked alternative with the complement of ‘s’, i.e., everything but ‘s’, means that this interpretation is explicitly restricted. As such, ‘Young performer, good writer, producer’ is enriched by the implicature ‘+> not a talented vocalist’. As explained by Huang (2017), the R-Principle takes precedence until a contrastive linguistic form gives rise to a Q-based implicature that blocks the applicability of the R-implicature at hand.

In summarising the governance of Horn’s principles in relation to Grice’s QI and MI, Rett (2020) explains that QI arises when utterance complexity is fixed but utterance informativity is varied by the speaker, e.g., Example 2-4’s utterance ‘I like Liam’ does not vary in complexity to the alternatives ‘Yes’ or ‘I like Liam and Noel’ (disregarding the increased complexity conjunction may entail) but is less informative than either alternative. In contrast, MI arises when utterance complexity varies but utterance informativity is fixed., e.g., in Example 1-2, the prolixity of Aretha’s utterance is more prolix than expected, yet it remains informative.

On the flip side, Horn does not delineate the Q-Inference and R-Inference by informativity and complexity, as Grice does with QI and MI. Instead, Q-Inference is defined by utterance *sufficiency* (a similar concept to informativity), while R-Inference is defined by utterance *necessity*, a concept that encompasses form-based prolixity, but also over-informativity and ambiguity. As such, the Q- and R-Principles do not directly correlate to Grice’s Maxims of Quantity and Manner. That said, the type of utterances that give rise to QI (e.g, Example 2-4) and MI (e.g., Example 1-2) often do contravene the Q-Principle and R-Principle respectively.

why R-implicature (i.e., implicatures derived via the contravening of the R-Principle) are Q-restricted in their meaning is somewhat thin, but the rationale appears to be that the Q-principle (*‘say as much as you can’*) restricts utterances that contravene the R-Principle from associating with unmarked meanings, as the R-Principle *‘say no more than you must’* R-associates unmarked language with unmarked meanings, meaning that, if a speaker were to instead use Example 1-2’s *‘singer’* to mean ‘+> not a singer in the stereotypical way’ they would not be abiding by the principle *‘say as much as you can’*, as a listener would likely R-associate *‘singer’* with an unmarked meaning and the speaker’s implicature would not be derived.

2.2.2. Levinson's Account of Conversational Implicature.

Levinson (1983, 1990, 2000) presents a similar reiteration of Grice's original principles, reconstituting Grice's conversational maxims as three heuristics, each of which comprises a Speaker's Maxim (SM) and a Recipient's Corollary (RC), the former, dictating what a speaker says, the latter, dictating what a listener should infer.

- i. Q-heuristic =
 - a. SM: Do not say less than is required.
 - b. RC: What isn't said, isn't.

- ii. I-heuristic =
 - a. SM: Do not say more than is required.
 - b. RC: What is simply described is stereotypically exemplified.

- iii. M-heuristic =
 - a. SM: Do not use a marked expression without reason.
 - b. RC: What is said in an abnormal way, describes an abnormal situation.

(Levinson, 2000)

Example 2-12

Levinson's proposal serves as a response to the observation that Horn's Q-Principle functions on both information expressed on a sentential level (i.e., form) and information on a lexical level (i.e., informativity). Therefore, Levinson provides an explanation of implicature that accounts for implicatures that stem from semantic informativeness and those from form separately, given their disparate nature. While lexical informativity is modulated by the ambiguity or obscurity associated with particular lexical items (e.g., 'some' vs 'all'), form-based informativity is modulated on a sentential level, by the length or complexity of an utterance. As such, Levinson accounts for the former, by the Q[quantity]-heuristic and the latter by the I[nformativeness]- and M[anner]-heuristic (Huang, 2009).

Here, Levinson's Q-heuristic subsumes the first subordinate of Grice's Quantity maxim, '*Do not say less than is required*', and gives rise to Q-Implicatures, which are the product of a

contrast between an informationally stronger alternative and an informationally weaker articulated expression (Levinson, 2000). In other words, the use of a weaker expression (e.g., ‘*some*’) implicates the inapplicability of stronger expression (e.g., ‘*all*’, in the Horn scale <*some, all*>). As such, the Q-heuristic is responsible for lexically-triggered scalar implicatures, the types of casual implicatures detailed by Gazdar (1979) and so-called *ad-hoc* (or *particularised*) scalars (e.g., the implicature in Example 2-6, ‘I like Liam +> I like Liam but not Noel’) (Stiller et al., 2015). The derivation of Levinson’s Q-Implicature can be illustrated as follows, adapted from Levinson (2000):

A: Do you like Oasis’ music?

B: I like some of it.

- i. A recognises that B has used an expression (e.g., *some*) for which there is an informationally stronger alternative (e.g., *all*) that, with the articulated expression, forms part of a Horn scale (e.g., <*some, all*>) such that the alternative entails the articulated expression.
- ii. The speaker infers that the assertion of the informationally weaker expression implicates that the speaker is not able to assert the informationally stronger alternative, as if they could, they would have (as per the Q-Principle’s Speaker’s Maxim).
- iii. The speaker therefore understands that the informationally stronger alternative is not the case (e.g., not ‘all’) (as per the Q-Principle’s Recipient’s Corollary).

Example 2-13

Levinson’s I-Principle accounts for the second subordinate maxim of Grice’s Maxim of Quantity ‘Do not say more than is required’ and results in the unmarked-to-unmarked association between brief, simple expressions and expected or stereotypical referents or situations. For instance, the utterance ‘I walked the dog yesterday’ can be said to I-Implicate ‘+> I walked my dog in the usual, normal way’ or, say, ‘+> I walked my dog Lucy to the top of the golf course and back’ should this be the contextually warranted interpretation. I-

Implicature is therefore accountable for the implicatures derived from instances of *negative strengthening* (e.g., the inference ‘+> I actively dislike chocolate’ from the utterance ‘I don’t like chocolate’), *possessive interpretation*, e.g., the inference ‘+> the dog that is my current pet (i.e., Lucy)’ from the utterance ‘the dog’ as opposed to competing interpretations like ‘+> the rescue dog that I sponsor’), and *conjunction buttressing* (e.g., the inference ‘+> I arrived home and then ate dinner’ from the utterance ‘I arrived home and ate dinner’), a.o.

According to Levinson, I-Implicature is differentiated from Q-Implicature by its directionality. I-Implicature transforms a lack of specification, (e.g., the expression ‘the dog’ is underspecified compared to ‘my dog Lucy’) to an interpretation of sufficient specification e.g., the implicature ‘+> my dog Lucy’ from ‘the dog’, remedies the articulated expression’s lack of specification. In contrast, Q-Implicature transforms a lack of informational strength into a restriction on the speaker’s ability to provide it, e.g., Q-Implicature restricts the interpretation of the informationally richer ‘+> all the cars are green’ in the case of the utterance ‘some of the cars are green’.

More fundamentally, different types of I-Implicatures are considered to be unified only by the instruction that licences them (i.e., the I-Heuristic) and not by a commonality in the cognitive process by which they are derived, unlike Q-Implicatures which are unified by a common mental process by which they are derived (detailed above in Example 2-13). Furthermore, the derivation of Q-Implicature (and M-Implicature) involve the processing of an alternative manner of speaking and are therefore considered metalinguistic, as they require knowledge external to what has been said, i.e., a knowledge that the speaker could have said something more informative, or less prolix, but did not. In contrast, I-Implicature is not considered metalinguistic, both by virtue of its lack of a standard derivational process and its appeal to a stereotypical interpretation, rather than a contextually-specific one.

Finally, Levinson’s M-Heuristic relates to the form of an utterance, rather than the content, as Q-Implicature and I-Implicature and subsumes the entirety of Grice’s Maxim of Manner. M-Implicatures act on marked superficial forms, defined by Levinson as those that are more morphologically complex, prolix and periphrastic, and less lexicalized, less common and less neutral in register (Levinson, 2000), and lead to inferences that target marked meanings. Here, the meaning-based markedness is somewhat ill-defined, likely due to the heterogeneity of M-Implicature. For Levinson, a marked meaning is one with additional connotations to

those of the unmarked alternative or is the complement of the connotations of the unmarked alternative, i.e., all but the corresponding I-Inference.

The process of M-Implicature derivation can be illustrated as follows:

- i. A listener recognises that a speaker has used an expression (e.g., '*Young performer, good writer, producer*') for which there exists a less marked alternative (e.g., '*singer*') with the same denotation (i.e., one whose profession is to sing [...].)
- ii. The speaker infers that the assertion of the unmarked expression would I-Implicate a stereotypical subset of the denotation (e.g., '+> one whose profession is to sing and is therefore good at it) and that the speaker would have used it, had they no reason not to (as per the M-Principle's Speaker's Maxim)
- iii. The speaker therefore takes the marked utterance to implicate the complement of the stereotypical subset of the denotation (i.e., '+> ¬ [...] and is therefore good at it') (as per the M-Principle's Recipients Corollary).

Example 2-14

As pointed out by Levinson, M-Implicature is therefore reliant on a corresponding I-Implicature. To compute M-Implicature, a listener must first access both the alternative linguistic form, e.g., '*singer*', and its denotation. Here, contextual information is needed in order to generate both the alternative and its denotation. The generation of the appropriate stereotypical denotation is important as it is requisite in the generation of the M-Implicature, e.g., in Example 1-2, if '*singer*' is taken to implicate '+> a singer who is paid or professional', the unintended M-implicature '+> Alicia Keys is not paid for her singing' could be derived. However, Levinson does not explicitly account for an integration of contextual information, nor specify whether this occurs in the generation of the appropriate I-Implicature, the generation of the appropriate complement of the I-Implicature (i.e., the M-Implicature) or both.

2.2.3. The (neo-)Gricean Accounts: comparison and conflict

Naturally, Horn and Levinson's accounts of the derivation of conversational implicature enjoy points of considerable overlap, both with one another and with Grice's seminal account. From this perspective, Horn, and Levinson's accounts of implicature are often considered to comprise an integrated 'Neo-Gricean' perspective (e.g., Russell, 2006). The Neo-Gricean approach is contrasted with 'Gricean' accounts, e.g., Grice's seminal work and other Gricean-inspired perspectives, like those of Russell (2006) and Geurts (2009), by its greater departure from Grice's original account of conversational implicature. For the purposes of this thesis, I will use the umbrella term 'Gricean-inspired' to describe a higher-order perspective that encompasses both Neo-Gricean and Gricean accounts, but not perspectives that depart significantly from Grice's seminal account, for example, that of Relevance Theory (Carston, 2005; Sperber & Wilson, 1995) – although, such accounts do maintain some crucial elements of Grice's perspective, they are more accurately termed 'Post-Gricean'.

As this thesis will use the Gricean-inspired perspective of conversational implicature as the foundation upon which the empirical investigation presented in this thesis is carried out, both the broad framework of the Gricean-inspired perspective must be defined, and the relevant theoretical differences between its constituents explained. Moreover, the theoretical discrepancies between the constituents of the Gricean-inspired perspective calls for a common definition of 'manner implicature' and 'quantity implicature' to be established. As Grice, Horn, and Levinson account for each differently, the empirical investigation of MI from a Gricean-inspired perspective is futile unless 'manner implicature' is first defined in these terms. Moreover, an account must be offered of how the terms 'manner implicature' and 'quantity implicature' extend to specific linguistic phenomena, i.e, what types of utterances or lexical forms are considered to trigger MI or QI, as there is no consensus between the Gricean-inspired perspective.

2.3. The Gricean-Inspired Framework

The foundation of the Gricean-inspired perspective of conversational implicature is the notion that speakers and listeners enter into a mutually assumed conversational pact that comprises of a set of maxims ('principles' for Horn and 'heuristics' for Levinson) that dictate how a speaker and listener should communicate. Adherence to the pact ensures a Zipfian-

inspired cost-benefit economy in which speakers and listeners try to maximise the efficiency and harmony of information exchange (e.g., Grice’s Cooperative Principle). For the purposes of summarization, Grice’s original maxims are rehashed by Horn (1984) and Levinson (2000) in the following way:

Grice (1975)	Maxim of Quantity (1)	Maxim of Manner	[Maxim of Quantity (2)] ³
Horn (1984)	Q-Principle/R-Principle	Q-Principle/R-Principle	R-Principle
Levinson (2000)	Q-Heuristic	M-Heuristic	I-Heuristic

Figure 2-1: A comparison of the different principles used by the Gricean-inspired perspective.

Common to the Gricean-inspired account is the notion that violations of these principles give rise to conversational implicatures, whereby articulated utterances are pragmatically enriched, such that more is meant by an utterance than is literally said. In the Gricean-inspired framework, *the alternative* plays an important role in the derivation of conversational implicature. Here, the alternative is an utterance a speaker could have made, but didn’t, that avoids contravening the conversational pact, see, Step (i) in Example 2-14.

For QI the alternative is an informationally stronger lexical item, while for MI it is an unmarked, but informationally coextensive, expression. In both, the resulting implicature restricts the utterance’s meaning away from that of the alternative. As detailed in Figure 2-2, there are differences between the sets of conversational principles posited by the constituents of the Gricean-inspired perspective, such that the implicatures they generate are described in disparate terms. To summarize, the equivalents of Grice’s (1975) original framework are presented in those of Horn (1984) and Levinson (2000) as follows:

³ N.B. The Gricean equivalent for Horn’s R-Principle and Levinson’s I-Heuristic, i.e., the Maxim of Quantity’s second subordinate maxim, is not presented as distinct from the overarching Maxim of Quantity in Grice’s framework. Consequently, its rehashing as a separate principle in Horn and Levinson’s framework’s and the subsequent presentation of the Maxim of Quantity (2) as an equivalent thereof, should not be taken as indicative of any conceptual distance between Grice’s Maxim of Quantity (2) and the supramaxim to which it belongs.

Grice (1975)	Quantity Implicature (QI)	Manner Implicature (MI)	<i>Conventional or Generalised Conversational Implicature</i>
Horn (1984)	Q-Inference	R-Inference	R-Association
Levinson (2000)	Q-Implicature	M-Implicature	I-Implicature

Figure 2-2: A summary of the equivalent concepts in the Gricean-inspired perspective.

It should be noted that the equivalence of these concepts is not a one-to-one mapping and nuance exists between each perspective’s conceptualisation of Manner-based and quantity-based implicature.

Nevertheless, henceforth, the terms QI and MI will continue to be used as generic and overarching terms that target the shared elements of Manner-based and quantity-based implicature in the frameworks comprising the Gricean-inspired perspective. For the class of called ‘implicature’ accounted for by Levinson’s I-Heuristic, Levinson’s original language will be used. This is because Levinson’s framework places on a greater focus on I-Implicature, while similar concepts play secondary roles in that of Grice and Horn. However, to understand the conceptual similarities between of QI, MI and I-Implicature in the Gricean-inspired perspective, the conflicts therewithin must first be acknowledged.

2.3.1. *QI in the Gricean-inspired perspective*

Overarchingly, QI are upper-bounding implicatures, to use Horn’s terminology, and therefore define the upper-bounds of an utterance’s meaning, e.g., ‘some’ means, at most, ‘some [but not all]’. Crucially, upper-bounding implicature give rise to maximally informative implicatures that restore the informativity of utterances that violate quantity-based principles. While Neo-Gricean QI taxonomy traditionally divides quantity-based inferences into clausal QI (e.g., ‘I think he is Dutch’ +> I do not know that he is Dutch) and *scalar* QI, *scalar* QI will be treated as the primary manifestation of QI for the purposes of this thesis. Scalar QI can be subdivided into ‘proper’ scalar implicature and *ad-hoc* scalar implicature. In the former, QI is triggered by a conventionalised semantic or lexical scale (e.g., the use of the word ‘some’, as comprising part of the Horn scale <*some, all*>), while in the latter, QI involves an ad-hoc, particularised scale that arises from the conversational context (e.g., the scale <*I like Liam, I like Liam and Noel*> from Example 2-4).

However, the degree to which Gricean-inspired accounts consider conversational context a component of QI derivation and therefore the extent to which QI is a pragmatic phenomenon is debated. Sauerland (2012) divides the Neo-Gricean perspectives into three camps: the Pragmatic Accounts, the Lexical Accounts, and the Grammatical account. As mentioned, the thesis presents a series of priming studies that attempt to prime MI using QI. While the primary purpose of Experiments 1 – 6 is to investigate the nature of MI, the different accounts of QI posit different relationships between MI and QI and therefore predict different outcomes for the Experiments 4 and 5. In short, understanding MI relies on an understanding of QI, while investigating the relationship between the two phenomena can shed a valuable light on their respective natures. As such, it is important to contextualise Experiments 4 and 5 with a comprehensive overview of the competing accounts of QI.

2.3.1.1. *The Pragmatic Account of QI*

The Pragmatic Account of QI follows the traditional Gricean theory of implicature derivation, that speakers calculate scalar implicature through a process of pragmatic reasoning whereby implicature derivation is triggered by the ostensive violation of a quantity-based conversational principle and resolved by the listener with appeal to the inferred motivations of a speaker. Grice (1975), Horn (1989) and contemporary Gricean theorists (e.g., Russell, 2006, Geurts, 2009) are proponents of such an account.

According to the Pragmatic Account, to reconcile a quantity-based violation with the belief that a speaker is abiding by a mutually assumed commitment to cooperative conversation (e.g., Grice's Cooperative Principle), a listener must consider the alternative way in which the speaker could have spoken but did not. Therefore, QI is considered post-compositional, given that it relies on information that is extralinguistic or not part of the semantic or syntactic composition of a sentence (e.g, the knowledge that a particular utterance, e.g., '*I like Liam*', or lexical choice, e.g., '*some*', is under informative).

Given that QI is considered to be derived after the utterance's formal composition, quantity enrichment must apply to the utterance globally, rather than locally, according to the Pragmatic Account. However, the existence of purportedly locally enriched scalar implicature proves an issue for global accounts of QI (Levinson, 2000; Chierchia, 2004); take the following example:

- i. 'Every car has some tires.'
- ii. Here, the maximally informative alternative is:
- iii. Every car has *all* its tires.
- iv. As (II) asymmetrically entails (I), the Gricean account predicts QI elicitation, i.e., that a listener reasons that, as the speaker chose not to say (II), they must believe \neg (II).
Therefore, arriving at the belief
- v. The likelihood is that Speaker X doesn't believe (II) is true.
- vi. Assuming that the speaker is opinionated, i.e., has an opinion about the truth-conditions of the alternative (Geurts, 2009; also van Rooji and Schulz, 2004), we arrive at:
- vii. Speaker X thinks (I) \vee \neg (II)
- viii. Taken together, (II) and (III) entail that the Speaker X believes \neg (II):
- ix. \rightarrow Not every car has some but not all of its tires.
- x. Or rephrased as:
- xi. \rightarrow Every car has some tires and not every car has all their tires.
- xii. However, it is contented that the following, local, implicature arises from (I):
- xiii. \rightarrow Every car has some but not all their tires.

Example 2-15

Crucially, the global implicature (VI) allows for a car to exist with all its tires, whereas a local implicature (VII) denies the existence of any fully wheeled automobile. While evidence is divided concerning the nature of local QI enrichment, e.g., see Chemla and Spector (2011) for evidence for and Geurts and Pouscoulous (2009) for evidence against, the fact that Pragmatic perspectives cannot account for the proposed QI in (VII) is taken by Lexical and Grammatical accounts of QI to suggest that the Pragmatic Account does not fully account for QI. As a result, the Lexical Account and the Grammatical account offer perspectives that do take into consideration the existence of embedded, local, QI enrichment.

2.3.1.2. *The Lexical Account of QI*

Beginning with Levinson (2000), the Lexical Account considers scalar terms to require an enrichment process modelled on Grice's (i.e., that of Example 2-5) but with the stipulation that this calculation need not be made on an ad-hoc basis with each encounter of a scalar term but rather that the product of this calculation (i.e., QI) is semantically incorporated into the lexical item (e.g., 'some') and is therefore the default interpretation of the lexical item. For instance, in the case of high frequency scalars (e.g., 'some', 'or' and numerals) and high frequency adjectival scalars (e.g., 'warm') the intuitive implicatures (e.g., '+> some but not all' and '+> warm but not hot') comprise the lexical item's semantic composition.

This account rests on two assumptions. The first, that each scalar term results in uniform, consistent enrichment and the second, that the derivation of QI implicature affects a cognitive burden and that lexicalising the result (i.e., the enriched meaning) thereof offers a less costly communicative route. As the enrichment of scalar terms is not obligatory and can be cancelled, as evidenced by acceptability of utterances like 'She ate some, in fact all, of the cake!', Levinson (2000) proposes that the semantically incorporated enrichment is caveated with a defeasibility marker (e.g., ^D*some but not all*); however, as per Levinson, defeasance is only permitted in instances in which the enrichment is explicitly negated. Similarly, Chierchia (2004) considers non-maximal scalar terms (e.g., *some*) to be concurrently representative of their unenriched and enriched (dubbed weak and strong respectively) meanings (e.g., *some_w* [at least some] and *some_s* [some but not all] respectively) until an interpretation is selected by a listener (Sauerland, 2012).

Given that QI is lexicalized, it can be processed by default (i.e., a listener need not calculate the enrichment, in the Gricean sense) and locally (i.e., the QI is derived at the point of the lexical item, rather than globally). However, the Lexical theory can be considered to both overshoot in generating instances of unwarranted Q-enrichment and undershoot in failing to derive instances of QI. As explained in Sauerland (2012), lexicalized accounts of scalar enrichment struggle to account for SI downward monotone environments, inferences in which stronger predicates can be substituted for weaker ones. This is prototypically exemplified by instances of negation, like the following (Potts, 2013):

‘My wife doesn’t like all of the decorating that I’ve done’

+> My wife likes at least some of the decorating that I’ve done

Example 2-16

Because ‘all’ is maximally informative and does not trigger scalar QI, the Lexical Account fails to predict the enrichment in Example 2-16. However, in contrast, the Pragmatic Account predicts the reversal of entailment relations in downward monotone environments, such that, ‘all’ is said to entail ‘some’, via the negation of the alternative ‘my wife doesn’t like some of the decorating I’ve done’, in this instance.

2.3.1.3. The Grammatical Account of QI

Following from Kadmon and Landman (1993), Chierchia (2006) and Fox (2007) present seminal Grammatical Accounts of QI derivation which extend beyond the Lexical Account’s semantic explanation of QI and instead posit syntax as the locus of QI derivation; as such, the limited consideration of pragmatic interference in QI postulated by the Lexical Account is done away with, in favour of the wholly syntactical explanation that the existence of a covert exhaustivity operator gives rise to scalar QI.

Such accounts are motivated by perceived failings of the Pragmatic and the Lexical Account; namely, the observation that local enrichment can occur both globally and embedded contexts (e.g., Landman, 1998, Recanati, 2003). While experimental evidence is divided in its support of the existence of embedded scalar implicature, e.g., see Chemla and Spector (2011) for evidence for and Geurts and Pouscoulous (2009) for evidence against, the appeal of the Grammatical perspective of QI derivation its ability to account for purported instances of embedded scalars, as illustrated in Example 2-17, below.

Broadly speaking, the Grammatical account attributes Q-enrichment to a phonologically silent but grammatically present exhaustivity operator, for Fox (2007) this is ‘Exh’ and for Chierchia (2004) ‘*O*’, that, as a syntactical operator, manifests in the syntactic representation of an utterance and comprises part of linguistic expression (Chemla & Spector, 2011); this is inspired by Levinson’s (1983) observation that SI can be laid bare by the explicit inclusion of the word *only*, i.e., I ate some of the chocolate vs. I ate *only* some of the chocolate. As per Fox (2007), *Exh*(P) can be understood as the conjunction of a proposition (P) (e.g., ‘I have some of the cards’) and the epistemically strong implicatures associated with it (e.g., +> I have some, but not all, of the cards) (Sauerland, 2012). Given that *Exh* (or, indeed, *O*)

comprises part of an utterance's syntactical representation, it need not warrant global application like SI that is derived on a post-compositional level (or, pragmatic level) does. The exhaustivity operator is therefore freely insertable in the syntax and can be inserted into embedded contexts whereby it asserts the conjunction of the prejacent (i.e., a prior proposition, e.g., p = 'some') with the negation of the excludable alternatives (e.g., $\neg q$ = 'not all') giving rise to an enriched meaning (e.g. $p \wedge \neg q$, or 'some and not all') (Bassi et al., 2021) – this provides a solution to the problem demonstrated in Example 2-16, which can be illustrated as demonstrated below:

Consider the following utterance:

- i. 'Every car has some tires.'

And the globally derived implicature:

- ii. \rightarrow Every car has some tires and not every car has all their tires.

As described, it is contended that the following, local, implicature arises from (i):

- iii. Every car has some but not all of its tires.

Here, we can apply Fox's freely-insertable exhaustivity operator to the scalar term locally, giving rise to interpretation (iii):

- iv. Every car has *Exh*[some of its tires]

Example 2-17

Which is equivalent to the implicature expressed in (ii). While the Grammatical Account works to account for instances of local enrichment, which are neglected by the Pragmatic Account, it also accounts for instances of global enrichment that are neglected by the Lexical

Account in its preference for localised scalar enrichment. This can be demonstrated using the implicature in Example 2-18 – see below:

My wife doesn't like all of the decorating that I've done

+> My wife likes at least some of the decorating that I've done

Example 2-18

As explained, 'all' cannot be locally enriched by the Lexical Account, owing to its status as a maximal i.e., it doesn't carry any semantically encoded SI. However, the implicature in Example 2-18 can be derived via global enrichment via the insertion of an exhaustivity operator:

Exh[My wife doesn't like all of the decorating that I've done]

Example 2-19

Here, the *Exh* functions to conjunct [[...]doesn't like all of the decorating[...]] with the alternative [[...] doesn't like any of the decorating[...]], leading to the enriched meaning expressed in (II).

In sum, the Grammatical Account favours a departure from pragmatics as the source of QI enrichment, instead accounting for QI as a syntactical phenomenon. Despite this, the availability of a grammatical operator generally relies on context and communicative reasoning, e.g., to help speakers to coordinate logical form (Potts, 2015). Therefore, the Grammatical Account is compatible with an adherence to higher-order pragmatic principles (e.g., Grice's Cooperative Principle), although not a product thereof. Broadly, the Grammatical Account can be considered a response to the Pragmatic and Lexical perspectives' failure to account for global and local enrichment in tandem. To this end, the Grammatical Account seeks to distinguish itself from other Neo-Gricean accounts; as explicated in Chierchia, Fox and Spector (2012), the goal of the Grammatical Account is to

counter the perspectives of the Neo-Gricean accounts by rejecting the former's appeal to conversational maxims.

Owing to the use of Horn's scales in the generation of alternatives (e.g., Chierchia, 2004) and the role pragmatic information plays in the availability of a grammatical operator, the Grammatical Account has been described as presenting a syntactic Neo-Gricean perspective (e.g., de Carvalho et al., 2016; Potts, 2015). Nevertheless, given the Grammatical Account's view that QI is a fundamentally syntactic phenomena, it does not comprise part of what is called the 'Gricean-inspired' perspective for the purposes of this thesis.

2.3.2. Post-Gricean Accounts of QI

As mentioned, the Pragmatic, Lexical and Grammatical Accounts of QI can be unified by their appeal to the Gricean-inspired principle of informativeness – although, this grouping may be rejected by proponents of the latter (e.g., Chierchia, Fox & Spector, 2012). On the contrary, Post-Gricean accounts extend beyond the boundaries of Grice's seminal account. For instance, Relevance Theory seeks to reduce the various conversational principles of the Gricean-inspired perspective into one overarching principle of Relevance, most notably, (Carston, 2002; Sperber & Wilson, 1995)

Modelled on the function of the Cooperative Principle, the Relevance Principle is as follows:

- i. Human cognition tends to be geared to the maximisation of relevance.
- ii. Every act of ostensive communication communicates a presumption of its own optimal relevance.

(Sperber & Wilson, 1995)

Example 2-20

The principle of Relevance can be defined as the balance of positive cognitive effects, i.e., the meaningful assumptions, or understandings, an utterance activates for the receiver and the effect it has in providing, rejecting, or updating beliefs, and the processing effort it requires (Sbisà, 2006). For instance, in Context A, whereby a brother and sister are discussing their

favourite childhood holiday destinations, the utterance ‘I loved going to Spain’, may give rise to varied positive cognitive effects (e.g., the knowledge that the speaker is referring to a particular hotel in Spain that they travelled to each year with their extended family and the assumption that this the situation the speaker is referring to) with little processing effort. In contrast, in Context B, whereby a speaker and listener are discussing their childhoods and are unacquainted with one another, the utterance ‘I loved going to Spain’ may produce few initial positive cognitive effects, and requires substantial cognitive processing on part of the listener to achieve them (e.g., the listener must determine whether the speaker intends to communicate that they travelled to Spain as a child, must determine whether the speaker is speaking about a single or regular occurrence, whether the speaker travelled to Spain on holiday, or for other purposes, etc.) As such, the utterance in Context A is considered more relevant than that in Context B.

For Relevance Theory, conversational implicature (in Grice’s sense), occurs in two forms: implicature and *explicature*. Explicatures are truth-conditional propositional interpretations inferentially developed from an utterance’s logical form. Through either semantic narrowing or semantic broadening, a *concept*, i.e., a mental representation (Fodor, 1975), is locally embellished and interpreted through the creation of a new, contextually bound, off-the-cuff, ad-hoc concept. In semantic narrowing, the specification of linguistic meaning occurs with the addition of extra information, e.g., the utterance ‘some of the cars are green’ is interpreted with the narrowed ad-hoc concept SOME*, which means ‘some but not all’. For semantic broadening, elements of semantic meaning are suspended, e.g., ‘He told a trillion lies’ suspends the strict numerical value of ‘trillion’ in the creation of the ad-hoc concept TRILLION* which may mean ‘a very large number’ (Carston, 2005; Sperber & Wilson, 1995; de Carvalho, et al, 2016). In contrast, for Relevance theorists, implicature refers to inferences that cannot be derived from an utterance’s logical form, or *what is said*, and are instead derived pragmatically, using contextual information. Revisiting Example 1-2, while the utterance ‘Young performer, good writer, producer’ may be interpreted with the ad-hoc concept PERFORMER*, i.e., an explicature, whereby PERFORMER* encodes the explicit exclusion of singing-based performance, the communicated meaning ‘+> Alicia Keys is not a talented vocalist’, is an implicature, as it cannot be generated from the utterance’s logical form.

Crucially, for Relevance Theory, QI is a made-to-order, context-dependent (i.e., particularised) inference, the processing of which necessitates significantly more cognitive

than the processing of QI does from the perspective of the Lexical or Grammatical Account. Moreover, given that Relevance Theory departs from multiple conversational maxims to that of just one, the boundaries between QI and other types of implicature, like MI, are less distinct.

2.3.3. MI in the Gricean-inspired perspective

As a comparatively understudied phenomenon, accounts of MI do not enjoy the same theoretical depth and diversity as that of QI. As a consequence, the Gricean-inspired perspective is less divided in the way MI is accounted for. At the very highest level, the Gricean-inspired perspective considers MI to be a type of particularised, context-dependent implicature that is triggered by the violation of conversational maxims that seek to restrict the manner in which utterances are expressed, MI is lower-bounded and therefore generates implicatures that define the lower limits of what an utterance means, i.e., derive the implicature than an utterance means more than 'x'.

For Grice and Levinson, MI is the result of maxims that govern utterance form (i.e., the Maxim of Manner and the M-Heuristic, respectively). In contrast, Horn's R-Principle, by which MI is generated, functions on both utterance form and content – in the latter, restricting the oversupplying of information. Nevertheless, the Gricean-inspired perspective is unified by the perspective that MI derivation involves the negation of an unmarked, but semantically equivalent, alternative. In this respect, both MI and QI are *enrichment-via-alternatives* (EVAs). In the Neo-Gricean accounts, the alternative in MI is explicitly defined as the unmarked linguistic alternative (i.e., what could have been said, but was not) and the stereotypical associations (or implicatures) it carries; as such, MI is described as a form of marked-to-marked concordance, whereby marked utterances are interpreted to infer marked meanings.

The Gricean-inspired perspective considers linguistic markedness to manifest as the use of less frequent or less lexicalized, more morphologically complex, more prolix, or more periphrastic forms. As a consequence, MI's context-dependency is two-fold. Unlike scalar QI, which is triggered by a lexical scalar, the trigger of MI is markedness, which has both a contextual element (i.e., certain linguistic forms are marked or unmarked in different contexts, e.g., the term *Equus ferus caballus* – the trinomial name for 'horse' – may be unmarked when used in a science journal but is otherwise marked in everyday use) and a

form-based element. Moreover, the derivation of MI is also context dependent, as the markedness of the MI's referent must be calculated, too, e.g., in Example 1-2, a listener must determine the manner in which Alicia Keys is not a conventional singer.

2.3.3.1. *What constitutes MI?*

The linguistic forms that give rise to QI have enjoyed extensive documentation – see Levinson (2000), Horn and Ward (2006), Hirschberg (1991) and Chierchia, Fox, and Spector (2012) for notable overviews. In contrast, far less attention is given to the typology of MI across the literature.

Grice's seminal account offers the following illustration of MI:

- i. Miss X produced a series of sounds that corresponded closely with the score of 'Home Sweet Home'.

Example 2-21

The type of MI presented in Example 2-21 is reliant on prolixity as a vehicle of markedness, and therefore of MI, much like that of Example 1-2. Beyond this, Grice provides no concrete explanation as to which types of utterances commonly violate Manner's submaxims and trigger MI. Similarly, Horn's (1984) division of the Maxim of Manner's submaxims between the Q- and the R-Principle means that MI, in Grice's original sense, is overshadowed by the discussion of Q-Inferences that are equivalent to Grice's QI, and R-Inferences that are equivalent to Levinson's I-Implicature and when MI is discussed, the mechanism of its derivation, rather than its surface form, is accounted for (Horn, 1984, p. 22), therefore, it is unclear what constitutes MI.

Litotes:

Nevertheless, Horn (1991) explores litotes as a possible manifestation of MI; litotes are the use of the negation (e.g., 'not bad') of an antonym (e.g., 'bad') of an alternative (e.g., 'good') that carries an implicature that either weakens the alternative (e.g., +> not good, but okay) or strengthens it (e.g., +> definitely good). Litotes are considered candidates for MI on the basis that the two negations cancel themselves out and that therefore the litotes and their alternative

are semantically equivalent, a requisite for MI to occur (Rett, 2020). As such, the following phrase can give rise to the following implicatures:

‘The penalty is not unjustified’

- i. +> The penalty is somewhat justified.
- ii. +> The penalty is definitely justified.

Example 2-22

While Jespersen (1924) and Horn (1991) account for litotes as ambiguous phenomena, for Levinson (2000), only Example 2-22 (ii) is considered M-Implicature. For Levinson, (i) indeed carries the MI that ‘the penalty’ is less ‘justified’ than the alternative (i.e., ‘the penalty is justified’) would I-Implicate, although it does not meet the criteria for MI and is considered an example of I-Implicature. Plainly, the implicature in the example is the result of a euphemistic, but unmarked, avoidance of the lexical item ‘justified’ and the alternative use of a less direct, but semantically equivalent, lexical item that has a coextensive I-Implicature to that of the alternative.

Periphrasis

McCawley (1978) and Katz (1972) identify *periphrasis*, i.e., the use of more words than necessary, as a trigger of a marked-to-marked Manner-based association, which is governed by Grice’s third Maxim of Manner, Levinson’s M-Heuristic and Horn’s R-Principle. As per the examples provided by McCawley and Katz, periphrasis manifests as a structural quality. Take Katz’s (1972) illustration of a *periphrastic causative*:

- i. He killed the sheriff.
⁴I+> He murdered the sheriff, e.g., with a gun.
- ii. He caused the sheriff to die.

⁴ As per Levinson (2000) ‘I+>’ means I-Implicates and ‘M+>’ means M-Implicates.

M+> He indirectly, accidentally, or unwillingly killed the sheriff.

Example 2-23

In instances like the above, McCawley (1978) points out that there is a general association between the lexicalized causative ('killed') and situations of direct causation and an association between the periphrastic causative and situations of indirect causation (Levinson, 2000). In Example 2-22 (ii) the M-Implicature implicates the negation of the corresponding I-Implicature, i.e., the stereotypical association of direct causation with lexicalized causatives. The perspective that periphrastic causatives comprise MI is supported by McCawley's (1978) observation that periphrastic causatives with no lexicalized counterpart do not give rise to M-Implicature, presumably owing to the lack of an available alternative in the generation thereof. Consider the following example:

i. He made her laugh.

+> He made her laugh in the usual way, e.g., by telling a joke.

Example 2-24

Cross-register doublets

Levinson (2000) observes that synonyms belonging to different registers (e.g., informal vs. formal) often display the marked-to-marked and unmarked-to-unmarked associations typical of M- and I-Implicature, respectively. This is demonstrated below:

i. I study Linguistics

I+> I study Linguistics in a regular way, e.g., at a regular university, with normal intensity.

ii. I read Linguistics.

M+> I study Linguistics in a scholarly, serious manner, possibly at an elite institution.

Example 2-25

As Levinson points out, various cross-register doublets exist in English (e.g., nouns like [horse, steed] and [residence, house] or pronouns [one, I]).

Lexicalized Expressions

MI is thought to be elicited by the contrast between the use of grammaticalized or morphologized expressions and lexical expressions. In the former, information that can be expressed grammatically or morphologically (e.g., possession) is expressed, while in the latter, lexical equivalents are used. For instance:

i. [...] his house.

I+> The house that he owns and lives in.

ii. [...] the house that belongs to him.

M+> The house that he owns in the legal sense, but may not live in.

iii. [...] her child's father.

I+> the man who fathered her child.

iv. [...] the father of her child.

M+> the man who fathered her child and is unknown or not in-the-picture, etc.

Example 2-26

Here, Levinson (2000) describes that in the grammaticalized or morphologized expressions, the resulting I-Implicature suggests a closeness or intimacy between the NPs, which is reversed by the corresponding M-Implicature in the lexicalized expression.

Zero versus Non-empty Morphology

Similarly, Levinson (2000) demonstrates that the inclusion of an articulated morpheme (e.g., ‘the’, ‘a’, or ‘he’) in instances in which zero-morphemes, i.e., morphemes that are syntactically presented, but not phonologically articulated (Gerner & Ling, 2020), or ellipsis is permissible, can give rise to MI. Consider the following examples:

i. I went home after the gig.

I+> I went home in a regular way.

ii. I went to my home after the gig.

M+> I went home for a specific purpose.

M+> I went home, unplanned.

iii. After work, he went home and slept.

I+> He went home to sleep in a regular way.

iv. After work, he went home and he slept.

M+> He went home, not with the express purpose of sleeping, but did.

M+> He went back home and slept as a result of exhaustion.

Example 2-27

Here, Levinson (2000) claims that a general principle of the pragmatic division of labour associates minimal forms, e.g., instances of zero-morphemes and ellipsis, with maximally

informative, preferred meanings, e.g., stereotypical associations. In contrast, forms that deviate from the minimal form implicate non-stereotypical associations, via M-Implicature.

Repetition and reduplication

Horn (1991) and Levinson (2000) document repetition or reduplication as a means of eliciting MI. In the case of adjectival reduplication, the consequential MI often serves to intensify the adjective's meaning. Take the following, for instance:

- i. He was a strong, strong man.

M+> He was an unusually strong man.

- ii. It was a big, big mistake.

M+> The mistake was very grave.

The reduplication of nouns is also demonstrated to give rise to MI, oftentimes giving rise to a MI that narrows, and perhaps intensifies, the meaning of the duplicates' constituent, e.g.:

- iii. They had a *fight*, fight.

M+> They had a particularly violent fight.

- iv. They're in *love*, love.

M+> They love one another in a particularly strong manner.

A similar effect is achieved by the duplication of verbs or verb-phrases:

- v. He cried and cried and cried.

M+> He cried with unusual length and intensity.

- vi. She lies and she lies and she lies.

M+> She lies with unusual frequency and persistency.

Example 2-28

In the case of reduplication, the types of MI that can be elicited may be somewhat narrower than those that can be elicited by markedness, e.g., intuitively ‘*fight, fight*’ cannot elicit the MI that the referenced fight was unusual in ways unrelated to its intensity.

Evaluatives

Based on the observations of Lehrer (1985) and Heim (2007) that negative antonyms, e.g., ‘*short*’, are marked in comparison to their positive counterparts, e.g., ‘*tall*’, Rett (2014) proposes that the use of such negative antonyms in evaluative constructions gives rise to MI. Take the following examples:

i. How short is Jane?

M+> Jane is not tall.

ii. How tall is Jane?

#M+> Jane is not short.

Example 2-29

Here, Rett argues that the use of the negative antonym ‘short’ in (i), implicates the negation of its alternative, ‘tall’, by virtue of its respective markedness, giving rise to the implicature, +> Jane is not tall. In contrast, the opposite is unavailable, as demonstrated in (ii). Here, Rett’s reading of (i) is motivated by the consideration of ‘short’, or indeed negative antonyms, generally, as marked and by the consequent negation of the positive antonym, rather than by the notion that MI implicates a markedness of the utterance’s referent; (i) does not implicate +> Jane is not tall in the stereotypical sense, i.e., as predicted by Horn (1984) and Levinson’s (2000) accounts, but rather simply implicates +> not tall, without implicating that there is a markedness to Jane’s height.

To this end, Rett's definition of MI focuses on a markedness of form and a negation of the marked form's alternative but decentres the role of the alternative's I-Implicature and therefore MI's typical association with a marked referent.

2.3.3.2. *Issues for the Gricean-inspired Account of MI*

As demonstrated, despite comparatively limited theoretical focus having been extended to MI, exploration of MI has revealed it to be a heterogeneous and somewhat elusive phenomenon; MI is triggered by a diverse array of surface forms and gives rise to a diverse array of implicatures, from the rather uniform implicatures that result from periphrasis and reduplication to the more ad-hoc implicatures that result from prolixity, as per Grice's (1975) seminal observation. As a result, the Gricean-inspired concept of MI experiences a degree of internal conflict and the nature of MI, therefore, remains somewhat fuzzy.

As observed by Wilson and Katsos (2023) the role of semantic equivalence and the role of markedness in the derivation of MI pose questions unanswered by the Gricean-inspired account. To recap, when a marked utterance form is used, for that which there is an unmarked semantic equivalent, MI will be triggered. Consequently, the derivation of MI is reliant on the articulated utterance having a salient, unmarked semantic equivalent. The nature of this equivalent, however, is unclear.

Let us revisit Grice's (1975) example:

- i. 'Miss X sang 'Home sweet home.'
+> Miss X produced a series of sounds that corresponded closely with the score of 'Home Sweet Home.'

Example 2-30

Here, the terms 'sang' and 'produced a series of sounds that corresponded closely with the score' are clearly not entirely semantically equivalent; for instance, the terms lack coextensivity, while 'sang' I-Implicates (or simply, *means*) that 'Home sweet home' was performed with the voice, the alternative presented in Example 2-7 could extend to situations in which 'Home sweet home' was played with an instrument, was hummed or was even burped.

Similarly, the utterance used in Example 1-2: ‘Young performer, good writer, producer’, clearly extends to other entertainment professionals (e.g., actors) while ‘singer’ does not. However, Wilson and Katsos (2023) point out that semantic equivalence is not always predicted by pragmatic theories; for instance, Clark (1996) postulates that lexical items possess conventional meanings that give rise to a contrastive function, such that the type of paraphrasing typical of MI necessarily gives rise to semantic inequivalences between the utterance and its alternative.

However, other types of MI exhibit higher degrees of coextensivity concerning the articulated utterance and the unmarked alternative. For instance, as described by Rett (2020), litotes are logically synonymous with their unmarked counterparts, e.g., the following example from Horn (1989, 1991):

- i. ‘It’s not impossible’
 - +> It was possible, but not easy.
 - +> It was very possible.

Example 2-31

Here, ‘not impossible’ and the alternative ‘possible’ are clearly logically coextensive, referring only to situations that surpass the threshold of possibility. However, Wilson and Katsos (2023) point out that litotes can be divided into cases involving contradictories, i.e., cases in which the negated lexical item, e.g. ‘impossible’ and its alternative, i.e., ‘possible’, are exhaustive, and contraries, those in which the negated lexical item and its alternative are mutually inconsistent, but not exhaustive. As contradictories are exhaustive (a situation must be either possible or impossible) the double negation of a contradictory cannot change its extension.

However, let us revisit an example of litotes involving contraries:

- ‘The penalty is not unjustified’
 - iii. +> The penalty is somewhat justified.

iv. +> The penalty is definitely justified.

Example 2-32

While the MIs generated are analogous, the nature of the alternatives are not. Here, ‘justified’ and ‘not unjustified’ are not clearly coextensive. As described by Sapir (1944), for antonymic sets comprising gradable adjectives (e.g., ‘justified’, ‘unjustified’) there exists a ‘zone of indifference’ whereby neither term applies – for instance, a penalty can be neither justified nor unjustified. Consequently, Wilson and Katsos explain that in MIs involving contradictories, there is a semantic inequivalence between the utterance, e.g., ‘not unjustified’, and its alternative ‘justified’, whereby the former’s extension, unlike that of the latter, includes the ‘zone of indifference’.

Wilson and Katsos (2023) conclude that the apparent disparity between the alternatives in MI warrants either the restriction of MI to the limited instances in which there is a true semantic equivalence between an utterance and its alternative or theorisation that the sufficient overlap, rather than semantic equivalence, is the threshold for MI. Although unaddressed by Wilson and Katsos, MI elicited by reduplication, lexicalized expressions and non-zero morphology are also examples in which there is a coextension between the utterance and its alternative. That said, the dilemma described by Wilson and Katsos may be somewhat of a false dichotomy; it seems to be that even between MIs involving a ‘semantic equivalence’ there is a spectrum of semantic overlap between the utterance and its alternative in the case of MI, such there exists no large subdivision of MIs for which the nature of the alternative is identical. Consequently, the only conclusion that can be made is that an utterance and an alternative must have reasonable overlap, but not semantic equivalence, for MI to be elicited.

2.3.3.3. *Markedness*

Given that MI derivation is reliant on a markedness of form, defining what *markedness* means in this context is important for developing a comprehensive theoretical understanding of the phenomenon.

In general, linguistic markedness is a well-researched concept, with rich traditions in disciplines of phonology, syntax and semantics (Andrews, 2012). Yet, as described by Wilson and Katsos (2023), the nature of markedness’ role in MI remains unclear. For the Gricean-inspired perspective, the nature of markedness appears to be two-fold, on one hand,

markedness can be conceived of as a product of utterance length, i.e., the degree of morphological complexity, prolixity, periphrasis (Levinson, 2000; Horn, 1984), and brevity (Grice, 1975), and on the other hand, markedness is conceived of as a context-bound quality, as existing in relation to the comparative unmarkedness of more usual or more frequent alternatives. As explained by Wilson and Katsos (2023), taken together these perspectives comprise a conceptualisation of markedness rooted in the conventionality of language – the idea that linguistic knowledge is defined by socially-constructed, arbitrary conventions (Diesendruck & Markson, 2011; Kalish & Sabbagh, 2007). Rett (2020) expands this view of markedness to include the qualities of narrowed breadth of cross-linguistic and language-internal distribution, later acquisition, greater processing cost and shorter diachronic lifespan.

Contrary to views of markedness that posit length-based qualities as vehicles thereof, Carston (2005) observes longer forms (e.g., *Can you pass the salt and pepper?*) are sometimes less marked than more perspicuous alternatives (e.g., *Can you pass the condiments?*). Similarly, Levinson (2000) and Horn (1989; 2007) point out that periphrastic causatives with no unmarked alternative (e.g., ‘He made her laugh’) are not considered marked. As explained by Wilson and Katsos (2023), such observations prove troublesome for accounts that define markedness in formal, but not distributive, terms (e.g., Franke, 2009) and for accounts that either offer too vague an account of markedness (e.g., Grice, 1975; Horn, 1989, 2007) or an account of ‘markedness’ as a plethora of heterogeneous concepts (e.g., Levinson, 2000; Rett, 2020), as such accounts will inevitably fall short in their ability to provide a comprehensive account of MI. Consequently, Wilson and Katsos (2023) call for clearer delimitation of markedness in theoretical accounts of MI and the demarcation of markedness as either a length-based or frequency-based phenomenon.

Another inconsistency of the Gricean-inspired programme’s conceptualisation of markedness is whether markedness exists as a cognitive effect or as a formal quality of an utterance. For instance, while Grice’s (1975) Maxim of Manner governs linguistic form (e.g., in the submaxims ‘be brief’ and ‘be orderly’), the submaxims ‘avoid ambiguity’ and ‘avoid obscurity’ seem to be defined in terms of the cognitive effects an utterance should avoid eliciting in a listener, e.g., unexpectedness, uncertainty, etc. As the violation of these maxims gives rise to markedness, markedness in the Gricean-inspired perspective must comprise, or at least elicit, some type of cognitive effect and can neither be wholly formal (i.e., length-based) nor distributive (i.e., frequency-based). Or, in other words, markedness is only markedness when realised as such by the listener. Consequently, the Gricean-inspired account

fails to give a clear locus for markedness; by some measures, markedness exists syntactically (e.g., *zero-* vs. *non-empty* morphology), semantically (e.g., *cross-register doublets*), pragmatically (e.g., litotes) and meta-linguistically (e.g., as the result of a cognitive effect induced by ambiguous or obscure utterance).

2.3.3.4. *What constitutes MI?*

Finally, another issue facing theoretical accounts of MI derivation is scope of what constitutes MI itself.

In Grice's seminal description of MI, MI is accounted for as comprising of the negation of the stereotypical associations of the unmarked alternative and the ensuing implicatures that follow thereafter. Take the following example:

'She caused the car to crash'

- i. +> She crashed the car in an unusual way.
- ii. +> She crashed the car with reduced agency, e.g., was not behind the wheel.

Example 2-33

Here, we have an example of a MI and two corresponding MIs; while the latter is the intuitive MI, the former is the 'first-level' MI, and is entailed by the latter, i.e., the implicature that the I-Implicature associated with the unmarked alternative is not applicable. As per Grice's initial analysis, both implicatures comprise MI and a distinction between the two is not made.

However, as per the derivational steps posited by the neo-Gricean perspective (refer back to the steps illustrated in Sections 2.2.1 and 2.2.2), only implicature (ii) is directly accounted for. However, Levinson (2000) is sensitive to this distinction, presenting the disparate implicatures as follows, albeit inconsistently:

- i. +> She crashed the car in an unusual way.
- ii. ++> She crashed the car with reduced agency, e.g., was not behind the wheel.

Example 2-34

Here, the symbol ‘++>’ denotes the sum of what is said, or the total communicated content. While the theoretical distinction between the two types of implicature is not explicated, Levinson (2000) alludes to the contribution of the Gricean notion of Relevance in the generation of the latter.

As a result, it is unclear whether MI is conceived of as simply a marked-to-marked association, or corollary to I-Implicature, or whether the concept of MI does extend to both the primary (+>) and secondary implicatures (++>) it intuitively generates. If the former, MI need not be conceived of as a particularised, ad-hoc phenomenon, but rather as more similar to QI in its derivation, but with a form-based, rather than lexical, trigger. If MI does indeed comprise both primary and secondary implicatures, then a comprehensive account of implicature must explain the process by which contextual information is integrated in the derivational process.

This observation leads to a second issue for the Gricean-inspired account of MI. While the type of contextual information needed to generate secondary implicatures could be integrated at some point in the process of MI, the optimal point at which it could be integrated is unclear. Take the latter two stages of the neo-Gricean derivation of MI, adapted from Horn (1984):

- i. The unmarked alternative tends to be associated with an unmarked situation (*s*), which represents the stereotypical or salient features or elements of the extension of the set E/E^* .
- ii. In contrast, the marked alternative, E^* , is associated with the complement of *s* with respect to the original extension of E/E^* .

Example 2-35

Here, there are two possible points of integration. On one hand, contextual information could be integrated in Step II. However, the referent(s) of E^* (i.e., the marked utterance) is already defined here, as the complement of the unmarked alternative – at this point, the implicature ‘+> not in the usual way’ can be generated.

Either there exists an undescribed Step (iii), whereby E^* is narrowed to include only contextually-relevant information (e.g., that the driver crashed the car with reduced agency)

or the situation (s) in Step I is narrowed such that contextually-relevant information is excluded from its complement (e.g., if salient features of stereotypical car crashing, i.e., (s), are 1) agency and 2) being behind the wheel of the car, these will be excluded from the complement of (s), i.e., E*). Nevertheless, a Step (iii) is not accounted for by the neo-Gricean perspective, nor is it clear how or why the stereotypical associations of the unmarked alternative would be pre-emptively modulated to generate the associations required to generate the MI at hand in the marked utterance. As such, neither Step (i) nor Step (ii) seem appropriate points as which the integration of contextual information could occur and it remains unclear how contextual information is integrated in the Gricean-inspired perspectives to generate implicatures like that of Example 2-34 (ii).⁵

To summarise, there remain theoretical grey areas concerning certain aspects of MI typology in the Gricean-inspired framework; namely concerning the nature of the semantic equivalence between the MI eliciting utterance and its unmarked alternative, the nature and role of markedness in the derivation of MI, the demarcation of MI and other pragmatic inference and the subsequent consequences this may have for the integration of contextual information in the derivation of MI. While this subsection serves only to illustrate and identify inconsistencies within theoretical approaches to MI, they are relevant to the investigation of MI presented in this thesis,

2.3.4. Post-Gricean Accounts of MI

MI is treated in much the same way as QI by Relevance Theory's (e.g., Carston, 2005; Sperber & Wilson, 1995) account of implicature derivation. As per Relevance Theory's account, MI is a particularised, ad-hoc explicature that is derived via a narrowing of the articulated utterance's logical form. However, much like the distinction between the so-called primary and secondary implicatures associated with MI described in *What constitutes MI?*, Relevance Theory distinguishes between explicatures that are derived from the utterance's

⁵ N.B. the derivation steps discussed, as presented in Example 2-8, for example, should not be understood as explicit, chronological cognitive steps, but rather as an illustration of the process required to derive MI. As explained by Levinson (2000), the process of implicature derivation involves the coarticulation of various other mechanisms, e.g., that of context sensitivity, that pattern together to derive what is understood by an utterance. That said, the absence of the explicit inclusion of a step that contributes contextual information to MI derivation means that the Gricean-inspired perspective's account of MI derivation fails to capture a critical element thereof. To that end, the critique presented is not that the Gricean-inspired perspective cannot account for how contextual information is integrated, indeed the steps presented in Example 2-8, for example, are neither fixed nor explicitly chronological, but rather that it does not.

logical form, e.g., ‘She caused the car to crash’ can be narrowed to encode only unusual car crashes, and implicatures that go beyond the narrowing of information encoded by the utterance’s logical form and require contextual information to be derived, e.g., the encoding of information pertaining to the subjects’ agency in the car crash.

The difference between implicature and explicature is broadly analogous to the distinction made by Levinson (2000), between what I have called primary and secondary MIs. Nevertheless, the Gricean-inspired approach relies on a metalinguistic alternative and therefore does not make the same distinction between implicature and explicature as the latter is reliant on the former, as per the Gricean-inspired perspective. Moreover, Relevance Theory does not consider MI derivation to occur via the use of an alternative.

2.3.5. I-Implicature in the Gricean-Inspired Perspective

Before taking stock of the similarities and differences between QI and MI, both within the Gricean-inspired perspective and between competing theories, I will briefly outline the nature of I-Implicature in the Gricean-inspired perspective.

Of the accounts comprising the Gricean-inspired perspective, Levinson (2000) gives the most comprehensive account of I-Implicature and the only account thereof that explicitly considers I-Implicature to be a type of conversational implicature. For Levinson, I-Implicature is the association between an unmarked expression and a stereotypical meaning. As previously described, the I-Heuristic gives rise to implicatures like the following:

- i. ‘He was dancing in the nightclub.’
+> He was dancing in a usual, contextually-appropriate way, e.g., not dancing the waltz.

Example 2-36

As with MI, Levinson (2000) accounts for I-Implicature as having both primary and secondary manifestations, although the boundaries between the two are somewhat ill-defined. Levinson’s conceptualisation of I-Implicature seems to encompass two, semi-distinct, types

of inference. The first is the *inference to stereotype* (Atlas and Levinson, 1981), whereby I-Implicature specifies that the use of unmarked language denotes a stereotypical or unmarked situation. Here, the I-Implicature implicates a stereotypical relationship between a lexical item and a semantic meaning. Given that the semantic meaning stereotypically associated with a lexical item must contextually defined (e.g., in some contexts, ‘dancing’ may stereotypically refer to waltzing), Levinson (2000) uses the ‘++>’ denotation to illustrate this type of implicature, suggesting that it extends beyond a basic, form-to-meaning association.

In contrast, Levinson (2000) presents a host of examples of I-Implicature that do not implicate a stereotypical relation between a lexical item and a semantic meaning, but rather the stereotypical relation between a form, or structure, and a type of meaning. These include the following, a.o.:

Noun-noun compounds

- i. ‘The karaoke microphone hire’
+> The place that hires out microphones that amplify the sound of karaoke.

Example 2-37

Preferred local coreference

- i. ‘Joe ate the cake and he was happy’
+> Joe₁ ate the cake and he₁ was happy

Example 2-38

Negative strengthening:

- i. ‘Joe didn’t like the karaoke’
- ii. +> Joe positively disliked the karaoke.

In these examples, the I-Implicature serves to implicate that there is a stereotypical relationship between the form or structure used, e.g., noun compounds, and a meaning, i.e., that the leftmost constituents of a noun phrase modify the rightmost. However, inferences to stereotypes implicate that there exists a stereotypical relationship between a lexical item, e.g. ‘dancing’ and a particular semantic meaning, e.g. the waltz. Here, the implicature does not select a stable association between form and meaning, e.g., the association between non-periphrastic causatives and agency or intentionality on the part of the utterance subject, but rather selects a contextually-dependent associate of the lexical item, e.g., that ‘dancing’ may refer to waltzing in the context of a particular dance class.

In one sense, all instantiations of Levinson’s (2000) I-Implicature can be boiled down to an implicature to the effect of ‘ $+> x$ ’, in the most specific and stereotypical way within reason’. On a more granular level, there seems to exist two types of I-Implicature, as described above; the first, an implicature that associates specific constructions, e.g., noun compounds, and forms to specific stereotypical types of meaning, and a second that associates non-specific lexical items (or, entire utterances) with non-specific stereotypical, semantic meanings. Although I draw a distinction between these examples of I-Implicature given by Levinson (2000) on the basis that the latter require the far greater use of contextual information in their generation, I tentatively suggest a way in which their derivation could be streamlined in 7.4.3.1.

R-association is the analogue of I-Implicature in Horn’s (1984) account, whereby the R-Principle’s maxim ‘Say no more than you must’ associates minimal, unmarked expressions with more specific, stereotypical meanings and encompasses a set of phenomena overlapping with those accounted for by Levinson (2000). I-Implicature (or indeed, R-association) has no clear equivalent in Grice’s original framework and the phenomena encompassed by the label are likely divided in their explanation between Grice’s *conventional implicature* (i.e., implicatures that are tied to certain lexical items, e.g., the contrastive implicature of ‘but’) and *generalised conversational implicature* (i.e., implicatures that are the product of conversational maxims, but are less reliant on context for their derivation).

Nevertheless, a crucial similarity between the treatment of I-Implicature within the Gricean-inspired perspective is that such inferences are not considered calculated, per se, and are

therefore not generated in the same way QI and MI are. As Levinson (2000) describes, this is owing to I-Implicature's status as a product of the adherence to conversational maxims (i.e., the I-Heuristic, the R-Principle and Grice's second Quantity Maxim and the Maxim of Manner), rather than the violation thereof, as QI and MI are.

2.3.5.1. *Issues for the Gricean-inspired account of I-Implicature*

The degree to which I-Implicature can be considered an example of conversational implicature, at least in the Gricean sense, is debatable and is acknowledged by Levinson (2000). Here, Levinson's and Grice's (1975) delineation of *what is said*, i.e., *the coded* for Levinson, and *implicature* is outlined, with Levinson ascribing a larger swath of phenomena to the latter. Grice's (1975) implicature typology describes the qualities of *defeasibility*, *non-detachability*, *context-dependency*, *non-conventionality* and *calculability* as the more or less essential qualities of conversational implicature. However, these qualities are observable to varying degrees when instances of I-Implicature are considered.

While I-Implicature is indeed defeasible, the properties of context-dependency, non-conventionality and calculability are far more problematic for I-Implicature. Below is an illustration of the defeasibility of Example 2-38's I-Implicature when additional information is supplied:

- i. 'Joe ate the cake and he was happy'
+> Joe₁ ate the cake and he₁ was happy

- ii. 'Joe ate the cake and he was happy – not Joe himself, but rather his father'
+> Joe₁ ate the cake and he₂ was happy – not Joe₁ himself₁, but rather his father₂.

Example 2-40

In contrast, the non-detachability of I-Implicature of is less evident. While implicature must not necessarily be non-detachable, indeed MI is not (Levinson, 1983), I-implicature demonstrates a rather consistent, multidirectional detachability. With MI, the vehicle of the implicature is utterance form, such that the switch from a marked form (e.g. prolix

description) to an unmarked form (e.g., a single word), renders the MI unavailable. That said, the wording of the MI eliciting utterance could be altered, and the MI would be still available, providing the utterance's markedness is sufficiently salient, see Example 2-41 below. In contrast, I-Implicature is more intimately attached to the constructions, forms and lexical items that elicit it, resulting in a far higher degree of detachability.

- i. 'He was wearing a piece of multi-coloured fabric tied around his collar'
+> He was wearing an unusual, or marked, tie.
++> His tie was scruffy, or homemade, or bohemian in style.
- ii. 'From his collar hung a colourful strip of silk'
+> He was wearing an unusual, or marked, tie.
++> His tie was scruffy, or homemade, or bohemian in style.

Example 2-41

As per Clark's (1996) perspective that synchronic synonyms are often near synonyms, rather than semantic equivalents, it follows that I-Implicature varies when lexical items are swapped out for one another. Take the following inference to stereotype:

- i. 'He chased the man'
- ii. 'He pursued the man'

Example 2-42

Here, it seems likely that (i) may implicate a more playful pursuit, while the use of 'pursue' in (ii) seems to intuitively allow for I-Implicatures that encode a sinisterness, a doggedness or a more goal-oriented approach to the subject's pursuit. In contrast, however, the difference between synonyms like 'drink' and 'beverage' is less clear in its impact on the types of I-Implicatures available. The 'horizontal' or sideways movement between synonyms can lead to the detachability of I-Implicature, which is not a quality shared by MI – indeed, providing a MI-eliciting utterance is unmarked, the nature of its unmarkedness does not lead to the

detachability of the MI. In contrast, I-Implicature and MI share a ‘vertical’ detachability, whereby movement along an axis of specificity leads to the detachability of the I-Implicature, e.g., if we consider the alternatives ‘John’, ‘the man’ and ‘the person who is male’, there is no shared, static I-Implicature, and the latter likely gives rise to MI, instead.

Context-dependency and non-conventionality are two intertwined properties not experienced by all forms of I-Implicature. In the case of Example 2-36’s use of ‘dancing’, a variety of I-Implicatures can be generated depending on the context in which the utterance is said, e.g., ‘+> he was dancing the tango’, ‘+> he was casually moving to the music’, or ‘+> he was fist-pumping enthusiastically’. In contrast, in less contextually-rich circumstances, e.g., if used in a general conversation, the I-Implicature associated with ‘dancing’ is likely rather uniform.

Take the following:

- i. ‘I like dancing’
+> expressed when discussing likes and dislikes with a date.
- ii. ‘Dancing improves people’s moods’
+> as mentioned by a therapist to their patient.
- iii. ‘There’s a programme about dancing on later’
+> a conversation between a couple.

Example 2-43

Here, while additional context might narrow the types of I-Implicatures generated, when said in the thinnest, most neutral contexts, it seems that lexical items can I-Implicate reference to the abstract, but prototypical, concept stereotypically associated with the lexical item, one that excludes fringe or marked cases thereof.

Consequently, while I-Implicature is clearly modulated by context, its elicitation is not dependent on it, as lexical items, even in contextually-thin contexts, still elicit I-Implicature. Moreover, when these implicatures are elicited, they do not rely on context for their calculation, but rather access a contextually-independent, generic or stereotypical meaning. Here, the non-conventionality of I-Implicature is called into question, as there is evidently an intimate link between the types of I-Implicatures generated in generic contexts, like those of

Example 2-43, and the conventional meaning of the lexical items that elicit them. In contrast, the following I-Implicature is clearly not a direct product of the conventional meaning of the lexical item:

- i. 'I'm walking to Grantchester tomorrow'
+> I'm walking to Grantchester tomorrow, in the usual way,
++> i.e., after the Sunday church service in town and with a pit-stop in the local pub for a pint.

Example 2-44

However, the types of I-Implicature detailed by Levinson (2000) that are elicited by certain constructions or forms, e.g., negative strengthening or conjunction buttressing, make a stronger case that I-Implicature should neither be considered context-dependent nor non-conventional given that the I-Implicatures are elicited by default, irrespective of context (unless cancelled) and have a one-to-one association with a specific convention, e.g., a lexical item(s) or grammatical construction, which suggest the I-Implicatures may comprise part of the forms' conventional meanings or are, at the very least, not entirely distinct from them.

Finally, the degree to which I-Implicature is calculated is uncertain. Levinson (2000) is clear in his claim that the type of implicatures derived via I-Implicature are not reliant on the same mental algorithms to compute them and are therefore not governed by those of QI and MI, which are, for the most part, analogous and rely on the consideration of an alternative way of speaking. On one hand, some calculation is required to generate contextually-specific I-Implicatures, like that of Example 2-44, otherwise it is unclear how the communicated, contextually-specific information would be integrated into the interpretation of utterances that generate such I-Implicatures, as a listener is unlikely to derive the I-Implicature without calculating the speaker's intentions.

On the other hand, as per Levinson's (2000) account, I-Implicatures do not seem to require the contemplation of an alternative way of speaking in their generation and, at points, seem to align with the preferred, default or conventional interpretations of certain forms or lexical items. Consequently, the type of calculation required to derive the I-implicature '+>[...] the normal pet dog that I own' from the utterance 'I love my dog' is unlikely to involve either the

same mechanisms, or the same level of cognitive processing, as those required to either derive QI or MI. As such, it is unlikely that I-Implicature, in the general sense, can be calculated in Grice's (1975) original sense, although, certain iterations thereof do seem to require the contemplation of the utterance in regard to its context.

As demonstrated, and as detailed by Levinson (2000), I-Implicature does not fit neatly into Grice's definition of conversational implicature. As identified in the application of Grice's diagnostics to Levinson's I-Implicature, and as described by Levinson (2000) himself, I-Implicature can be viewed from two perspectives; one that concerns the most basic implicature, as denoted by the indication '+>', and the second, the more contextually-dependent implicature, as denoted by the indication '++>'. This said, Levinson's distinction between the referents of each denotation is rather inconsistent and some I-Implicatures that are clearly reliant on a great deal of contextual information, e.g., inferences to stereotype, are considered the more basic form of I-Implicature. Consequently, as described with regard to MI, it is unclear what constitutes I-Implicature in the Gricean-inspired perspective. As such, the degree to which I-Implicature is considered *implicature* is unclear, as the typology of I-Implicature is heterogeneous, ranging from conventional to seemingly context-dependent; this presents a problem in the pursuit of a comprehensive account of I-Implicature.

Further to the extent to which I-Implicature comprises conversational implicature, Levinson (2000) fails to account for two other aspects of I-Implicature typology. Firstly, the role that markedness, or a lack thereof, plays in the generation of I-Implicature is unclear. On one hand, for Levinson (2000), I-Implicature is not metalinguistic and therefore does not involve the contemplation of how a speaker has spoken. On the other hand, I-Implicature is a product of a lack of markedness, as it is elicited by unmarked forms or utterances. Given these two perspectives, it is unclear how unmarkedness functions as both an unprocessed, or unnoticed, linguistic quality and a trigger of I-Implicature in I-Implicature eliciting utterances.

That said, contrastive uses of I-Implicature contradict this, e.g.:

- i. A: Do you identify as female?
B: I am a woman, yes.
+> I am a woman in the stereotypical sense.

Example 2-45

If we take Speaker B's use of 'woman' as it contrasts to A's description, Speaker B clearly seeks to assert the I-Implicature associated with the term 'woman'. Clearly, this must be calculated by A, which requires an acknowledgement of the difference in markedness between B's utterance and A's use of the unmarked 'woman', i.e., the comparative unmarkedness of 'woman' must be salient. Consequently, as described in the case of MI, an account of markedness is therefore required in the construction of a theoretically robust account of I-Implicature. Such an account cannot simply account for markedness, however, but also unmarkedness as a separate theoretical concept, given its importance in the generation of I-Implicatures like that of Example 2-45 and unclear importance in the generation of I-Implicature in general. A comprehensive account of unmarkedness, as it pertains to I-Implicature, should seek to identify whether unmarkedness is simply the lack of observable qualities of markedness, e.g. as per many traditional perspectives (Andrews, 2012), or whether unmarkedness should perhaps be thought of as a type of conventionality whereby certain lexical items, forms or constructions possess salient, identifiable qualities that triggered I-Implicature. If the latter is the case, then these qualities should be a taxonomized.

Depending on whether or not I-Implicature requires a trigger, e.g., a salient unmarked expression, and whether this comprises part of a process of implicature calculation, a second issue presented by Levinson's (2000) account of I-Implicature is the extent to which I-Implicature occurs in natural conversation. If we consider examples like Example 2-44, it is clear that I-Implicatures do manifest in ad-hoc, contextually-dependant instantiations and serve a similar communicative function to that of QI and MI – in this sense, presumably following a similar distribution and use to that of conversational implicature, in general.

However, if we take the alternative perspective that I-Implicature is not (or is less) context-specific than other types of implicature, e.g., focussing on form-based I-Implicature like *periphrastic causatives*, it is unclear whether or not this means that all linguistic expressions give rise to I-Implicature and carry the type of generic, prototypical implicatures by Example 2-44 at the bare minimum. Given that I-Implicature is a product of the adherence to conversational maxims, rather than the violation thereof, it would not be surprising if they were very frequent. If this is the case, I-Implicature should be reconceived as something separate to QI and MI and afforded the same treatment that the adherence to other conversational maxims is (or is not).

For instance, the basic I-Implicature, ‘+> [...] in the usual, or stereotypical way’ is analogous what one could imagine an R-Implicature (Relevance), e.g., ‘+>[...] in a way that pertains to the goals of this conversation’, or a Q-Implicature (Quality), e.g., ‘[...] in a way that accurately represents my beliefs’ looks like. Yet, these are not afforded the same theoretical weight extended to I-Implicature by Levinson (2000). Alternatively, there may be merit in better delineating I-Implicature that functions as the undetected adherence to the I-Heuristic (or a similar maxim) and from I-Implicature that functions more ostensively and requires more calculation and ultimately behaves more similarly to other types of conversational implicature.

In sum, a comprehensive account of I-Implicature must explicate the extent to which I-Implicature is considered an example of conversational implicature and whether this extent differs to that of QI and MI, account for the phenomena that constitute I-Implicature (e.g., the extent to which contextually-bound I-Implicature are *I-Implicature*), formalise the role of markedness and clarify the frequency and role of I-Implicature in general communication.

2.3.6. The Similarities and Differences Between QI and MI in the Gricean-Inspired Perspective.

Having taken stock of the nature of QI and MI in the Gricean-inspired perspective, I will briefly summarize the differences and similarities between the typology of QI and MI and that of I-Implicature.

At the broadest level, QI and MI are both considered types of conversational implicature and are elicited by the violation of conversational maxims. Generally, QI and MI are governed by separate maxims, the former, the product of a conversational maxim that seeks to ensure sufficient information is given, e.g., Grice’s Maxim of Quantity, the latter, a product of a conversational maxim that governs the manner in which utterances should be expressed, for the purposes of ensuring clarity, e.g., Grice’s Maxim of Manner. Given that the Gricean-inspired perspective accounts of QI and MI as a consequence of maxim violation, they are calculable and are subsequently derived by the listener with appeal to a consideration of what the speaker could have said but chose not to. To that end, QI and MI are derived in the following way:

<u>Manner Implicature</u>	<u>Quantity Implicature</u>
1) Speaker said 'x'	1) Speaker said 'x'
2) Speaker chose to say 'x' instead of a more <i>conventional</i> alternative (y)	2) Speaker chose to say 'x' instead of a more <i>informative</i> alternative (y)
3) If Speaker meant 'y', they would have said it.	3) If Speaker meant 'y', they would have said it.
4) Assuming the Speaker is cooperative, the fact that they did not say the alternative means '¬ y'	4) Assuming the Speaker is rational, the fact that they did not say the alternative means '¬ y'
5) Speaker means 'x and ¬ y'. In other words, the Speaker seeks to communicate the explicated utterance with the implicature that they mean it 'not in a <i>conventional</i> way'.	5) Speaker means 'x and ¬ y'. In other words, the Speaker seeks to communicate the explicated utterance with the implicature that they mean 'not the more <i>informative</i> alternative'.

Figure 2-3: MI and QI implicature derivation in the Gricean-inspired perspective.

Beyond their calculability, QI and MI are also considered to be defeasible, context-dependent and non-conventional (Grice, 1975). Moreover, Rett (2020) supplies *reinforceability* (e.g., that reiterating the implicature's content is non-redundant, e.g., 'I like some, but not all, of the episodes' (Horn, 1972) and *variation with discourse status*, e.g., the notion that implicatures cannot be cancelled if they are pertinent to the Question-Under-Discussion (QUD) as further implicature diagnostics. Although MI is undoubtedly more beholden to the input of contextual information in its derivation, *ad-hoc* QI also requires the integration of contextual information at some point in the derivation process, namely in the generation of the contextual-bound alternatives. On the flip side, MI and QI are considered different in that MI is detachable while QI, at least by Grice (1975), is considered non-detachable. Nevertheless, Rett (2020) cites evidence that suggests that QI, too, may exhibit features of detachability, e.g., scalars in the case of downwards-entailing contexts (Ladusaw, 1979) and inconsistencies between the types of QI generated by tautology, ultimately suspending detachability as a property of conversational implicature.

Another source of conflict between accounts of QI and MI is the nature of the alternative in each. In the former, the alternative is more informative than, and entailing of, the articulated utterance, in the latter, the alternative is semantically equivalent, but marked in comparison, to the articulated utterance. Moreover, while there is a semantic equivalence between the articulated utterance and its alternative in the case of MI, the alternative is associated with (or implicates) a stereotypical subset of qualities that are associated with the concept referenced by both the articulated utterance and its alternative (e.g., *dog* by ‘four-legged canine’ and ‘dog’). Consequently, in order to generate the appropriate alternative, i.e., both the correct lexical item and the correct stereotypical associations, the derivation of all MI requires appeal to contextual information. In contrast, in the case of scalar QI, the generation of the more informative alternative is conceived of as a product of a Horn scale, and is accessed semantically, rather than as a product of contextual information.

A consequence of the divergent nature of the alternatives in QI and MI derivation is that the implicatures thereof are divergent in their boundedness. In the Gricean-inspired perspective, QI is upper-bounded and therefore defines the upper-bounds of an utterance’s meaning, e.g., ‘some’ means, at most, ‘some [but not all]’. Here, an upper-bounded alternative, i.e., ‘all’ is negated, which sets the upper-bounds of what ‘some’ can mean. In contrast, MI is lower-bounded. As such, the MI generates implicatures that define the lower limits of what an utterance means because they negate a lower-bounded alternative, e.g., ‘The man who is my father’ may negate the unmarked, lower-bounded alternative ‘my dad’ to implicate that the utterance means more than just ‘my dad’, in the stereotypical sense.

Moreover, MI is considered a type of particularised conversational implicature (PCI), i.e., an implicature that arises on an ad-hoc basis, in a specific conversational context. Conversely, scalar QI is considered a form of generalized conversational implicature (GCI) by the Gricean-inspired perspective, with ad-hoc QI considered to exist in a borderline space between GCI and PCI, by Levinson (2000). Broadly, the Gricean-inspired perspective considers MI to be a non-default, made-to-order phenomenon that occurs when contextually warranted. In contrast, Horn (1972, 1984, 2007) and Levinson (2000), alongside other Neo-Gricean theorists, e.g., Chierchia (2004), Landman (1998), who consider scalar QI to be generated by default, in a manner distinctly less cognitively effortful to that of MI derivation. Nevertheless, in Grice’s (1975) original framework, QI is considered a non-default phenomenon, a view maintained by contemporary Gricean accounts, e.g., Geurts, 2009, Russell, 2006, a.o. A related distinction is the degree to which the enrichment of MI and QI

occurs locally. As per the Gricean-inspired perspective, MI must occur globally, whereas Levinson (2000) and Horn (1989; 2007) account for the local enrichment of QI, although this does not concur with Gricean accounts thereof (e.g., Grice, 1975; Russell, 2006; Geurts, 2009).

Below is an overview of the similarities and differences between QI and MI, as per the Gricean-inspired perspective (N.B., contentious qualities are asterisked):

<u>Qualities of Implicature</u>	QI	MI
Defeasible?	✓	✓
Non-Detachable?	✓	✗
Non-Conventional?	✓	✓
Context-dependent?	✓	✓
Generalised?	✓ *	✗ *
Lower-bounded?	✓	✗
Locally enriched?	✓ *	✗
Semantic equivalence of the alternative?	✗	✓

Figure 2-4: Table of qualities shared or not shared between QI and MI as per the Gricean-inspired perspective.

2.3.1. The Similarities and Differences Between QI, MI and I-Implicature in the Gricean-Inspired Perspective.

The differences between QI and MI and I-Implicature have not experienced a large degree of theoretical exploration beyond that of Levinson’s (2000) explanation of the nature of I-Implicature; this is not unexpected, however, as I-Implicature is not afforded the same importance by other Gricean-inspired and Post-Gricean accounts as it is by Levinson.

QI and I-Implicature, unlike MI, share an upper-boundedness. I-Implicature defines the upper-bounds of what an utterance can mean, i.e., ‘x’ means, at most, ‘x in the stereotypical sense’. However, unlike QI, the inferences generated by I-Implicature are not inferences to a more precise interpretation. While Levinson (2000) believes that QI implicature generates more precise interpretations, e.g., ‘+> some, but not all, of the cars are green’ from ‘some of the cars are green’, that comprise a subcase of *what is said*, this is not the case for I-Implicature and MI. Instead, I-Implicature and MI are not considered to generate more

precise interpretations of *what is said*, but rather implicate specialisations of *what is said*. Simply, Levinson considers QI to increase the exactness of *what is said*, while I-Implicature and MI move it away from a generic meaning towards a specialised meaning, i.e., a stereotypical one or an atypical one. In sum, QI and I-Implicature and MI interact differently with *what is said*.

Unlike QI and MI, with I-Implicature is positive in character, the extension of what is implicated, e.g., ‘+> I’m walking to Grantchester tomorrow, in the usual way’, is a proper subset of the extension of *what is said*, e.g., ‘I am walking to Grantchester tomorrow’; this subset is restricted positively, i.e., it is defined by what is included in the subset, e.g., stereotypical associations. In contrast, QI and MI are negative in character, while the extension of what is implicated is a proper subset of *what is said*, this subset is negatively restricted, i.e., defined by what is excluded, because the complement thereof is negated by the implicature, e.g., scalar QI’s restriction of *all* in the implicature ‘+ some but not all’ and MI’s negation of the corresponding I-Implicature, i.e., ‘+> *x* but not in the stereotypical way’ (Levinson, 2000).

I-Implicature and MI both rely on a contextually-based understanding of the stereotypical associations of particular lexical items, constructions or forms. Therefore, I-Implicature and MI require non-linguistic, or at least sociolinguistic, information to be derived, e.g., the knowledge of the culturally-bound stereotypical associations of *dog*. In contrast, QI does not rely on any appeal to comparable information in its resolution, rather the lexically stored scale-mates of the scalar term, e.g., knowledge of the presence of ‘all’ in the Horn scale <*some, all*>. Consequently, QI is metalinguistic, as it requires reference to what could have been said, but was not, e.g., the alternative *all*, to the scalar *some*. MI is equally metalinguistic, as its generation requires knowledge of the associations of the articulated utterance’s alternative, i.e., the corresponding I-Implicature, as these are negated to derive the MI.

Below is an overview of the similarities and differences between QI, MI and I-Implicature, as per the Gricean-inspired perspective (N.B., contentious qualities are asterisked, while undocumented qualities are marked with ‘?’):

<u>Qualities of Implicature</u>	QI	MI	I-I ⁶
Defeasible?	✓	✓	✗
Non-Detachable?	✓	✗	✓
Non-Conventional?	✓	✓	✓
Context-dependent?	✓	✓	✗
Generalised?	✓ *	✗ *	✓
Lower-bounded?	✓	✗	✓
Locally enriched?	✓ *	✗	✓
Semantic equivalence of the alternative?	✗	✓	N/A
Negative in character?	✓	✓	✗
Context-based (i.e., knowledge of stereotypical associations)	✗	✓	✓
Metalinguistic?	✓	✓	✗

Figure 2-5 Table of qualities shared or not shared between QI, MI and I-Implicature as per the Gricean-inspired perspective

2.4. Conclusion

I have summarised the nature of QI, MI and I-Implicature in the Gricean-inspired perspective and related perspectives, e.g., the Post-Gricean framework. I have detailed the implications and inconsistencies associated with the Gricean-inspired perspective's conceptualisations of the phenomena. The purpose of the aforementioned section is to lay the theoretical groundwork requisite for understanding the experimental work that comprises the bulk of this thesis. Given that Experiments 1 – 6 seek to elicit MI and prime MI with both MI and QI and I-Implicature, understanding the different perspectives of the similarities and differences between the phenomena is of utmost importance.

2.5. Probabilistic Accounts of Conversational Implicature

As a footnote to the summary of the Gricean-inspired perspective of conversational implicature, I will give an overview of the constituents of the Probabilistic perspective of conversational implicature.

⁶ These qualities are an uncontroversial interpretation of Levinson's (2000) perspective of I-Implicature, rather than my own.

While these accounts are inspired by the Gricean model, they seek to account for implicature derivation in probabilistic terms and do not centre the same ostensive-inferential mechanism in their account of conversational implicature and are therefore to be considered a distinct camp of perspectives. Nevertheless, the Probabilistic perspective does not seek to supplant, but rather complement, mechanistic accounts of implicature derivation, e.g., that of the Gricean-inspired perspective, by providing formalizations of Gricean reasoning that account for pragmatic reasoning that occurs on a more foundational level, i.e., below that of the adherence to conversational maxims (Franke & Jäger, 2016; Bergen et al., 2016). Such formalizations are valuable in their modelling of Gricean implicature, namely QI and, to a lesser degree, MI and I-Implicature (Rett, 2020; Wilson & Katsos, 2023).

The Probabilistic Account is considered to encompass the interconnected perspectives of game-theory approaches (e.g., Lewis, 196; Parikh, 2001; Franke, 2009; Blutner 2000; Bergen, et al., 2016, a.o.) Bayesian approaches (e.g., Lassiter and Goodman, 2015; Potts et al., 2015) and constraint-based approaches to pragmatic reasoning (e.g., Frank & Goodman, 2012; Goodman & Stuhlmüller, 2013; Degen & Tanenhaus, 2015, 2019) and such accounts can be defined by a set of common properties (Franke & Jäger, 2016). Such accounts are *probabilistic* in that they account for the inherent uncertainty associated with pragmatic communication by positing models thereof that represent subjective uncertainty as probabilistic distribution,

Moreover, they are *interactive*, as they represent the distribution of likelihoods (i.e., the probability of certain outcomes) through the explicit interaction between speaker, listener and context, rather than collapsing the representation of these constituents into one, e.g., as the Grammatical Account's exhaustivity operator does. Pragmatic communication is modelled as *rationalistic* and is therefore assumed to be optimally adapted in ensuring that pragmatic meanings are communicated, therefore the Probabilistic Account seeks to model rationalistic or optimal pragmatic reasoning. The Probabilistic Account is, to varying degrees, *Bayesian* in its approach. Broadly, pragmatic communication is modelled with a sensitivity to the listener's prior knowledge and the likelihood of the speaker's particular intention and a mechanism that updates the beliefs of the listener in response thereto. Such accounts are *computational* and *data-oriented*, and in the case of the former, accounts seek to offer a formalised, mathematical model of pragmatic reasoning which therefore makes empirical predictions about the outcomes of pragmatic communication which are ripe for experimental investigation, as per their data-oriented nature (Franke & Jäger, 2016).

The Rational Speech Act (RSA) framework is one such probabilistic model that formalises Gricean implicature derivation (Frank & Goodman, 2012; Goodman & Stuhlmüller, 2013). The RSA is a game-theory account of pragmatic reasoning whereby language use is considered a signally game (Lewis, 1969). The RSA framework models pragmatic reasoning as a recursive process involving a *literal listener*, modelled as interpreting an utterance literally (i.e., semantically), a *pragmatic speaker*, modelled to select an utterance based on the *literal listener*'s likely interpretation, therefore selecting for maximal communicative utility (i.e., a cost-benefit calculation of informativeness vs. utterance cost), and a *pragmatic listener*, who is modelled to reverse-engineer the *pragmatic speaker*'s intentions, taking into account their subjective beliefs about the *pragmatic speaker*'s communicative intentions with sensitivity to the conversational context (Degen, 2023). The RSA framework has been used to model scalar QI (Frank & Goodman, 2012; Peloquin & Frank, 2016), I-Implicature and MI (Bergen et al., 2016; Rett & Bumford, 2021). A benefit of the RSA framework as a model of implicature derivation is that the RSA models the *alternative* as contextually defined, rather than, say, lexically, i.e., as per Levinson (2000), and that the both the 'world knowledge' of the listener and a sensitivity to the Question Under Discussion comprise part of the RSA framework, therefore providing a clearer representation of the role of context in pragmatic reasoning than that provided by the Gricean-inspired perspective

2.5.1. The Question-Under-Discussion (QUD)

In short, QUDs are questions explicitly or implicitly addressed by interlocutors in a discourse, either by their answering, or the positing of questions that help answer the QUDs, e.g., according to von Steutterheim and Klein (1989) a narrative typically addresses the successive QUDs 'What happened at t_1 / t_2 / t_3 ' whereby t corresponds to chronological point (Benz & Jasinskaja, 2017).

The notion of the QUD plays a role in both the RSA and constraint-based probabilistic models of pragmatic reasoning (Degen & Tanenhaus, 2015, 2019; Degen, 2023) and is useful in understanding the nature of implicature derivation more generally (Wilson & Katsos, 2023). The derivation of scalar implicature is demonstrated to be sensitive to the QUDs (Zondervan, Meroni, & Gualmini, 2008; Degen & Goodman, 2014; Cummins & Rohde, 2015); given the competing QUDs 'Did he eat all of the cookies?' and 'Did he eat any of the cookies?' the scalar term *some* is more likely to be enriched to '+> some but not all', in the

former (Degen, 2023). As noted by Wilson and Katsos (2023), scalar implicature derivation is reliant on the relative degree to which *what is said*, e.g., ‘some’, and the semantically equivalent alternative, e.g., *all*, address the QUD. If the semantically equivalent alternative does not address the QUD, then suspending scalar enrichment will be licensed.

Conversely, as the nature of the alternative is different in the case of MI (i.e., not semantically equivalent, but instead similar), the implicatures derived are sensitive to the QUD in a different way. With MI, the primary implicature ‘+> not in the usual way’ is uniform between all MI, however, a variety of contextually-bound secondary implicatures are often available, take the following MI:

- i. He was thrashing rhythmically to the music.
 - +> He was dancing in an unusual way.
 - ++> He was dancing drunkenly.
 - ++> He was dancing badly.
 - [...]

Example 2-46

Here, we have two secondary MIs that could be elicited by Example 2-46’s utterance. The respective derivation of each can be viewed in terms of the QUD each answers, for instance, supposing in the example’s context a QUD is ‘Did Dad get too drunk last night?’ the MI ‘++> He was dancing drunkenly’ may be elicited, in contrast, the QUD ‘Did Dad embarrass himself’ might elicit the implicature ‘+> He was dancing badly’. As such, QUD can be used as a framework for understanding how certain conversational contexts license certain directions of MI enrichment, i.e., by licensing implicatures that address the QUDs to a greater extent. (Wilson & Katsos, 2023). Moreover, the QUD can be applied to I-Implicature, take the following:

- i. I’m going clubbing tomorrow.
 - +> I am going clubbing in the usual way.
 - ++> [...] therefore, drinking alcohol and ending up hungover and tired.
 - ++> [...] therefore, I will be with my raver friends.

[...]

Example 2-47

Supposing in the context of Example 2-47 a QUD were ‘Are you able to give me a lift to work on Saturday morning?’ the implicature that ‘clubbing’ stereotypically involves both alcohol consumption and a subsequent hangover is derived. In contrast, if a QUD were ‘Have you seen your friends recently?’ the implicature that ‘clubbing’ stereotypically involves a particular group of friends may be derived. As such, in a similar fashion to MI, QUDs can be used as a framework for understanding the particular secondary enrichments I-Implicatures may implicate. Moreover, like with scalar implicature, it appears as though QUDs can license or suspend I-Implicature:

- i. James took the gift and he was happy
 - +> James₁ took the gift and he₁ was happy
 - +> James₁ took the gift and he₂ was happy

Example 2-48

In a neutral context, local coreference is likely implicated by the utterance in Example 2-48; nevertheless, if a QUD were ‘Did James rebuff the birthday present from his estranged father?’, the I-Implicated local conference may be suspended to allow an interpretation of the utterance that better addresses the QUD.

2.5.2. Conclusion

To conclude, Probabilistic accounts of pragmatic reasoning, owing to their interactive, context-sensitive nature and compatibility with Gricean pragmatics, offer a useful perspective for understanding the nature of Gricean-inspired implicature derivation. Namely, Probabilistic accounts can explain aspects of pragmatic communication not explicitly accounted for by the Gricean-inspired perspective, e.g., how contextual information is integrated and updated in the derivational process. Similarly, the notion of the Question-Under-Discussion proves a useful analytical tool in accounting for similar gaps in the Gricean-inspired approach, e.g., how contextually-bound implicatures arise, how the

alternative is generated in the case of MI. As such, the relevance of these perspectives will be revisited in the experimental investigation of conversational implicature and the discussion of the results thereof.

3. Experimental Background

3.1. Experimental Investigation of QI

Pragmatic inference has been a central focus in the field of Experimental Pragmatics, particularly in the investigation of Gricean *conversational implicature*. Scalar implicature, a type of QI, has been the subject of widespread investigation, with research addressing various theoretical questions. These include whether scalar implicatures are generalized or particularized (e.g., Levinson, 2000; Hirschberg, 1991), whether they are derived by default (e.g., Chierchia, 2004; Levinson, 2000) or depending on context (e.g., Grice, 1975; Geurts, 2009; Russell, 2006), and whether their derivation occurs locally (e.g., Chierchia, 2004; Levinson, 2000; Carston, 1998) or globally (e.g., Grice, 1975; Russell, 2006; Geurts, 2009, 2010).

A growing body of empirical evidence contributes to these debates. For example, studies have explored the generalization of scalar implicature (e.g., Papafragou & Tantalou, 2004; Katsos & Bishop, 2011), the possibility of local enrichment (e.g., Geurts & Pouscoulous, 2009; Chemla & Spector, 2011), and the influence of QUDs on implicature derivation (e.g., Yang et al., 2018; Kursat & Degen, 2020; Ronai & Xiang, 2021). Other research examines the role of alternatives in implicature derivation (e.g., Breheny et al., 2006; Skordos & Papafragou, 2016; Bott & Frisson, 2022) and the extent to which Theory of Mind and Executive Function contribute to the process (e.g., Fairchild & Papafragou, 2021).

Experimental investigations into scalar implicature have utilized a wide variety of paradigms. For instance, mouse-tracking studies (e.g., Tomlinson et al., 2013) and eye-tracking studies (e.g., Politzer-Ahles & Husband, 2018), structural priming paradigms (e.g., Bott & Chemla, 2016; Rees & Bott, 2018; Marty et al., 2022), event-related potential (ERP) studies (e.g., Politzer-Ahles, 2020), and truth-value judgment tasks (e.g., Degen & Tanenhaus, 2011; Spychalska et al., 2016; Chemla & Singh, 2014; Khorsheed & Gotzner, 2023). Additionally, scalar QI research spans diverse populations, including adults (e.g., Bott & Frisson, 2022),

individuals with Autism Spectrum Disorder (e.g., Chevallier et al., 2010; Hochstein et al., 2018), children (e.g., Foppolo et al., 2012; Tieu et al., 2015; Stiller et al., 2015; Foppolo et al., 2021), and second-language learners (e.g., Khorsheed et al., 2022a, 2022b; Zhang & Wu, 2022).

Experimentation has provided important insights into how QI is processed. Evidence suggests that QI is derived with a delay and may rely on working memory resources (Chemla & Singh, 2014; Khorsheed & Gotzner, 2023). Moreover, data from developmental studies suggests that the trajectory of QI acquisition is consistent (Papafragou & Skordos, 2016). On the flip side, the evidence suggests that QI derivation varies between individuals (e.g., Antoniou et al., 2016; Fairchild & Papafragou, 2021) and between different types of scalar implicature (e.g., van Tiel et al., 2014; Bott & Chemla, 2016; Gotzner et al., 2018; Noveck et al. 2022) as do the associated processing costs (Chemla & Singh, 2014). Moreover, these observations appear to be effected by the type of experimental methodologies that are used (Khorsheed & Gotzner, 2023).

Conclusions from the experimental investigation of QI are often framed in support of Relevance Theory's account of QI derivation, owing to processing cost associated with QI that is predicted by relevance-theoretic accounts and observed experimentally (Noveck, 2018; Breheny, 2019) (Khorsheed & Gotzner, 2023). However, there exist perspectives that seek to challenge these conclusions (e.g., Grodner et al., 2010; Degen & Tanenhaus, 2015; Barbet & Thierry, 2018) (Khorsheed, et al., 2022a).

3.2. Experimental Investigation of MI

In contrast to the well-researched nature of QI, MI has received little experimental investigation; this is likely a result of the comparatively small degree of attention given to theoretical accounts of MI and the subsequently scant debate concerning MI's nature. Consequently, the degree to which MI exists 'in the wild' and the readiness with which such inferences are derived is, for the most part, unknown. That said, a small body of emergent experimental work investigates the nature of Manner-based inferences (Wilson & Katsos, 2023; Wilson, 2017).

As accounted for by Wilson and Katsos (2023), there exists some experimental investigation involving different types of MI. For instance, *negative strengthening* is investigated by

Gotzner et al., (2018b, 2018a). While utterances like ‘he is not a genius’ are said to I-implicate a negative strengthening, e.g., ‘+> He is stupid’ (e.g., Levinson, 2000) Gotzner et al., (2018b, 2018a) consider such utterances as occasions of Manner-based implicature, with this stance justified by the notion that utterances like ‘not a genius’ are more prolix than their alternatives, e.g., ‘stupid’. While this does not fulfil the usual criteria for Gricean-inspired MI, the study aims to investigate the interplay between Manner-based and quantity-based (i.e., ad-hoc scalar) inferences. Because negated scalar terms (e.g., ‘not a genius’ or ‘not all’) implicate a weaker scale mate (e.g., ‘intelligent’ or ‘smart’, and ‘some’, respectively) but negative strengthening excludes weaker scales mates to implicate a stronger term of an alternate scale (e.g., ‘not a genius’ may implicate the stronger term ‘stupid’ in the alternate scale < *slow, simple, stupid*>), the implicatures generated by Quantity- and Manner-based inference are in direct conflict.

Gotzner et al., (2018b, 2018a) observe that when confronted with utterances containing both scalar implicature and negative strengthening, participants were less likely to derive scalar implicature when they had derived the negative strengthening associated with the negated scalar term. While Gotzner et al., (2018b, 2018a) conclude that these findings demonstrate that different scalar terms can give rise to distinct patterns of inference, the dubious role Manner-based inference plays in the construction of Gotzner et al., (2018b, 2018a)’s paradigm limits the project’s contribution to the experimental investigation of MI – although, it should be noted that the experimental investigation of MI is not an express aim of the study. Wilson and Katsos (2023)’s overview outlines similar studies (i.e., Leffel, Cremers, Gotzner & Romoli, 2019; Tessler & Franke, 2019) that investigate manner-like, negation-based inferences, but conclude that such studies are complicated by questions concerning whether or not the inferences investigated are indeed examples of MI and the influence, or lack thereof, of other pragmatic factors like prosody.

Alongside the observation that manner-like, negation-based inferences associate with particular meanings, a small body of evidence exists in support of the suggestion that periphrastic causatives give rise to MI-like interpretations (Wilson & Katsos, 2023; Wilson, 2017). Song and Wolff (2005) and Wolff (2003) observe that in a picture-to-sentence matching task participants prefer to associate periphrastic causatives with unintentional actions and lexical causatives with intentional actions, both productively and comprehensively. Kline, Muentener and Schulz (2013) studied the memory recall of direct causal scenes when paired with either transitive sentences (e.g., ‘He stretched the toy’) or

periphrastic sentences (e.g., ‘He made the toy stretch’) finding that participants were better able to recall the results of the action (e.g., the length of the stretch) in the former and the manner with which that action occurred (e.g., the directionality of the stretching) in the latter. While Kline, Muentener and Schulz (2013) aimed to investigate a linguistic impact on memory recall, the study evidences a meaning-based distinction between periphrastic and non-periphrastic sentences, as does the data provided by Song and Wolff (2005) and Wolff (2003) (Wilson, 2017).

Similar evidence for such associations can be provided by insight from artificial language paradigms and the experimental investigation of Referential Pacts (Wilson, 2017). For instance, Bergen, Goodman, and Levy (2016) and Bergen et al., (2016) present studies in which participants take part in a communication game whereby they must use particular symbols (as a ‘non-linguistic’ proxy for language) to reference and interpret reference to different objects. The studies observe that participants consider the frequency of the object and *costliness* of communication (operationalised in monetary terms, i.e., by particular symbols being or less costly) when assigning reference. This is taken as evidence for participants engaging in a Manner-like reasoning process by assigning less costly reference to frequent objects and more costly reference to infrequent objects, mimicking the marked-to-marked language-to-meaning association of MI.

Similarly, the suggestion that speakers and listeners engage in a sensitivity to Manner-like conventions can be observed in the experimental investigation of Referential Pacts, the notion that interlocutors enter into implicitly agreed upon discourse-specific patterns of reference (Wilson, 2017), e.g., say, referring to Darwin College, University of Cambridge as ‘Darwin’. For instance, Barr and Keysar (2002) find that listeners rely on referential precedents when interpreting utterances, while Metzger and Brennan (2003) observe that the flouting of these pacts acts against listeners’ expectations, leading to a response delay.

Referential Pacts are observed to occur both on a partner-specific basis, particularly in interactive dialog (Brown-Schmidt, 2009), and on a generic basis, as per general expectations listeners may have about language use (Kronmüller & Barr, 2007). Evidence gleaned from the experimental investigation of Referential Pacts suggests that speakers and listeners do adhere to mutually assumed referential conventions, as predicted by the Gricean-inspired perspective, although the degree to which this approximates an adherence to a Manner-like

conversational maxim is unclear (Wilson, 2017). Two studies, Katsos and Wilson (2016) and Wilson (2017) target MI more explicitly.

3.2.1.1. Wilson & Katsos (2016)

Katsos and Wilson (2016) present a study investigating the availability of MI and I-Implicature, the former operationalised as both particularised MI (i.e., prolix-based MI) and generalised MI, i.e., *modals*, whereby the inclusion of a modal verb, e.g. in ‘John had the ability to solve it’, implicates that although possible, the action did not occur, *periphrastic causatives* and *litotes*. In the study’s Experiment 1, participants were provided with a conversational context and either an MI or I-Implicature eliciting utterance and tasked with matching the utterance to the most similar caption, either with a meaning congruent with an enriched meaning (i.e., an MI in the case of the MI-eliciting caption) or a conventional meaning compatible with an I-Implicated meaning. An example is illustrated below:

Conversational context:	Ending: MI Condition:	Ending: I-Implicature Condition:	M-Implicated Meaning	OR	I-Implicated Meaning
<i>Ben and Jude were acting in a play. Ben asked, "What happens at the end of the scene?" Jude answered:</i>	<i>"The villain leaves the room with a swish of his cloak and causes the door to close."</i>	<i>"The villain leaves the room with a swish of his cloak and closes the door."</i>	<i>The villain closes the door by swishing his cloak and making a draft</i>		<i>The villain leaves the room and closes the door, by pulling it shut.</i>

Figure 3-1: An example of the experimental stimuli used by Wilson & Katsos (2016) across conditions with both possible answers.

In the study’s follow-up experiments, participants were asked to rate the similarity between the M- and I-implicated meanings and the utterances in the MI and I-Implicature conditions respectively and were tasked with providing written feedback on the meaning of the utterances in the MI and I-Implicature conditions.

Across all measures of MI and I-Implicature derivation (i.e., binary choice, Likert scale-style rating and qualitative feedback), the derivation of MI in the MI condition was low (e.g., Experiment 1: 56%), while the derivation of I-Implicature was far higher (e.g., Experiment 1: 92%), with no significant difference between the derivation of the particularised MI and the generalised MI. Katsos and Wilson (2016) conclude that the experimental evidence for the frequency of Manner-based contrasts is not as robust as the Gricean-inspired perspective would predict, but acknowledge that the methodological constraints of operationalising MI experimentally (e.g., in this case, lacking any prosodic component) may preclude the observation of more ready MI derivation, yet take the observations as preliminary evidence to suggest that MI comprises part of adults' pragmatic repertoire. Moreover, while Katsos and Wilson (2016) do acknowledge the heterogeneity of MI, the operationalisation of MI is, necessarily, rudimentary (e.g., the MI subtype 'periphrastic causatives' includes only two examples thereof) which may prevent a more nuanced understanding of between-subtype derivational differences.

3.2.1.2. *Wilson (2017)*

Wilson (2017) builds upon the work of Katsos and Wilson (2016) and the limited prior investigation of MI from a developmental perspective (e.g., Antoniou, Grohmann, Kambanaros, & Katsos, 2013; Okanda, Asada, Moriguchi & Itakura 2015; Antoniou & Katsos, 2017) to present a set of picture-matching experiments investigating age-related differences (adults vs. children) in MI implicature derivation. In the study, participants were provided with either a MI-eliciting caption, e.g., 'The girl used a knife to make small pieces of apple' or a I-Implicature-eliciting caption, e.g., 'The girl cut the apple' and asked to match the caption to one of two images, one depicting a scene compatible with the meaning associated with the MI and another compatible with the I-Implicature (See Figure 3-2: The visual stimuli used in Wilson (2017) for the item associated with the caption: 'The girl cut the apple' for a visualisation). The captions were either GCIs, operationalised as MIs involving periphrastic causatives, or PCIs, operationalised as MIs involving prolixity.

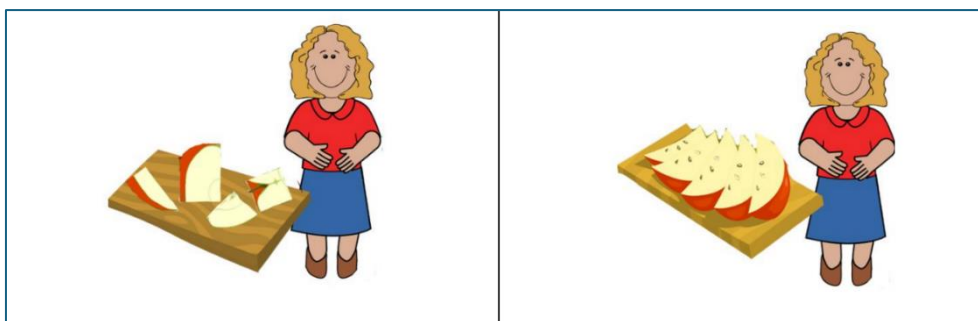


Figure 3-2: The visual stimuli used in Wilson (2017) for the item associated with the caption: ‘The girl cut the apple’

Wilson (2017) observed no difference in the rate at which adults and children matched the MI-eliciting captions to the MI-compatible scenes, with both showing higher rates of MI-compatible scene selection in trials with MI-eliciting captions. Moreover, Wilson (2017) reports no difference between the rates at which generalised and particularised MIs are derived by participants, although the child participants derived the generalised MIs at a higher rate than adults did, explained by Wilson (2017) as an adult preference for the selection of the typical image. Wilson (2017) provides corroborating evidence for Wilson and Katsos’ (2016) observation that adults interpret utterances with a sensitivity to Manner-based conventions and the observation that the distinction between GCIs and PCIs formulated by Levinson (2000) may not be as stark as it is considered theoretically.

3.3. Conclusion

Overall, studies concerning manner-like, negation-based inferences (i.e., Leffel, Cremers, Gotzner & Romoli, 2019; Tessler & Franke, 2019; Gotzner et al., 2018b, 2018a) and those concerning periphrastic causatives (i.e., Kline, Muentener & Schulz, 2013; Song & Wolff, 2005; Wolff, 2003) evidence the existence of a MI-like association between particular constructions and particular meanings, in keeping with the Neo-Gricean accounts of Quantity- and Manner-based inference (Levinson 2000; Horn, 1984, a.o.). Similarly, evidence provided from artificial language paradigms (e.g, Bergen et al., 2016; Bergen, Goodman, & Levy, 2016) and the experimental investigation of Referential Pacts (Brown-Schmidt, 2009; Kronmüller & Barr, 2007; Metzging & Brennan 2003; Barr & Keysar, 2002) provides evidence to suggest that manner-like reasoning guides utterance interpretation. While Katsos

and Wilson (2016) and Wilson (2017) provide evidence for the use of Manner-based contrasts as part of adults and children's pragmatic repertoires more explicitly, none of the aforementioned experimental investigation evidences the existence of the type of cognitive, inferential mechanism fundamental to the Gricean-inspired perspective. Consequently, while there exists a small body of evidence to suggest that MI exists superficially, the nature of the mechanisms governing MI derivation is unexplored. From this perspective, evidence for the existence of MI in the Gricean sense, i.e., as distinct type of pragmatic inference, remains weak.

3.4. The Current Study

As described in the preceding sections, there exists contention and ambiguity concerning the theoretical nature of MI and its relation to QI and I-Implicature. Principally, the extent to which MI and QI and I-Implicature share a derivational mechanism is unclear. Moreover, the nature and role of the *alternative*, *markedness* and contextual information in MI is uncertain, as is the scope of MI (i.e., does it account for 'secondary' inferences).

While some of these inconsistencies have been explored theoretically, a dearth of evidence exists to account for the existence of Manner-based inference empirically. Further to that, the small body of experimental investigation into MI that does exist (e.g., Wilson, 2017; Wilson & Katsos, 2016) only provides the phenomenological observation of Manner-based contrasts, rather than providing evidence for a derivational mechanism⁷ that governs Manner-based inferences.

To address the gap in empirical evidence for MI and its nature, the current study investigates MI, QI and I-Implicature using *structural priming* techniques, seeking to better understand 1) the extent to which Manner-based implicature can be elicited experimentally 2) the nature of the mechanism that governs MI derivation and 3) the extent to which MI, QI and I-

⁷ As described, the Gricean-inspired perspective frames conversational implicature in cognitive terms, broadly, as the result of a type of ostensive-inferential mechanism of linguistic cognition. Levinson (2000) and Horn (1989) provide structured accounts of how this mechanism might govern implicature derivation, e.g., by considering alternative utterances a speaker could have said, but did not and negating them. In this thesis the term 'mechanism' or 'derivational mechanism' is used semi-generically to reference the Gricean-like inferential process that derives conversational implicature, without a specific commitment to the structure, nature and scope of the mechanism, as such qualities are the subject of the investigation of this thesis. However, the assumptions of a cognitive mechanism of implicature derivation and conclusion from the current study are presented in 7.3.1.

Implicature derivation is governed by the same mechanism, as hypothesised by the Gricean-inspired perspective.

Before the exact methodological details of the current study are presented, I will present an overview of the experimental techniques and concepts used in the current study and present a detailed account of the prior studies that have inspired the specific methodological approach of the experiments comprising this thesis. In the broadest terms, the current study comprises a set of experiments that establish a baseline for MI and I-Implicature derivation (Experiments 1a and b), investigate a within-category MI-to-MI priming effect (Experiments 2a, 2b and 2c), an I-Implicature-to-MI between-category priming effect (Experiment 3) and a QI-to-MI between-category priming effect using two types of QI, *ad-hoc QI* and *adjectival scalars*. Finally, a ‘neutral’ non-pragmatic prime is tested to contextualise these results (Experiment 6). In order to understand the findings of these studies, a theoretical and methodological background to the paradigm is described.

3.4.1. Priming

In the broadest terms, *priming* is psychological phenomenon whereby the prior exposure to a particular stimulus, e.g., an image or a word, modulates a participant’s response to a successive stimulus, usually either influencing the direction of the response, or hastening a response time. See Pickering and Ferreira (2008) and Branigan and Pickering (2017) for overviews. Priming is a widely-used technique across the domains of psychology, neuroscience and linguistics (Weingarten et al., 2016; Bargh & Chartrand, 2000; Schacter & Buckner, 1998); for instance, Meyer and Schvaneveldt (1971) presents the widely cited observation that participants are quicker in recognising the word ‘doctor’ as a word after having seen the word ‘nurse’ than they are after having seen an unrelated word, a prototypical demonstration of linguistic priming, specifically *semantic priming*.

The paradigms used to elicit linguistic priming are diverse (Pickering & Ferreira, 2008; Tooley, 2023). From a cognitive perspective, priming tasks can be perceptual or conceptual (Ward, 2023; Roediger & Blaxton, 1987). In perceptual priming tasks, the same stimuli are presented in the priming stage as the target stage, sometimes in a degraded form (e.g., the word ‘doctor’ may appear before a ‘d___or’ target) and priming is evidenced by faster or more accurate recognition, i.e., perception (Ward, 2023). In contrast, conceptual priming tasks require the non-perceptual processing of a stimulus (e.g., the pragmatic processing of a

QI) and priming is measured as the faster or more accurate retrieval or re-application of the same mechanism or process (e.g., the derivation of QI in succeeding trials). Priming paradigms can investigate productive or comprehension-based priming effects (Tooley, 2023). In the former, priming is observed as the productive repetition of primed structures, while in the latter, priming is observed in comprehension, i.e., whether or not prior exposure to a particular structure modulates the subsequent comprehension of the same structure. Alongside more traditional production and comprehension-based tasks, like description elicitation or picture-to-sentence matching tasks, linguistic priming can be observed in eye-tracking (e.g., Marjanovic & Crepaldi, 2018; Boudewyn et al., 2014; Hoedemaker & Gordon, 2014), Event-Related-Potentials (e.g., Coderre et al., 2017; Franklin, et al. 2007) and fMRI data (e.g., Nakagawa et al., 2022).

The application of priming paradigms is observed across different levels of linguistic representation, with application to phonetic (e.g., Baxter et al., 2017; Dufour & Frauenfelder, 2016) and phonological levels (e.g., Zhang, 2023; Lim & Christianson, 2023; Sevald, Dell & Cole, 1995), lexical and semantic levels (e.g., Dufour & Meynadier, 2019; Snoeren et al., 2008), pragmatic levels (e.g., Rees et al., 2023; Rees & Bott, 2018; Bott & Chemla, 2016; Bott & Chemla, 2013) and conceptual domains (e.g., Ward, 2023; Tytus & Rundblad, 2016). Moreover, priming paradigms have found application across populations, for instance, in the areas of language acquisition (e.g. Van Zwet & Unsworth, 2024; Wada et al., 2020; Garrod & Clark, 1993; Huttenlocher et al., 2004), bilingualism (e.g., Hartsuiker & Bernolet, 2017; Schoonbaert et al. 2007; Loebell & Bock, 2003), language disorder (e.g., Keen & Lee, 2022; Chahboun et al., 2017; Ferreira et al., 2005; Hartsuiker & Kolk, 1998) and language and aging (e.g., Van Boxtel et al., 2023; Lee et al., 2022).

Consequently, linguistic priming is considered a well-evidenced phenomenon suitable for use across a wide range of experimental applications.

3.4.1.1. *Structural Priming:*

Structural priming (also referred to as *syntactic priming*) is a type of conceptual repetition priming whereby the use of a particular linguistic structure prompts its use in a subsequent experimental trial. Bock (1986) presents a seminal structural priming paradigm whereby participants were required to first repeat active or passive transitive sentences and double-object or prepositional-object dative sentences and then describe a picture. As per the

predictions of structural priming, participants were more likely to employ the syntactical form they were exposed to in the priming trial (e.g., a double object) in their subsequent description. Structural priming paradigms are widespread and well attested in their efficacy, both in productive and comprehensive priming paradigms (Mahowald et al., 2016; Tooley, 2023), demonstrating a general utility in understanding the nature of language (Branigan & Pickering, 2017). As mentioned, given that priming effects are observed across diverse levels of linguistic representation, it follows that structural priming techniques may prove useful in the investigation of pragmatic phenomena (Rees & Bott, 2017).

While limited in its breadth, there exists a small body of research that applies structural priming techniques to pragmatic phenomena. For instance, de Carvalho et al. (2016) employ priming techniques to investigate Horn scales. In the study, participants were presented with a lexical decision task in which scalars (e.g., ‘some’ or ‘all’) were preceded by stronger scale-mates (e.g., ‘all’ when ‘some’ is the target), weaker scale-mates (e.g., ‘some’ when ‘all’ is the target), identical words or unrelated non-words. de Carvalho et al. (2016) found that lexical decision reaction times were faster in the stronger prime condition than in the weaker prime and the non-word conditions and therefore concluded that the increased reaction times were indicative of the structural priming of QI, which itself attests to the psychological validity of Horn scales.

3.4.1.2. *Between-category Priming*

Given that priming effects can be observed across levels of linguistic representation, priming paradigms provide a useful tool for the investigation of the interconnectedness of these representational levels and the sharedness, or distinctness, of mechanisms underlying specific linguistic phenomena. Attempts to prime between phenomena with varying degrees of relatedness can be referred to as *between-category priming* and is a widely used methodology in linguistic priming paradigms.

For instance, semantic priming studies (e.g., McNamara & Holbrook, 2003; Neely, 1991; Meyer & Schvaneveldt, 1971) have evidenced the semantic interrelatedness of conceptually approximate lexical items, which suggests that superficially disparate lexical items enjoy a shared pre-lexical conceptual or semantic processing (Dyson et al., 2021). More specifically, the between-category priming of different semantic classes has been used to investigate differences in the memory-based organisation of concrete and abstract nouns (Ferré et al.,

2015). Moreover, bilingualism research made use of between-category priming to investigate the sharedness of levels of representation cross-linguistically (Hartsuiker & Bernolet, 2017), e.g., to evidence the sharedness of grammatical representations between different languages (Son, 2020; Kutasi et al., 2018) and to evidence language-independent information processing (Fleischer et al., 2012). Similarly, between-category priming studies have been used to investigate a shared grammatical representations between morphological and lexical items (Chen & Hartsuiker, 2021; Crepaldi et al., 2014). From a cognitive perspective, priming between linguistic and mathematical processing has been used to investigate the domain-generalty of cognitive processes (Zeng et al., 2021).

In terms of enrichment processing, a topic pertinent to this thesis, between-category priming has been used to investigate the sharedness of mechanisms of semantic and syntactic enrichment (e.g., Ito & Husband, 2022; Meyer & Feiman, 2021). Prior to this, Bott and Chemla (2016) employed techniques of structural between-category priming to investigate the sharedness of pragmatic mechanisms; Bott and Chemla attempted to prime four types of *enrichment-via-alternatives* (or ‘EVAs’, i.e., pragmatic enrichments that involve processing of an alternative), *plurals*, *numerals*, *ad-hoc QI* and QI involving scalar terms. Bott and Chemla (2016) observed between-category priming effects between numerals, ad-hoc QI and scalar QI, but not plurals, indicating that there exists a sharedness in the processing of these EVAs, but not plural enrichment. Moreover, Bott and Chemla’s (2016) study and those inspired by it (e.g., Meyer & Feiman, 2021) evidence the utility of between-category priming paradigms in investigating the extent to which derivational mechanisms are shared between different types of enrichment.

3.4.1.3. *What is priming?*

There exists contention concerning the mechanism that underpins priming, as described in Pickering and Ferreira’s (2008) review of structural priming. From one perspective, linguistic priming can be understood as a mechanism of dialogue alignment (e.g., Pickering & Garrod, 2004; Garod, Tosi & Pickering, 2018), whereby linguistic priming functions to facilitate communicative success by aligning the linguistic features of interlocutors involved in a dialogue. Another perspective considers linguistic priming to be a facilitator of linguistic fluency. Given that priming is oftentimes framed in terms of processing speed (e.g., Segaert et al., 2016; 2011; Corley & Scheepers, 2002, Smith & Wheeldon 2001), priming could be mechanism through which the cognitive load associated with the planning and encoding of linguistic structures is lightened (Cai & Zhao, 2024).

Alternately, priming can be considered a mechanism of implicit learning (e.g., Bock & Griffin, 2000), as evidenced by the longevity of structural priming effects (Hartsuiker et al., 2008) and the fact that less frequent structures give rise to stronger priming effects (Jaeger & Snider, 2013). In turn, this can be taken as evidence of error-based implicit learning (Cai & Zhao, 2024). Compatible with all of these perspectives, is the view that priming is the result of the residual activation of representations, for instance Pickering and Branigan's (1998) account considers a structural network comprising of combinatorial nodes that remain activated after they have been accessed (for instance, a node might represent a semantic concept or syntactical structure) such that their activation is expedited if they are re-accessed – see Branigan and Pickering (2017) for an overview. Bott and Chemla (2016), Rees and Bott (2017) and Meyer and Feiman (2021) extend the account of residual-activation to *mechanisms* and explain that, like representations, the use of particular cognitive mechanisms⁸ used in language comprehension, e.g., enrichment derivation, can be primed by the prior activation of the same mechanism.

Whether or not priming is an effect of explicit or implicit memory is debated, while some evidence suggests that the explicit memorisation of a particular structure (e.g., Zhang et al., 2020) can augment the priming effect it may have, most priming studies do not engage explicit memorisation and priming is shown to occur in circumstances in which explicit memory cannot be readily engaged (e.g., Ferreira et al., 2008; Heyselaar et al., 2021). This may suggest that priming is not driven by explicit memory, but may be helped by it (Chang et al., 2012) (Cai & Zhao, 2024). While activation-based accounts suggest that priming is an effect of the residual activation of a particular structure, the longevity of certain priming effects (see Pickering & Branigan, 1998) and the theoretical improbability of extended residual activation suggests that structural priming does function as the result of memory, at least in part.

3.4.2. Conclusion

As demonstrated, linguistic priming, and more specifically structural priming, is well-evidenced across a variety of linguistic phenomena and range of participant populations⁹. As

⁸ The nature of the mechanism primed in these studies is discussed in more detail in 7.2 of the thesis.

⁹ It should be acknowledged that the use of psychological priming techniques has been met with a degree of incredulity in recent years but noted that the cynicism with which priming is met is the result of the replicability crisis concerning social and behavioral priming, not structural priming (see Kahneman, 2012).

explained, given that structural priming occurs across and between levels of representation, it presents a useful tool for the experimental investigation of MI and is particularly useful for the investigation of the extent to which phenomena are related, e.g., MI, QI and I-Implicature. As mentioned in, the set of experiments presented in this thesis are inspired by a small body of pre-existing experimental investigation into conversational implicature that uses structural priming techniques. Before presenting the current study, I will provide brief synopsis of this body of research and the methodologies used to elicit priming.

3.4.3. Methodological Background

3.4.3.1. Bott & Chemla (2016)

Bott and Chemla (2016) employ a structural priming paradigm to investigate the extent to which there exists a shared derivational mechanism between QI-like phenomena. Bott and Chemla investigate *some* QI (i.e. ‘+> some but not all’), *ad-hoc* QI (e.g., ‘+> there is one cat’ from ‘There is a cat!’), *numeral* enrichment (e.g., ‘+> I have exactly four sweets’ from ‘I have four sweets’) and *plural* enrichment (e.g., ‘+> Jane has more than one child’ from ‘Jane has children’), based on the observation that all of the phenomena can be considered enrichment-via-alternatives, but that debate exists concerning the degree to the phenomena are derived by the enrichment-deriving mechanism.

While lexical quantifier scales (e.g., *some* QI) and numerals are both considered EVAs by some accounts (e.g., Chierchia, 2004; Horn, 1989; Gazdar, 1979), other perspectives (e.g., Breheny, 2008; Geurts, 2006) account for numerals as exact numbers, whereby numerical exactness is a product of semantics, rather than pragmatic implicature. Similarly, while ad-hoc QI is considered to involve the negation of an alternative (at least, as per the Gricean-inspired perspective), the nature of the alternative is different to that of lexical scales, as it is contextually defined. Similarly, while plurals are considered to exemplify some qualities of EVAs (e.g., Spector, 2007), they are not traditionally considered instances of QI, unlike ad-hoc scalars, lexical scalars, and numerals, by some perspectives (Bott & Chemla, 2016).

Bott & Chemla (2016) employ a structural priming paradigm based on Huang, Spelke and Snedeker’s (2013) ‘hidden box’ paradigm, whereby participants are asked to choose between an existent, visible option, and a hidden, speculative answer (i.e., the hidden box) in the

critical trial (see Figure 3-3). In Bott and Chemla’s paradigm, the ‘hidden box’ critical trials were preceded by two prime trials. In all trials, participants were presented with two cards and an EVA-containing caption (e.g., ‘some of the symbols are clubs’) and were asked to match the caption to the best representation. The prime trials were presented in either a *weak* or *strong* condition that were designed to either force or prevent the pragmatic enrichment of the caption. In the strong condition, a card compatible with an enriched understanding of the caption was presented (e.g., a card in which some, but not all, of the symbols were of the referenced shape) alongside a card that represented the unenriched interpretation of the caption compatible with a semantic, but not pragmatic, understanding (see Figure 3-3).

In contrast in the weak prime condition, participants were presented with a card that represented the unenriched interpretation of the caption and a false image, such that they were forced to select the card compatible with the unenriched interpretation of the caption. In the critical trials, participants could either select the card compatible with the unenriched caption, or select a ‘Better Picture?’ card, which suggested that they perceived there to be an image that better matched the caption (i.e., one that represented the pragmatically enriched interpretation of the caption). In sum, the *strong* primes involved pragmatic enrichment, while the *weak* primes did not, and the selection of the ‘Better Picture?’ card in either was taken to indicate QI derivation.

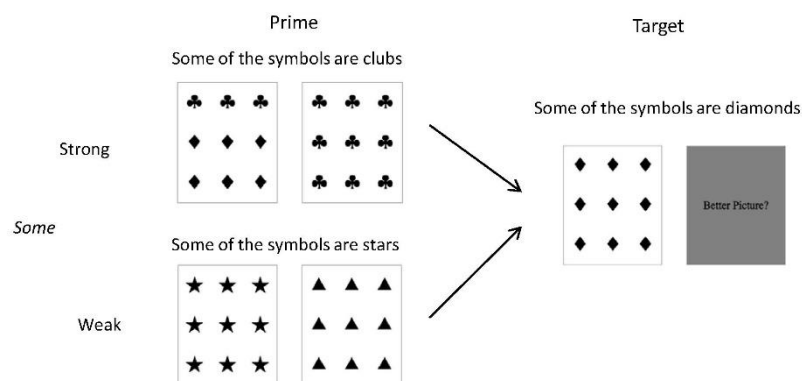


Figure 3-3: The ‘strong’ and ‘weak’ trials and the ‘Better Picture?’ target trial used in Bott & Chemla (2016)

The study aimed to investigate the extent to which there exists a derivational mechanism for EVAs that functions domain-generically, i.e., is responsible for the derivation of different EVA subtypes. Consequently, the prime trials involved all four EVA conditions (*some* and ad-hoc QI and numeral and plural enrichment) and were fully crossed with critical trials

involving the same four conditions. Therefore, a permutation of each EVA prime and EVA critical trial was achieved, both within-category (e.g, ad-hoc → ad-hoc) and between-category (e.g., ad-hoc → numeral). Bott and Chemla (2016) predicted that if there exists an EVA-specific derivational mechanism, then the derivation of a previous EVA should influence the subsequent derivation of another type. The *weak* prime trials aimed to reduce activation of the EVA-specific mechanism, while the *strong* trials served to increase activation of the mechanism and prime the derivation of subsequent EVAs.

Bott and Chemla (2016) present three experiments: Experiment 1 investigates the within-category and between-category priming of *some* and ad-hoc QI and numeral enrichment, Experiment 2 investigated the within-category priming of different numeral enrichments (i.e., ‘six’ and ‘four’), while Experiment 3 addressed the between- and cross-category priming of *some* QI and numeral and plural enrichment. Between the studies, Bott and Chemla report a strong within-category priming effect for all EVA subtypes, with the selection of the ‘Better Picture?’ card higher after the *strong* priming trials than after the *weak* priming trials. Between-category priming was observed between *some* QI and ad-hoc QI and numeral enrichment. However, plural enrichment was not primed significantly by any of the non-plural primes. In the between-category effects, Bott and Chemla observed no effect of prime direction (e.g., ad-hoc → numeral vs. numeral → ad-hoc) and no interaction between EVA types (i.e., no one EVA prime to EVA target combination conferred a priming advantage). The priming effect was stronger in within-category condition than in the between-category condition. Moreover, the observed priming effect was stronger in earlier trials than later trials, suggesting that the priming effect’s potency is related to its contextual novelty, in some way.

Bott and Chemla (2016)’s account of the findings is activation-based, and they interpret the priming effect to be the residual activation of a shared EVA mechanism that is likely shared between EVA subtypes (or, at least, between *some* and *ad-hoc* QI and numeral enrichment) but is not shared by plural enrichment. Bott and Chemla (2016) offer a two-fold explanation for the locus of priming effect, suggesting that the mechanism responsible for EVA enrichment could be primed in its entirety – explaining the between-category priming effect – and that there could be the added benefit of an alternative-specific priming effect, whereby the specific alternative requisite for the particular EVA is primed – explaining the within-category priming effect – see Waldon and Degen (2020) for a similar claim.

Bott and Chemla (2016) demonstrates a seminal application of a structural priming paradigm to QI and provides some important insights for the nature of QI. Firstly, that QI can be primed in experimental settings. Secondly, that QI subtypes can be primed by other QI subtypes. Thirdly, that priming can be used as a methodology for discerning the degree to which mechanisms of implicature derivation are shared between different types of pragmatic enrichment. Bott and Chemla's (2016) findings are compatible with the Gricean-inspired account, Post-Gricean accounts and Grammatical accounts of QI derivation, given that across all three, *some* and ad-hoc QI and numerals are considered to be enriched in the same way, albeit by drastically different mechanisms.

3.4.3.2. Rees & Bott (2018)

Rees and Bott (2018) investigate the extent to which the priming effect observed in Bott and Chemla (2016) is a product of the residual activation of QI's derivational mechanism, or if it is the result of the increased salience of the alternative itself. In investigating this question, Rees and Bott (2018) seek to determine whether the presence of a salient alternative leads to implicature priming automatically, or whether the search for an alternative is a discrete step in the derivation of QI and whether implicature can be primed independently, i.e., via a non-specific usage mechanism – as predicted by the *salience model* and the *combination model* of implicature derivation, respectively. The competing models of implicature derivation are presented below in Figure 3-5. Broadly, the study seeks to better understand the locus of the priming effect observed in Bott and Chemla (2016).

To investigate these questions, an adaptation of Bott and Chemla's paradigm is used. The paradigm utilises a *priming-via-the-alternative* condition, whereby EVA alternatives (e.g., 'all' to 'some') are used to prime the derivation of their respective enrichment. Here, Rees and Bott's *priming-via-the-alternative* condition does not involve the activation of the usage mechanism as no QI is derived. Instead, the prime trials independently promote the salience of the EVA alternative. In this condition, participants are presented with an EVA-containing caption, e.g., 'All of the symbols are crosses', and two images, one representative of the EVA alternative (e.g., a configuration representative of 'all') and a false image (i.e., one that does not represent the caption) – see Figure 3-4 for an illustration. Alongside the priming-via-alternative condition, the *strong* and *weak* prime conditions presented in Bott and Chemla (2016) were reused, as were the same critical trials.

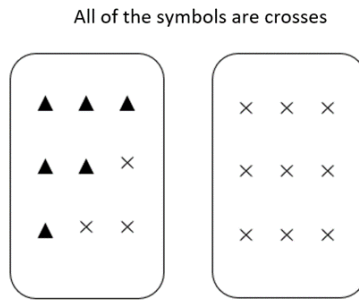


Figure 3-4: The priming-via-the-alternative prime trial in the 'some QI' condition from Rees & Bott (2018)

Rees and Bott's (2018) findings replicated those of Bott and Chemla (2016). The study found an increase rate of 'Better Picture?' selection, i.e., QI derivation, in the *strong* prime condition, compared to the pseudo-baseline – the rate of QI derivation in the *weak* condition. Moreover, Rees and Bott observe that the same effect is observed in the *priming-via-the-alternative* condition and that this increase is not significantly different to that of the *strong* primes. Rees and Bott (2018) therefore conclude that the observed priming effect is the result of the alternative having been primed, rather than the usage mechanism more broadly. This conclusion is based on the observation that the *priming-via-the-alternative* condition does not involve the QI-specific usage mechanism, but that the *strong* prime condition does involve the alternative. Rees and Bott (2018) attribute this finding to two possible causes. On one hand, the *strong* and *priming-via-the-alternative* conditions may raise levels of alternative activation and the alternative may remain activated in subsequent critical trials. On the other hand, the increased importance of context and informativity required by the *strong* and *priming-via-alternative prime trials* may modulate the participants' degree of contextual awareness in the subsequent critical trial, while the *weak* trials may cause them to suspend it.

Rees and Bott (2018) conclude with two main findings. Firstly, that interlocutors are sensitive to the salience of alternatives when deriving QI. Secondly, that there is no evidence to suggest that the mechanism responsible for QI comprises of discreet steps whereby the salience of the alternative can occur without the subsequent derivation of the QI. As such, Rees and Bott conclude that the findings of the study are better accommodated by a *salience* model of implicature derivation, whereby sufficiently activated alternatives give rise to automatic implicature derivation. However, the explanation of the priming effect in Rees and Bott (2018) is somewhat inconclusive, and it is unclear whether identifying a salient alternative involves the activation of the alternative-specific mechanism (i.e., one responsible for the search for an alternative) or whether the lexical representation of the alternative (e.g., ‘all’) is activated by the prime trials.

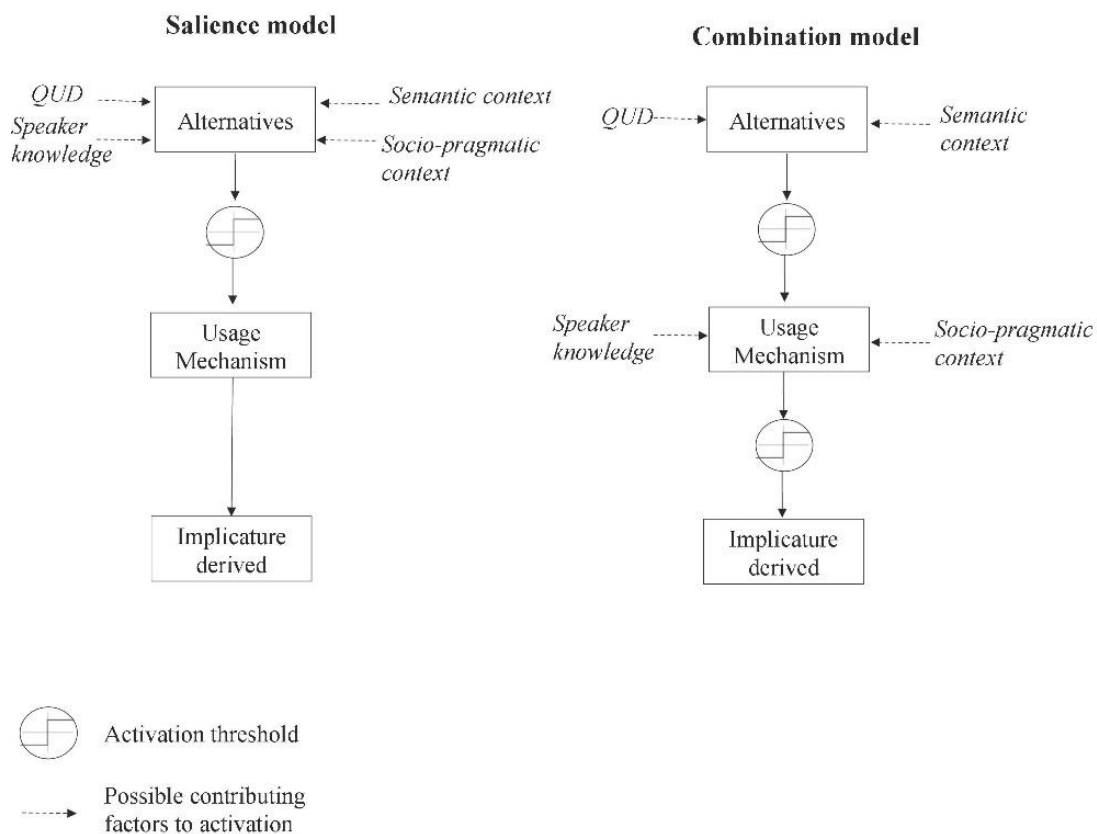


Figure 3-5: An illustration of the salience and combination model taken from Rees & Bott (2018) licensed under CC BY-NC-ND.

3.4.3.3. Marty et al 2022 & 2021

Marty et al., (2021) point out that Bott and Chemla (2016) and Rees and Bott (2018) both use a pseudo-baseline, i.e., the rate of QI in after the weak primes, rather than a true baseline, i.e., the rate of QI after no primes. Consequently, Marty et al., (2021) argue that the pseudo-

baseline may be lower than a true baseline due to an *inverse-preference* effect or ‘reverse-priming’ effect, whereby the usual *weak* interpretations prime the non-derivation of QI in the critical trials. Therefore, the true effect of the *strong* and the *priming-via-the-alternative* conditions may be obscured.

Marty, et al. (2021) presents a re-run the priming paradigm used by Bott and Chemla (2016) and Rees and Bott (2018) using an unprimed baseline condition. Marty et al. find that when the *weak* and the *strong* prime conditions are compared to the unprimed baseline, an increase in QI derivation can only be seen in the ad-hoc QI condition. In contrast, the *strong* primes in the *some* QI and numeral enrichment conditions do not increase the rate of enrichment in subsequent critical trials. Marty et al. conclude that this likely because of the higher rate of baseline rates of *some* QI and numeral enrichment derivation (~ 70%). In contrast, the baseline rate for ad-hoc QI is far lower (~ 10%). This suggests that a ceiling-effect may be responsible for the *some* QI and numeral enrichments resistance to priming.

On the flip side, the *weak* priming condition lowered rates of QI derivation below baseline level across all three QI subtypes. In Marty et. al’s *priming-via-the-alternative* condition, a priming effect for ad-hoc QI is observed, but not for other implicature subtypes. Marty et. al (2022) retests the *strong* and *weak* prime conditions and replicate those of Marty et al. (2021). In sum, the *strong* primes increased the rate of QI derivation in the ad-hoc QI condition, but not in the *some* QI or numeral enrichment conditions, whereas the *weak* primes lowered enrichment rates across all three.

Departing from the usage-based or activation-based accounts discussed above, Marty et. al (2022) consider the results from perspective of *context adaptation*. From an adaption-based perspective, priming is the result of a probabilistic, adaptive mechanism of implicit learning whereby fast-cycling, reactive and flexible predictions about words or constructions are made by interlocutors based on their prior exposure to them in a conversational context – see Fine et al. (2013) and Jaeger and Snider (2013) for similar accounts. Given that QIs are defeasible and cancellable depending on conversational context, Marty et al. (2022) reason that the disparate rates of pragmatic enrichment in the baseline condition are indicative of decontextualised assumptions about the likelihood of a QI subtype being warranted, e.g., ad-hoc QI is less likely than *some* QI, which is less likely than numeral enrichment).

Marty et al. (2022) considers the within-participant presentation of both the *strong* and *weak* primes to require a dynamic, updating consideration of contextual factors. Consequently,

optimal informativeness is utilised in the *strong* trials, but is disregarded in the *weak* trials in response to the changing context. In turn, this gives rise to quick-cycling contextual adaptations. On the other hand, Marty et al. (2022) argue that the relatively static linguistic context established in the baseline trials results in a less dynamic consideration of current contextual factors and prior expectations of implicature probability and distribution are accessed instead.

Marty et al.'s (2022) account considers the *strong* ad-hoc QI primes to elicit an inverse-preference effect, such that the infrequent and unexpected enriched ad-hoc QI primes require a more radical adjustment from baseline. In turn, the mechanism of *context adaptation* is more sensitive to them and ascribes them more probabilistic weight, such that they are more readily enriched in the subsequent critical trials. In contrast, the mechanism is less sensitive to expected interpretations (i.e., enriched *some* QI or numerals) and therefore they do not lead to a significant priming effect.

3.4.4. Conclusion

Evidently, Bott and Chemla's (2016) paradigm is useful in eliciting within-category and between-category priming effects for subtypes of QI. Consequently, it provides a useful blueprint for the development of other experimental paradigms that seek to use structural priming as a tool for the investigation of other types of conversational implicature. Therefore, a variation of Bott and Chemla's (2016) paradigm will be used to prime MI using other MI, I-Implicature and different types of QI in the current study.

However, the nature of priming effect elicited by the paradigm is unclear. Bott and Chemla (2016), Rees and Bott (2018) and Marty et al. (2021) favour explanations involving the residual activation of either the usage mechanism or the alternative. In contrast, Marty et al. (2022) suggest that the observed priming effect is rather the result of a mechanism of *context adaptation*. While Bott and Chemla (2016), Rees and Bott (2018) and Marty et al.'s (2021) perspectives assume that QI is computed as an EVA (i.e., in the Gricean sense), Marty et al.'s (2022) perspective of priming is more ambivalent to this question. While *context adaptation* may prompt interlocutors to derive more QIs in the critical trials, the priming effect does not directly target the mechanism responsible for implicature derivation.

Whether the priming effect observed between Bott and Chemla (2016), Rees and Bott (2018) and Marty et al. (2021; 2022) is the result of the residual activation of QI's derivational mechanism, a part of it, or the result of *context adaptation*, the paradigm is useful in

discerning the boundaries between theoretically related phenomena. Nevertheless, the nature and extent of this relatedness hinges on the nature of the priming effect observed. For instance, if priming occurs at the level of the alternative, this may suggest that QI subtypes share alternative-deriving mechanisms or alternative representations. In contrast, if priming occurs at the level of the usage mechanism, then fewer levels of sub-computation may be shared¹⁰. In contrast, an account of *context adaptation* does not suggest that different QI subtypes involve any shared, domain-specific derivational mechanisms.

Consequently, the application of this paradigm to MI, QI and I-Implicature is in itself experimental and findings from the between-category priming of less theoretically adjacent phenomena have the capacity to contribute to both our theoretical understanding of the extent to which phenomena share a derivational mechanism and contribute to our understanding of the nature of the priming effect between conversational implicatures.

4. Priming MI with MI

4.1. Overview of Experiments

The following section will present the series of experiments that aim to investigate whether not MI can be primed by other MI and what factors may modulate the strength of any priming effect. To investigate this question, I used the same priming paradigm as Bott and Chemla (2016) as part of an online study testing native English-speaking adults. In Section 3.44, I present an overview of the construction of the paradigm and the two pre-experimental item screening tasks.

Experiments 1a and 1b establish a baseline for MI and I-Implicature respectively, to establish whether the critical trials I have constructed elicit different response patterns depending on whether or not prolix captions are used, i.e., whether they elicit MI. In Sections 5 and 6, I present three experiments. Experiment 2a presents a first attempt at priming MI with other

¹⁰ Meyer and Feiman (2021) uses a variation of Bott & Chemla's (2018) paradigm to investigate the extent to which different levels of sub-computation are shared between different EVA types, concluding that priming can occur at different levels of shared sub-computation, which suggests that Rees & Bott's (2018) conclusion may not capture the whole picture of the priming effect between the EVA subtypes.

MI, while Experiment 2b and Experiment 2c investigate whether the inclusion of a pre-trial training block and the inclusion of guided primes augments any priming effect.

4.2. Pre-Experimental Item Screening

The experimental design of the current study was based on the structure of the picture-matching task presented in Bott and Chemla's (2016) original paradigm, with the study's critical trials that included a caption containing a possible enrichment (i.e., MI) and two images, one corresponding to a pragmatically enriched interpretation of the caption and another corresponding with a semantic, unenriched interpretation of the caption. Participants were tasked with matching one of the images to the caption by answering the question 'What image can Alan [the experimental character] see?' The experiment was constructed and hosted on Gorilla Experiment Builder (www.gorilla.sc) an online experiment building platform, and participants interacted with the experiment using a mouse, trackpad, or touchscreen (i.e., clicking).

For Experiments 1a and 1b, the paradigm involved critical trials that elicited MI and I-Implicature, respectively. MI was operationalised as prolixity-based MI, whereby the use of a prolix utterance M-Implicates a marked referent, while markedness was operationalised as an unconventionality or 'weirdness'. On the flip side, I-Implicature was operationalised as implicature-to-stereotype (See 2.3.5), whereby the use of a conventional term (e.g., 't-shirt') indicates a stereotypical referent. The paradigm involved MIs and I-Implicatures that referred to items of clothing, such that the MI trials aimed to associate MI-containing captions with images of *marked* articles of clothing, while I-Implicature trials aimed to associate I-Implicature containing captions with *unmarked* articles of clothing. The critical trial for the experimental item 't-shirt' in the MI condition is illustrated below:

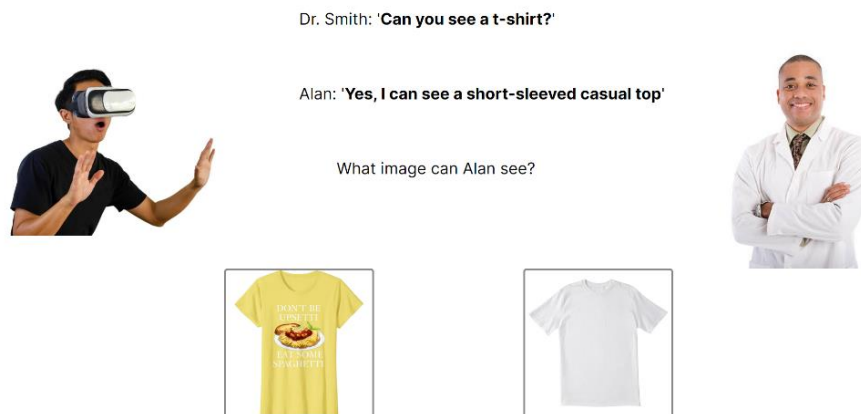


Figure 4-1: MI critical trial for the item 't-shirt'

4.2.1. Images

To adapt Bott and Chemla's (2016) paradigm to MI and I-Implicature, I compiled a stock of marked-unmarked image pairs that corresponded to a single lexical item. I chose to only use images depicting articles of clothing or accessories (e.g., 'sunglasses'). My reasoning was two-fold.

Firstly, picking items from the same semantic field allows for a uniform sense of item markedness, as related items are more likely to be marked in a similar manner. In contrast, disparate items (say, 'hoover' and 'cat') can share fewer manifestations of item markedness. This decision was motivated by a desire to maintain the level of inter-trial similarity present in Bott and Chemla's (2016) trials. Moreover, MI and I-Implicature can be considered to have primary and secondary extensions, i.e., as denoted by Levinson (2000) with the symbols '+>' and '++>', respectively.

While the primary implicature can be considered uniform between all instantiations of an implicature (e.g., all instances of I-Implicature implicature-to-stereotype implicate '+> [...] in the stereotypical way'), secondary implicatures (e.g., the I-Implicature '++> with a fork' from the utterance 'I ate the spaghetti') can diverge quite drastically from one another. Moreover, it is unclear whether or not primary and secondary implicature are the product of the same mechanism. To ensure the highest chance of successfully priming of MI and I-Implicature, I decided to create trials with as uniform a primary and secondary implicature as possible.

Secondly, articles of clothing are highly familiar to the general population, such that establishing *marked* and *unmarked* qualities is easier. In contrast, less commonly encountered objects (say, ‘hovercraft’ or ‘angler fish’) have intuitively fewer salient stereotypical associations and therefore are difficult to find marked and unmarked examples of.

I constructed the lexical item list intuitively, with the guiding notion that the items should correspond to unmarked, familiar articles of clothing (e.g., avoiding items like ‘bustle’). This process resulted in the following list of 42 items:

Apron	Bra	Gloves	Jeans	Shirt	T-shirt
Bathing Suit	Cap	Goggles	Necklace	Skirt	Tuxedo
Bathrobe	Cowboy Boot	Handbag	Polo Shirt	Socks	Vest
Beanie	Dress	Hat	Puffer Coat	Sunglasses	Wallet
Belt	Earring	Headband	Purse	Swimming Trunks	Wedding Dress
Bowtie	Flip Flops	Headphones	Ring	Tie	Wetsuit
Boxer Briefs	Glasses	High Heels	Scarf	Tights	Yoga Pants

I used Google Images to collect the item’s image pairs. For the unmarked images, I googled ‘[item]’ and for the marked images googling variations of ‘weird [item]’ and ‘ugly [item]’. This approach allowed me to gather a variety of appropriate images and I used my intuition to select images that appeared sufficiently marked and sufficiently unmarked and also had a sufficient degree of contrast with their marked or unmarked counterpart.

Qualitatively, the typicality of the unmarked images manifested as a plainness or ordinariness, while the marked objects involved shapes, colours, materials, patterns and textures atypical of their item class. The images were edited to have a plain white background and the images were resized to ensure neither the marked nor the unmarked image was noticeably larger. A full list of the 84 images can be found in 9.1, while the image pairs ‘bra’, ‘bathing suit’, ‘gloves’ and ‘dress’ are depicted below:









Item:	'bra'	'bathing suit'	'gloves'	'dress'
Marked				
Unmarked				

Figure 4-2: marked and unmarked image pairs for different MI critical trials.

4.2.2. Item Screening Task 1

Item Screening Task 1 task aimed to assess whether both the marked and unmarked items were identified with the labels I associated with them (e.g., if the marked 'bra' image was considered to depict a bra).

Establishing whether the images were considered to correspond to the labels was important given the alternative's role in the generation of MI. When a prolix description is used (e.g., in the MI-condition critical trials) its ability to elicit a MI relies on a salient, unmarked, alternative way of speaking, i.e., I-Implicature. For instance, 'dress' is the unmarked alternative to a prolix description thereof. If a marked image is not associated with a particular, unmarked item label (e.g., 'dress'), then there is no salient alternative with which to derive a MI. Moreover, if the marked and unmarked images pairs do not correspond to the same label, they are not in competition referentially. Consequently, a participant could answer the question 'What image can Alan see?' semantically rather than pragmatically, if only one of the images is thought to correspond to referenced item.

Similarly, if the marked images are not readily identified with the item label, the critical trials may inadvertently trigger an ad-hoc QI in which a nonce Horn scale, e.g., < bra*, bra >, is used. Here, 'bra*' represents bra-like objects and is entailed by the use of the word 'bra' (i.e., to be a 'bra' is to be bra-like). If processed as a QI, the circumvention of the lexical item 'bra' may imply '+>bra*but not a bra' (c.f., '+> some but not all'). That is, imply 'objects that are bra-like but not bras' and may give rise to selections of the marked image in the critical trials. Nevertheless, if the depicted marked items are independently associated with the item labels,

this provides assurance that participants are not selecting the marked image as the referents of the prolix captions due to QI, i.e., because they believe them to be ¬*item label*.

4.2.2.1. *Participants*

Participants (N = 74) were recruited from Prolific (www.prolific.com) [2024] an online recruitment platform for paid experimental participants and were selected on the basis that they were adults (≥ 18 years old) and self-identified as native English speakers.

4.2.2.2. *Procedure*

The screening procedure involved a small-scale online survey in which the 84 images were depicted next to a textbox and participants tasked with labelling the item in the images. The task was between-participants, such that a given participant was either randomly assigned to provide a response to all 42 of the marked images or all 42 of the unmarked images; 36 participants were assigned to the marked condition and 38 to the unmarked condition. The images were presented separately to avoid encouraging language that sought to contrast the marked and unmarked images (e.g., ‘the normal [item]’) and to avoid contrastive labelling when the images corresponding to the same item were displayed, e.g., ‘hat’ for the unmarked *hat* image, and a description or a hyponym for the marked image.

After an instruction screen, participants were presented with all 42 trials, without fillers, in a randomised order. In the pre-trial screen, participants were given the following instructions:

“In this task, you will be asked to label the object that appears in an image. You will see an image on the left and an empty text box on the right. You will be tasked with giving the object a label. Try not to think too hard about your answer - the shorter the better!”

In the critical trials, participants were presented with an image and the question ‘What object is shown in the image?’ and their responses were recorded in a text box. The trial for the marked ‘hat’ image is depicted below:

What object is shown in the image?

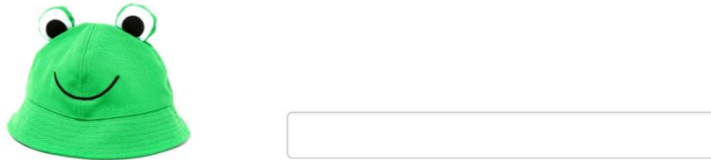


Figure 4-3: trial screen for the item 'hat' in Item Screening Task 1

4.2.2.3. *Data & Conclusion*

For all images, over 70% of the descriptions mentioned the item label or a synonym thereof, which was used as the threshold for continuing to use the images in subsequent experimental trials. While a threshold of 70% may appear low, this allows for the inclusion of images that were described using features of the depicted item, rather than the item label, for instance the marked 'bathing suit' image was variously labelled as an '*avocado*' and an '*avocado seed*' such that 73% of responses labelled the object as a 'bathing suit'.

Moreover, as the images will be presented alongside a caption containing either the item label or a description of the item, identifying the items will be easier in the subsequent experimental trials than in the pre-trial screening task. Nevertheless, the average successful label rating was above 90% for all images, marked and unmarked. However, the marked image descriptions were longer and contained more adjectives, e.g., 'weird', 'funny' and 'ugly'.

In conclusion, the screening task revealed that the items depicted in the marked and unmarked images were associated by the experimental participants with the item labels I had ascribed to them and were appropriate stimuli for use in the succeeding experiments.

4.2.3. Item Screening Task 2

The second screening task involved rating the markedness of the articles of clothing depicted in the marked and unmarked images in relation to their assigned item label (i.e., answering the question ‘How typical is this [item label]?’).

The task was designed to provide an empirical measure of each item’s markedness and to assess the relative degree of contrast between the markedness of the marked and unmarked pairs. Establishing this is important for two reasons. Firstly, considering the images were chosen based on intuition, it is essential to understand whether or not they are considered by participants to be particularly marked or particularly stereotypical. Secondly, given that MI relies on an implicit contrast between a marked referent and its stereotypical alternative, having an empirical measure of the markedness contrast between the paired images allows for me to investigate whether the extent of this contrast influences the availability of MI.

4.2.3.1. Participants

Participants (N = 38) were recruited from Prolific (www.prolific.com) [2024] and were selected on the basis that they were adults (≥ 18 years old) and self-identified as native English speakers.

4.2.3.2. Procedure

The item rating task was conducted within participants, whereby each participant provided a markedness rating for all 84 images, both marked and unmarked. The marked and unmarked images were presented together in order to generate a sense of contrast between them. Given that markedness is contextually defined (Rett, 2020; Wilson, 2017), if presented with a batch of only marked images, for example, participants may rate the less-marked marked images as ‘unmarked’ compared to the preceding images and vice versa.

The task consisted of 84 trials, presented without fillers and in a randomised order, which was presented after a pre-trial screen that displayed the following instructions:

*“In this task you will be asked to provide a typicality rating for an object presented on the screen. You will be presented with an image and asked, ‘how typical is this [*object]*?’ For instance, if you are presented with an image of a horse, you will make a judgement on how*

typical (usual/normal/stereotypical) or atypical (unusual/weird/non-stereotypical) the horse is for a 'horse'. You will position your answer on a scale of 'atypical' to 'typical'. Try not to think too hard - there is no right or wrong answer”

In the critical trials, participants were presented with the caption ‘How typical is this [item label]?’ alongside an image corresponding to the referenced object and a response button. Participants’ responses were recorded using a sliding scale between the labels ‘atypical’ and ‘typical’ and participants were able to move on to the next trial after having selected a point along the scale and clicking a button labelled ‘Next’. Note, the terms ‘marked’ and ‘unmarked’ were not used in the task due to ‘markedness’ being an intuitively less familiar concept than ‘typicality’. The trial for the marked ‘bathrobe’ image is depicted below:

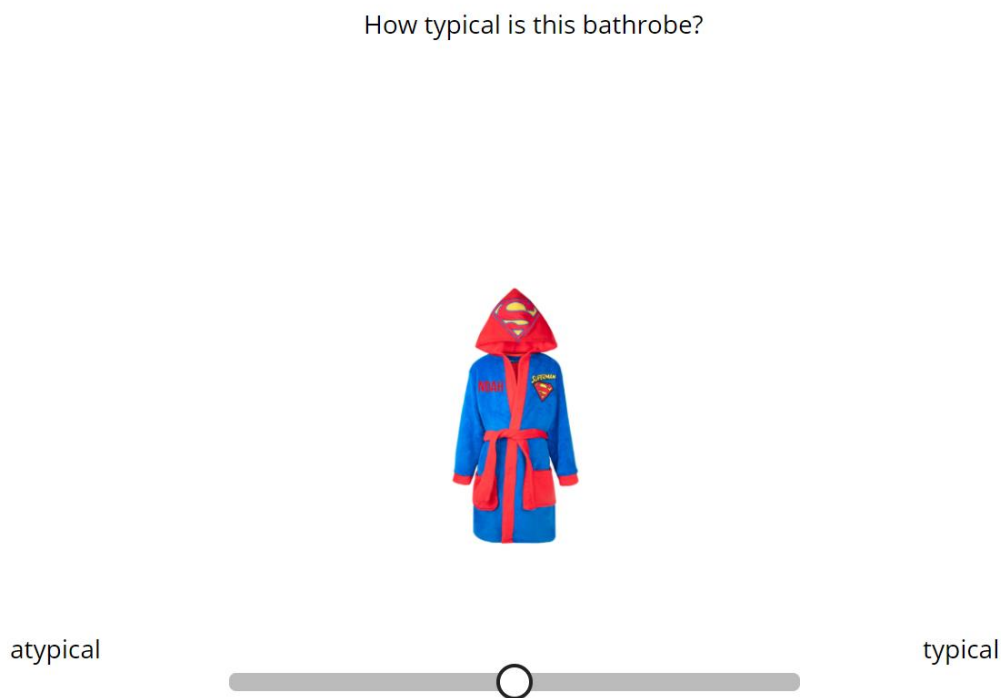


Figure 4-4: trial screen for the item 'bathrobe' in Item Screening Task 2

4.2.3.3. Data & Conclusion

The participant’s sliding scale responses were recorded numerically as a figure between 0 and 100, with 0 representing the selection of the leftmost point of the sliding scale (i.e., most

atypical) and 100 the rightmost (i.e., most typical). Consequently, each trial provided a ‘typicality’ score out of 100 for each item.

The mean typicality score for the marked items stood at 69 (SD = 3.39, range = 27.2 – 55.7) and 40 (SD = 6.01, range = 59.4 – 72.3) for the unmarked items, which suggests that the marked and unmarked images contrast along the lines of markedness. The mean difference between the typicality judgements of the marked and unmarked images stood at 28. While all item pairs demonstrated the expected pattern of typicality contrast, typicality score differences between the marked and unmarked image pairs ranged between 15 and 38 (SD = 6.35), suggesting that the item pairs varied in their degrees of contrast.

In the absence of item pairs that failed to demonstrate the expected pattern of contrast, I used no particular exclusion criteria at this stage of the item screening process. This is because it was unclear how the degree of contrast within the item pairs should affect the availability of MI and I-Implicature when the items are used as experimental stimuli. For instance, it could be the case that a smaller degree of markedness-based contrast between the competing referents increases the importance of linguistic information in discerning the caption’s referent (i.e., leading to more implicature derivation). In contrast, a larger degree of markedness-based contrast between two potential referents may aid the selection of the ‘correct’ item by reducing ambiguity. Consequently, none of the 42 items screened in the task were excluded as potential stimuli in the current study.

In sum, the second item screening task revealed that the marked and unmarked images did indeed contrast in their markedness, but that the degree of contrast varied between item pairs.

4.3. Experiment 1 - MI and I-Implicature Baselines

Experiment 1 aimed to establish the degree to which M-Implicature and I-Implicature were derived in neutral, baseline contexts. Establishing the baseline rate of MI and I-Implicature is important for two reasons. Firstly, the establishment of an experimental baseline is important for subsequent experimentation as a means of measuring the scale of any priming effect.

Secondly, while Wilson (2017) and Wilson and Katsos’ (2016) studies do provide preliminary evidence to suggest that MI can be elicited in experimental settings, Experiment 1 seeks to establish this across a wider range of stimuli. Moreover, the establishment of baseline rates of parallel MI and I-Implicature (e.g., a set comprising of a MI that targets marked ‘shoes’ and

the inverse I-Implicature that targets unmarked ‘shoes’) is useful for investigating the interplay between MI and I-Implicature, helping answer questions addressed later in thesis, for instance, whether MI and I-Implicature are countervailing forces or can facilitate one another’s derivation.

4.3.1. Experiment 1a – MI Baseline

4.3.1.1. Participants

Participants (N = 58) were recruited from Prolific (www.prolific.com) [2024] and were automatically approved for participation on the basis that they were adults (≥ 18 years old), self-identified as native English speakers and had not participated in any of the study’s previous experiments (e.g., the item-screening tasks). Data from participants who did not complete the task and from those who had taken longer than a pre-prescribed time frame (10 minutes) was automatically rejected from inclusion in the study (N = 0).

Note, Prolific.ac. allows any participant who self-elects to participate in an experiment providing they pass the participant screening requirements. However, Gorilla, the experiment builder and host website, can impose additional requirements for the collection of experimental data from these participants, for instance, experiment time limits. To avoid the inclusion of poor-quality data in the study, each of the experiments employed a time limit (between 10 – 15 minutes, depending on the study) after which the data from any experimental trial blocks that had been abandoned or were incomplete were rejected from inclusion in the study. For instance, in Experiment 2c, N = 145 started the study, but N = 150 participants provided the complete datasets that were used in the experimental results’ final analysis. The amount of participants excluded varied depending on how long the time restraint were.

4.3.1.2. Design

As mentioned, the experimental design was based on the ‘Better Picture?’ paradigm used in Bott and Chemla (2016), Rees and Bott (2018) and Marty et al. (2022; 2021). In its basic form, the paradigm involves matching a caption to either of two images, one representing a semantic interpretation of the caption and the other representing the pragmatic interpretation of the caption. Unlike the ‘Better Picture?’ paradigm used in Bott and Chemla (2016) the

current study's paradigm did not include an obscured 'Better Picture?' card representing the pragmatic choice in the critical trials. In contrast, the critical trials are constructed as per Bott and Chemla's priming trials, whereby neither the image representing the semantic interpretation of the caption, nor the pragmatic caption, is covered.

I made this decision on the basis that MI is likely more elusive than other types of EVAs, owing to its context dependency. Consequently, if faced with a choice between a semantically compatible unmarked image and a 'Better Picture?' card, participants might struggle to imagine a marked, MI-compatible image and instead select the unmarked image at higher rates than if both images were present. Unlike Bott and Chemla's (2016) paradigm, I aimed to contextualize the linguistic basis of the experimental trials with a narrative; the paradigm's trials involved the depiction of a conversation between 'Alan', a character participating in a VR-based experimental study and 'Dr Smith', a fictional experimenter conducting the study.

The experiment comprised 126 trials: 42 of which were critical trials corresponding to the 42 marked-unmarked image pairs and 84 of which were filler trials (1:2 ratio). The experiment was conducted within participants, such that each participant provided a response for every possible critical trial. The trials were divided into 42 *filler* → *filler* → *critical trial* blocks, the constituents of which were completely randomised and did not re-enter the selection pool for randomization once they had appeared. As Experiment 1a sought to establish a baseline rate of MI derivation across the stimuli, all the critical trials were in the MI condition.

4.3.1.3. *Materials*

Critical Trials

Experiment 1a's critical trials consisted of the following components: 1) a caption reading "**Dr Smith: 'Can you see a [item label]?'**" that contained one of the 42 item labels presented above, e.g., 't-shirt', 2) a second caption, reading '**Alan: 'Yes, I [can] see [...]'**' 3) the marked and unmarked images representative of the mentioned object and 4) static images of both 'Alan' (a man in a black shirt wearing a VR headset) and 'Dr Smith' (a man dressed in a lab coat). The left-right configuration of the images was randomised in each trial in all trial types across all experiments. Below, shows the critical trial for the item 't-shirt' is depicted:

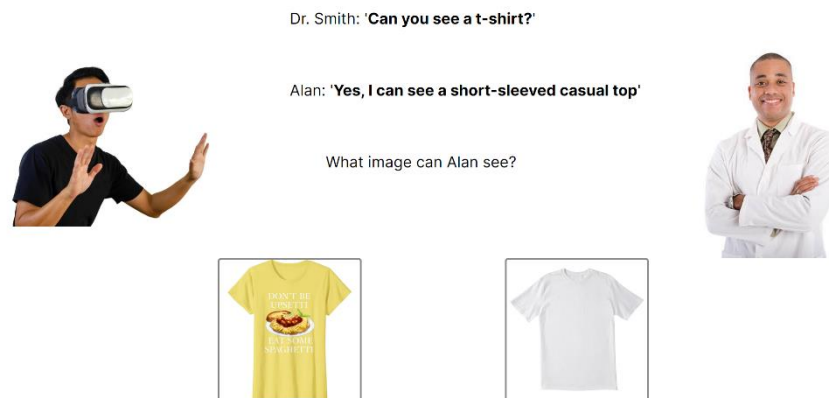


Figure 4-5 MI critical trial for the item 't-shirt'

In the critical trials, the second caption reading “Alan: ‘Yes, I [can] see a [...]’”, contained the use of a prolix description e.g., ‘a short-sleeved casual top’, which aimed to elicit a prolixity-based MI that associated with the marked image of the item mentioned in the caption. The trial’s construction sought to warrant the application of the M-Heuristic’s Recipient’s Corollary ‘What is said in an abnormal way, describes an abnormal situation’ (Levinson, 2000) – or indeed the similar inferential principles posited by Grice (1975) and Horn (1984).

While the item labels used in the ‘Dr Smith’ captions (e.g. ‘t-shirt’) were chosen arbitrarily and before the sourcing of the item images, the prolix descriptions captions used in the ‘Alan’ captions were created after the selection of the item labels and the sourcing of the corresponding marked and unmarked images. The prolix descriptions were based on abridged versions of the item labels’ English-language Google dictionary definitions (Oxford Languages, 2023). I abridged the dictionary definitions on an intuitive, ad-hoc basis, following the principle that the captions should be as ‘conversational’ as possible, i.e., not overly formal or technical, while still being marked in length and complexity.

Consequently, some definitions were paraphrased in favour of ‘everyday’ language. In some instances, the definitions were not marked enough, so were paraphrased to increase their prolixity, e.g., by adding words. Overall, the MI captions had a mean length of 9.9 words; the shortest, ‘I see a close-fitting hat’ (corresponding to the item ‘beanie’) contained 6 unique words and the longest, ‘I can see a piece of clothing that sits on the waist and hangs down the legs’ (corresponding to the item ‘skirt’) contained 17. Half of the captions started with the

phrase ‘I see’, while the other half started with the phrase ‘I can see’ in order to avoid repetitiveness and to try to emulate the types of phrasing variation seen in everyday language use. A full list of the captions can be found in the 9.3.

As mentioned, the MI critical trials aimed to elicit implicature by creating a prolixity-based contrast between the caption (e.g., ‘I can see a piece of clothing that sits on the waist and hangs down the legs’) and the unspoken alternative (e.g., ‘I can see a skirt’). Prolixity-based MI was used over other types of MI for two main reasons. Firstly, Grice’s (1975) seminal illustration of MI presents MI as the product of a prolixity-based contrast, while Levinson’s (2000) and Horn’s (1984) perspectives similarly centre prolixity as a prototypical source of Manner-based inference – although other types of MI are presented, of course.

As such, investigating prolixity-based MI appears a logical starting point for the novel experimental investigation of MI, broadly. Secondly, the type of prolixity illustrated by Grice (1975) is circumlocutory in nature, as it involves the ostensive circumvention of the alternative, e.g, a label. Given that the circumvention of the alternative is the basis of prolixity in prolixity-based MI, the construction of stimuli is relatively simple (i.e., can be governed by the single heuristic ‘omit mention of the item label’) and is also effective for ensuring inter-item uniformity. While the prolixity-based MI-eliciting captions are diverse in length and content, they are uniform in that they all avoid the alternative in favour of a multi-word description.

To address Wilson and Katsos (2023; 2016) and Wilson’s (2017) concerns that a lack of a life-like conversational context may inhibit the elicitation of MI in experimental settings, Experiment 1a’s trials involved a fictional conversational context. Moreover, evidence suggests that implicatures are more likely to be derived if pertinent to the QUD(s) (e.g., Degen, 2023) and Experiment 1a’s fictional context sought to warrant a salient QUD. Therefore, the critical trials depicted a scene in which Dr Smith, an experimenter, asked Alan, an experimental participant, whether he could see a particular item in his VR headset. As part of the experimental instructions, the context of the scene was presented to the experimental participants as follows:

“Alan is participating in a visual recognition experiment and has been tasked with telling the researcher what image he can see in the VR headset. The researcher will show Alan two images in the VR headset, one will then disappear. In this task you will be given sets of two

images and a caption. You will be tasked with matching an image to the caption. Alan is tasked with communicating which image stays, i.e., which image he can still see”

Although it is unclear to what degree the experimental participants consider the experimental narrative and engage with the perspectives of the experimental characters when responding to the trials, the fictional context was constructed to motivate the character of Alan’s use of MI i.e., to disambiguate between two visible, contrasting images. As such, the trials’ context was designed to depict the character of Alan’s adherence to the M-Heuristic’s Speaker’s Maxim ‘Do not use a marked expression without reason’ (Levinson, 2000); Alan was depicted as using the prolix form to contrast with the alternative – the un-prolix form – and its implicature, i.e., the stereotypical image, to reference the atypical image.

In all instances, the ‘Alan’ caption was affirmative, indicating that Alan saw the item mentioned by Dr Smith. I constructed the critical trials to create a sense of contrast between the MI elicited by the ‘Alan’ caption and the alternative (i.e., the countervailing I-Implicature) by having the ‘Dr Smith’ caption contain the corresponding the item label. This decision was motivated by both the theoretical importance of the alternative in MI derivation and experimental suggesting that the salience of the alternative plays a role in implicature derivation (e.g., Rees & Bott, 2018, Bott & Frisson, 2022). The affirmative nature of the ‘Alan’ captions (i.e., the inclusion of ‘Yes [...]’) was designed to ensure that any perceived contrast between the prolix description and the use of the item label in the ‘Dr Smith’ caption was not interpreted as a semantic contrast, but rather a pragmatic contrast. Without including an affirmative, the ‘Alan’ caption could be interpreted by the participants as a negative response to Dr Smith’s question, or an attempt to elicit a QI involving a nonce Horn scale (e.g., <*t-shirt, t-shirt>), as discussed. Consequently, participants might fail to derive the MI in the ‘Alan’ caption and instead select the marked image based on the assumption that it is less likely to be the referent of the ‘Dr Smith’ caption than the unmarked image.

As discussed, the extent to which such MI is used ‘in the wild’ is unclear, and MIs seem to be more often used in situations involving irony, humor or scorn (e.g., as in Example 1-2) than they are in situations involving referential ambiguity. Moreover, MIs are unlikely to appear with much frequency in natural conversation but appear frequently in Experiments 1 – 6. However, to operationalize MI as a vehicle of irony, humor or scorn would be to introduce an additional pragmatic element into the paradigm that may make the effect of the type of MI described by the Gricean-perspective, i.e., a formal *linguistic* phenomenon, more difficult to

analyze. While the addition of features like prosody, body language and facial expression to the paradigm may raise the rate of MI derivation, this deviates from the Gricean-inspired account’s perspective of MI as a linguistic phenomenon, first and foremost, which the focus of Experiments 1 – 6. As such, a relatively neutral type of MI was used in the study, although future research that seeks to increase the rate of MI may benefit from including extralinguistic factors, too.

Filler Trials

Similar to the critical trials, Experiment 1a’s filler trials consisted of the following components: 1) a caption reading “**Dr Smith: ‘Can you see a [item label]?’**”, 2) a second caption, reading ‘**Alan: ‘Yes, I [/can] see [...]**’ 3) two images and 4) static images of both ‘Alan’ (a man in a black shirt wearing a VR headset) and ‘Dr Smith’ (a man dressed in a lab coat). Figure 4-6 depicts the filler trial for the item ‘fork’:

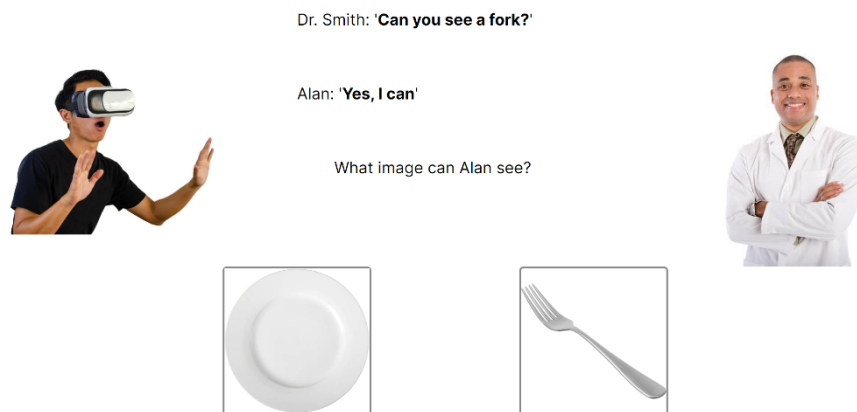


Figure 4-6: 'Yes' filler trial for the item 'fork'.

Unlike the critical trials, component 2 (i.e., the ‘Alan’ caption’) did not contain a prolix description but either the phrase ‘Yes, I can’ or ‘No, I can’t’ and the item labels were related to the semantic fields of clothing (e.g., blouse), inanimate household objects (e.g., fork), animals (e.g., dog) and foodstuffs (e.g., burger).

Although the filler trials involved separate images to those of the critical trials, I chose to include fillers involving articles of clothing to avoid having contrasting semantic fields in the filler and critical trials, which could alert participants to the fact that they were separate types of trial. At the highest level, the filler trials comprised two types, ‘Yes’ fillers and ‘No’ fillers.

The latter was made up of two different trial types, which shall be referred to as *No-False-False* and *No-Marked-Unmarked* trials.

In the *No-False-False* trials, the item mentioned in the ‘Dr Smith’ caption (e.g., fork) was not present in either image and, instead, two images depicting unrelated unmarked items were presented (e.g., a sweatband and some trainers). In contrast, the *No-Marked-Unmarked* presented a pair of same-item marked and an unmarked image unrelated to the caption (e.g., an image of typical dog and an atypical dog and a caption referencing the item ‘fork’). In both, Alan’s response read ‘No, I can’t’ and participants were therefore free to select an image at random.

The ‘Yes’ fillers were trials in which the ‘Alan’ caption read ‘Yes, I can’, and an unmarked image related to the caption (e.g., a typical fork for a ‘fork’-containing caption) was presented alongside an unrelated, unmarked object. The ‘Yes’ trials only included unmarked images to avoid associating marked images with non-prolix item labels, as this is the opposite configuration to the MI trials. Moreover, the ‘Yes’ trials did not involve articles of clothing as the item label or target image. This was to avoid eliciting clothing-related I-Implicature in the filler trials, which could influence the participants’ responses in the critical trials.

In the broadest sense, the filler trials were included to break up the critical trials and were constructed in line with the filler trials presented in Bott and Chemla (2016). Like in Bott and Chemla’s paradigm, the content of the filler trials was constructed to increase the availability of implicature in the critical trials. Including filler trials with plain language ensures that the prolix language of the ‘Alan’ captions in the MI critical trials remains infrequent. As described, linguistic markedness can be defined with regard to its frequency within a given context (Rett, 2020).

Consequently, if the prolixity-based ‘Alan’ captions comprised too large a portion of the language used in the paradigm, there is the risk that participants’ opinions of its degree of relative markedness change, which could affect the availability of MI in the critical trials. Moreover, if the character of Alan was not shown to use language ‘normally’ in trials outside of the critical trials, then participants may begin to assume that prolix language is the only available type of language in the paradigm and may therefore cease to perceive it as having a pragmatic function. In addition, because the ‘Yes’ trials have an objectively correct and an incorrect answer, they have a secondary function as attention checks.

Example filler trials are presented in 9.11

4.3.1.4. Procedure

The experiment was hosted on Gorilla Experiment Builder (www.gorilla.sc) an online experiment building and hosting platform and was accessed by participants via Prolific (www.prolific.com) [2024], a participant recruitment website. After completing a form that collected participant consent and demographic data (see 9.12 and 9.13), the participants were presented with two instruction screens. The first read:

“Welcome to the task! In this task you will be given sets of two images and a caption. You will be tasked with matching the caption to one of the images - you will do this by using clicking on the image”

While the second introduced the fictional context of the scenario depicted in the trials:

“In this experiment, you will be presented with a fictional scenario involving Alan and a researcher. Alan is participating in a visual recognition experiment and has been tasked with telling the researcher what image he can see in the VR headset. The researcher will show Alan two images in the VR headset, one will then disappear. Alan is tasked with communicating which image stays, i.e., which image he can still see. Try not to think too hard about your answer!”

The first two screens navigated with a ‘Next’ button, after which the main experimental trial block began. Each participant gave responses to 126 trials. Trial responses were recorded by participants clicking on either the marked or unmarked image and a new trial was automatically presented after participants gave a response. To prevent participants clicking through the trial blocks too quickly, a 0.5 second buffer was implemented between each trial during which the screen briefly flashed white. Once all of the trial blocks were completed, the participants were redirected back to Prolific (www.prolific.com) [2024] to collect their payment; Participants were paid ‘per hour’ for their participation and average length of the experiments ranged between 5 – 15 minutes. The participants were usually paid around £1.50 – £2.50 each, with an average pay per hour of £8.53, although this varied drastically between participant depending on the advertised length of the experiment and the amount of time they took to complete it.

4.3.1.5. *Predictions*

Given that MI is considered to be context dependent, it is unlikely that MI is derived without a sufficiently robust conversational context. Moreover, MI's 'trigger' (i.e., marked form) is also contextually defined, such that even if MI is contextually warranted, an insufficiently marked form could reduce its availability. Secondly, considering that there exists no accepted blueprint for the experimental implementation of MI and that the current study's methodology is novel, the average rate of MI derivation between items is likely to be low. Moreover, similar studies (e.g., Marty et al., 2022) found baseline rates of enrichment for ad hoc enrichment types.

Considering the Gricean-inspired perspective that MI is a PCI, MI is therefore contextually bound and thus both ad-hoc and heterogenous. As the critical trials of Experiment 1a vary in terms of linguistic and visual input (i.e. context), the degree of item-specific MI availability may not be uniform.

I make the following predictions for Experiment 1a:

- I predict the mean rate of MI to be relatively low across the items.
- I predict there to be a high degree of inter-item variation in the mean rate of MI derivation.
- If MI is not elicited, I expect the rates of marked image selection to be somewhat uniform.

4.3.1.6. *Coding*

The selection of the marked image was interpreted as indicative of MI derivation in the critical trials and participants' responses were coded as either 'correct' or 'incorrect' depending on whether the marked image was selected. The responses were assigned a numerical score, 1 for 'correct' responses and 0 for 'incorrect' responses.

The responses to the filler trials were also coded assigned 'correct' or 'incorrect' values. In the 'Yes' filler trials, the selection of image depicting the mentioned object coded as 'correct'. In the No-Marked-Unmarked trials, responses that selected the unmarked image were coded as 'correct'. While in the No-False-False filler trials, one of the response items was randomly designated as the 'correct' answer and responses that selected this item coded as 'correct' and

those that did not as 'incorrect'. Note, for No-Marked-Unmarked filler trials and No-False-False filler trials, there is no truly correct response.

4.3.1.7. *Results & Discussion*

Across the 'Yes' filler trials for each participant, the percentage of correct responses was calculated. The 'Yes' filler trials were used as proxy attention checks, given that they were unambiguous. Participants (N=0) who gave a correct response to < 95% of the 'Yes' filler trials were excluded from the data analysis on the basis that this was indicative of either poor attention to the study or a misunderstanding of the task's instructions.

The rate of marked image selection (MI derivation) was 15% across all of the MI stimuli. For the filler trials, the rate of correct image selection was 95% for the 'Yes' fillers, 56% for No-Marked-Unmarked filler and 47% for the No-False-False fillers.

As predicted, the mean rate of MI derivation was low across the experimental items. However, the items (or, MIs) showed the anticipated degree of heterogeneity ($r = 0.48$); on the lowest end, the item 'purse' had a 0% rate of MI derivation, while on the highest end, 'cowboy boot' had a 48% rate of MI derivation. That said, the standard deviation ($SD = 089$) indicates that the MI derivation rates across the items did not vary dramatically from the mean, suggesting that the type of prolixity-based MI elicited by the experimental paradigm is generally rather elusive and achieved at far lower rates across the experimental items in Experiment 1a than in Katsos and Wilson (2016)'s study, whereby MI derivation rates of 56% were observed.

Below, Figure 4-7 depicts the mean rate of MI derivation across all of the experimental items.

A more in-depth analysis of the inter-item variance is presented in 4.9, with reference to the data provided by Experiment 1b, the baseline rates of the comparative I-Implicatures.

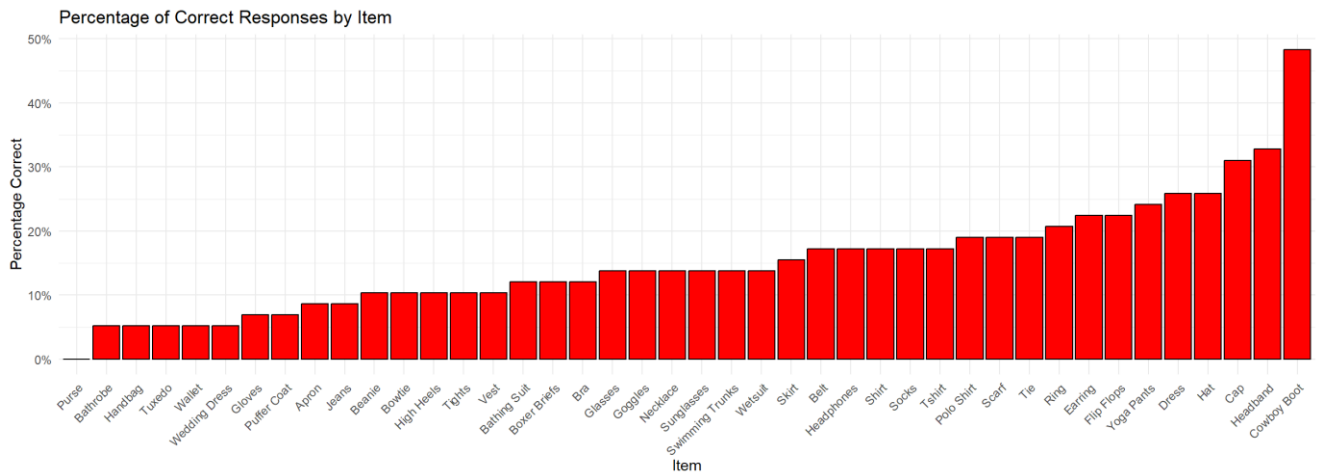


Figure 4-7: The mean rate of MI derivation in Experiment 1a by item ordered smallest to largest.

4.3.2. Experiment 1b – I-Implicature Baseline

Experiment 1b established the baseline rates of I-Implicature across the set of 42 experimental stimuli to complement the MI baseline established in Experiment 1a. Moreover, I-Implicature prime trials are used in Experiment 3. Therefore, establishing a baseline for the I-Implicature stimuli is important for understanding their efficacy when used in prime trials.

The I-Implicature baseline can be viewed from two perspectives. Firstly, the results of the I-Implicature baseline can be considered to represent the rate of I-Implicature derivation across the set of experimental stimuli and, therefore, be considered indicative of the general availability of I-Implicature. That said, the concept of ‘I-Implicature’ is restricted to Levinson’s (2000) account of conversational implicature and the Gricean-inspired perspective does not clearly acknowledge the existence of similar ‘plain language’ implicatures. From this perspective, the ‘I-Implicature’ baseline can alternatively be considered a neutral baseline (i.e., a ‘no implicature’ baseline) that can contextualise the rates of MI derivation in Experiment 1a.

4.3.2.1. Participants

Participants (N = 25) were recruited from Prolific (www.prolific.com) [2024] and selected on the basis that they were adults (> 18 years old), self-identified as native English speakers and had not participated in any of the study’s previous experiments. N = 1 of the participants was

excluded from the analysis for providing incomplete data, therefore data was collected from $N = 24$ participants.

Fewer participants were recruited in Experiment 1b than 1a for two reasons. Firstly, the MI critical trials, but not the I-Implicature critical trials, are used in the priming studies in later experiments. Consequently, a more robust baseline is needed for the former. Secondly, the rates of I-Implicature derivation are likely more stable than those of MI, given that I-Implicature is reportedly generated by default, whereas MI is dependent on context and its availability likely varies more between item and between participant, warranting a larger participant pool.

4.3.2.2. *Design*

The overarching structure and content of Experiment 1b's critical trials were identical to those of the marked-unmarked image pairs presented in Experiment 1a and 84 filler trials (1:2 ratio). The experiment was conducted within participants, and each participant provided a response for all possible critical trials. Like Experiment 1a, the trials were divided *into filler* → *filler* → *critical trial* blocks, and the constituents of these blocks were completely randomised, with each only appearing once. Experiment 1a. Experiment 1b consisted of 42 critical trials that corresponded to those of Experiment 1a. See 9.1 and 9.4 for a list of the items.

4.3.2.3. *Materials*

Critical Trials

Experiment 1b's critical trials consisted of the same components as Experiment 1a: 1) a caption reading “**Dr Smith: ‘Can you see a [item label]?’**” which contained one of the 42 item labels presented above, e.g., ‘t-shirt’, 2) a second caption, reading ‘**Alan: ‘Yes, I [can] see [...]**’ 3) the marked and unmarked images representative of the mentioned object and 4) static images of both ‘Alan’ (a man in a black shirt wearing a VR headset) and ‘Dr Smith’ (a man dressed in a lab coat). Below, Figure 4-8 shows the critical trial for the item ‘vest’:

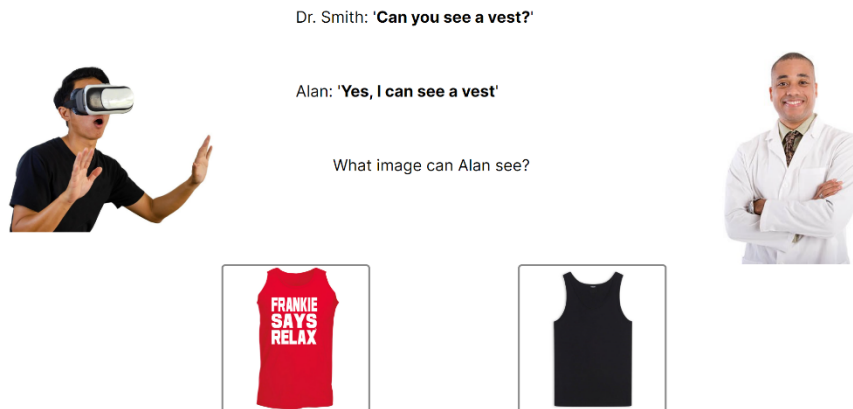


Figure 4-8: The I-Implicature critical trial for the item 'vest'

While Experiment 1a's critical trials sought to elicit MI, Experiment 1b's were designed to elicit I-Implicature such that the caption "Alan: 'Yes, I [/can] see a [...]"", contained the use of the item label, e.g., 'vest', rather than a prolix description thereof. Nevertheless, the trials contained the exact images as those used in Experiment 1a. The critical trials in Experiment 1b were designed to engage the I-Heuristics Recipient's Corollary 'What is simply described is stereotypically exemplified' (Levinson, 2000), such that participants associated the unmarked language (e.g., 'vest') with the corresponding unmarked image. The I-Implicature critical trials aimed to elicit an implicature-to-stereotype via the use of the item label. Levinson's Implicature-to-stereotype can be considered the I-Heuristic's analogue to prolixity-based MI, given the role implicature-to-stereotype plays as the alternative in the derivation of prolixity-based MI (Levinson, 2000).

According to the Gricean-inspired perspective, I-Implicature is not an enrichment via alternative (EVA) and is conceived of as more-or-less default in its generation. Consequently, I did not construct the I-Implicature critical trials to increase the likelihood of I-Implicature derivation (i.e., involve contrast between the item label and another phrasing). Rather, I sought to create a sense of concordance between the 'Dr Smith' caption and the 'Alan' caption by repeating the same I-Implicature in both. Nevertheless, the repetition of the item label in the 'Alan' caption was inevitable, as a response without the item label (e.g., 'Yes, I can') cannot be said to elicit I-Implicature in the same way.

Because Experiment 1b involved only the set of pre-established item labels, the captions did not require additional construction, and they were shorter than the prolix captions used in Experiment 1a, with an average length of 4.7 words. Half of the captions started with the phrase ‘I see’ and the other half ‘I can see’ to achieve a more naturalistic utterance variation.

Fillers Trials

Experiment 1b involved the same filler trials as Experiment 1a, the ‘Yes’, No-Marked-Unmarked and No-False-False filler trials.

As the filler trials do not involve pragmatic ambiguity (i.e., either a single image semantically matching the ‘Dr Smith’ caption is presented, or none are), the ‘Alan’ captions are either ‘Yes, I can’ or ‘No, I can’t’. In the ‘Yes’ filler trials, the absence of a caption containing the item label (e.g., ‘Yes, I can see a fork’, below) highlights the presence of the item label in the I-Implicature critical trials. Although I-Implicatures are not computed in the same fashion as other implicatures and may not require contextual input (Levinson, 2000), I constructed the ‘Yes’ filler trials to contrast with Experiment 1b’s critical trials to focus the disambiguating use of the item label in the latter.

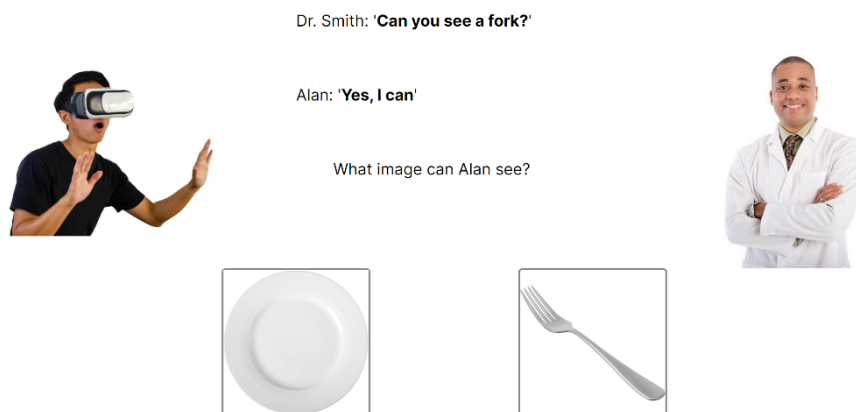


Figure 4-9: Filler trial for the item 'fork'

4.3.2.4. Procedure

As with Experiment 1a, the task experiment was hosted on Gorilla Experiment Builder (www.gorilla.sc) and was accessed by participants via Prolific (www.prolific.com) [2024]. After providing consent and demographic data, the participants were presented with the same two instruction screens as Experiment 1a, repeated below.

Instruction Screen 1: *“Welcome to the task! In this task you will be given sets of two images and a caption. You will be tasked with matching the caption to one of the images - you will do this by using clicking on the image”*

Instruction Screen 2: *“In this experiment, you will be presented with a fictional scenario involving Alan and a researcher. Alan is participating in a visual recognition experiment and has been tasked with telling the researcher what image he can see in the VR headset. The researcher will show Alan two images in the VR headset, one will then disappear. Alan is tasked with communicating which image stays, i.e., which image he can still see. Try not to think too hard about your answer!”*

Each participant gave responses to 126 trials, which were recorded by participants clicking on either of the item images.

4.3.2.5. Predictions

According to Levinson (2000), I-Implicature is generated by default and, therefore, should be readily derived in Experiment 1b’s critical trials. Moreover, as per the I-Heuristic’s Recipients Corollary, listeners operate on the assumption of referential parsimony and, therefore, do not attempt to derive implicatures for which there is no appropriate linguistic cue (e.g., they should not interpret ‘bra’ to M-Implicate ‘+> an unusual bra’). Moreover, the low rates of MI in Experiment 1a’s MI critical trials suggest that participants avoid selecting the marked image, even in contexts that warrant it.

Considering this, I make the following predictions for Experiment 1b:

- I expect a larger degree of unmarked image selection in the critical trials of Experiment 1b compared to that of Experiment 1a.
- I predict a higher rate of I-Implicature in Experiment 1b than the rate of MI in Experiment 1a.

- Given the low rates of MI in Experiment 1a, the rates of implicature in Experiment 1b may appear complementary to those of Experiment 1b – that is, there may be a similar proportion of unmarked image selections in both experiments.
- If I-Implicature is not derived, results that indicate the random selection of either image, i.e. a 50:50 ratio, are expected.

4.3.2.6. Coding

The selection of the unmarked image was interpreted as indicative of I-Implicature derivation in Experiment 1b’s critical trials, and responses were coded as 1 (i.e., ‘correct’) if they selected the unmarked image and as ‘0’ (i.e., ‘incorrect’) if they selected the marked image. The responses to the filler trials were coded in the same way as those of Experiment 1a.

4.3.2.7. Results

The rate of correct image selection was > 95% for each participant across the ‘Yes’ filler trials. Therefore, no participants were excluded from the analysis. The rate of unmarked image selection was 87% across the I-Implicature trials. For the filler trials, the rate of correct image selection was 95% for the ‘Yes’ fillers, 56% for the No-Marked-Unmarked filler trials and 47% for the No-False-False fillers.

The rate of I-Implicature derivation varied between items ($r=0.33$, $SD=0.08$) at a similar distribution to that of MI in Experiment 1a. The lowest rate of I-Implicature was observed in the item ‘beanie’ at 67%, while the items ‘bathing suit’, ‘tights’ and ‘wedding dress’ all had

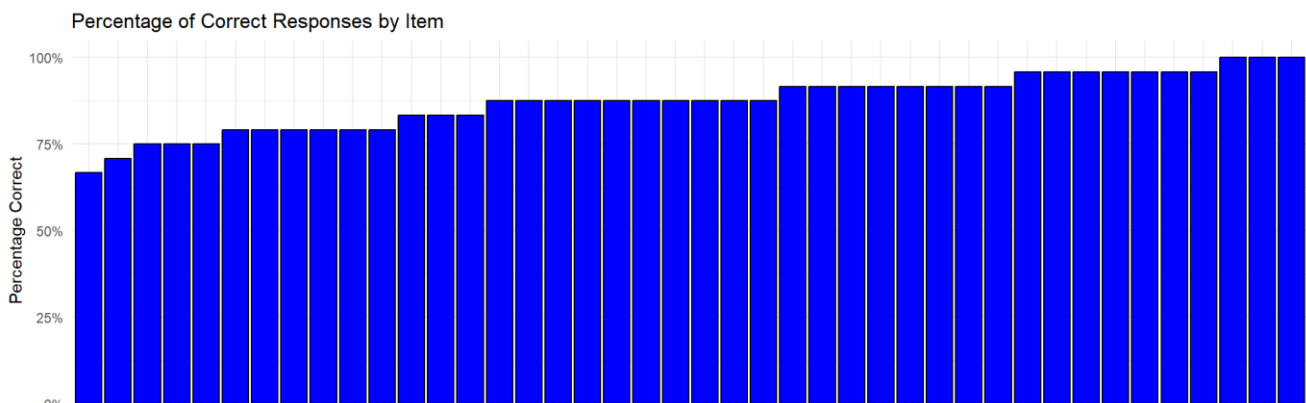


Figure 4-10: The by-item rates of unmarked image selection in Experiment 1b.

100% rates of I-Implicature derivation. Figure 4-10 demonstrates the rates of I-Implicature across all the experimental stimuli:

4.4. Experiment 1: Combined Analysis

The purpose of Experiment 1 was to establish a baseline rate of MI and I-Implicature derivation. Establishing a baseline rate for MI derivation in Experiment 1a was important for two main reasons. Firstly, the baseline confirmed whether the paradigm could elicit MI. Secondly, the baseline is necessary to ascertain the magnitude of the priming effects in the later priming studies (Experiments 2 – 6). Establishing the baseline rate of I-Implicature in Experiment 1b was necessary to contextualise the baseline of MI – after all, without comparing the MI baseline to a second ‘neutral’ baseline, it is impossible to assess whether MI is derived in the critical trials. Experiment 1b also allowed me to collect an independent baseline for ‘I-Implicature’ derivation, which is useful in exploring the nature of I-Implicature and its relatedness to MI – a secondary aim of this thesis.

Alongside this, the individual baseline rates of MI and I-Implicature across the 42 experimental items allowed me to investigate item-specific questions, which are important for understanding how MI and I-Implicature are elicited. Firstly, I explored how the rates of MI differed across the items, aiming to understand which items likely elicited MI in Experiment 1b. Secondly, I investigated the effect the typicality of the images had on MI and I-Implicature availability. As such, the discussion below focuses on three additional questions:

- Does the difference in marked image selection between the conditions vary by item?
- Are the rates respective rates of I-Implicature and MI per item related?
- Does the typicality of the respective images and their interaction influence the availability of MI and I-Implicature?

4.4.1. Data analysis

As per Barr, Levy, Scheepers and Tily’s (2013) recommendations concerning confirmatory hypothesis testing using mixed-effects models, I began with the maximal random effects model justified by the experimental design and simplified this model if it did not achieve

convergence¹¹. For the main effect of the experimental condition for Experiments 1 – 6, I used a generalised, linear mixed regression model with a binomial distribution and a logit function (i.e., a logistic regression), with the condition (e.g., Experiment 1a vs. Baseline) as a fixed effect and random intercepts for experimental item and individual participant, as well as random slopes for experimental item¹². The model predicted the likelihood of marked image selection (i.e., MI derivation). The models were fitted in RStudio using the lme4 package (Bates et al., 2015; R Core Team, 2016).

4.4.2. Main Effect of Condition

While the rate of implicature derivation was ostensibly very different between Experiment 1a and Experiment 1b (i.e., 15% and 87%, respectively), the rate of unmarked image selection was similar between the experiments – in Experiment 1a, the rate of unmarked image selection was 85%. Although I predicted the rates of marked vs. unmarked image selection to be complementary, the response patterns in both conditions were very similar. Despite their numerical similarity, analysis using the mixed-effects logistic regression model described above revealed that the MI critical trials, i.e., Experiment 1a, were associated with the significantly increased odds of selecting the marked image compared to Experiment 1b ($\beta = -0.22$, $SE = 0.11$, $z = -1.97$, $p = .05$).

4.4.3. Item Analysis

As observed, the response patterns in Experiment 1a and 1b vary between the items in both Experiment 1a and 1b – see Figure 4-10 and Figure 4-7. Assuming that the significant increase in the selection of the marked image in Experiment 1 indicates MI derivation, it is important to understand whether MI is elicited across all of the items or whether some items drive this effect but not others. The degree to which MI is elicited across the experimental

¹¹ N.B. In the body of the ‘Data Analysis’ section for Experiments 1 – 6, I have indicated where convergence has not been achieved, and the model has been simplified. Where this is not indicated, the models successfully converged.

¹² The models did not include random slopes for participant. As each participant provided only one critical trial response per item type, and because the model is dominated by large inter-item variance, participant-level variability is difficult or impossible to estimate. Initial attempts to include slopes for participant caused convergence issues and I therefore decided to omit the inclusion of participant slopes in favour of an intercepts model.

items can be estimated by calculating the similarity between the rates of unmarked image selection in Experiment 1a and Experiment 1b. Figure 4-11 presents the sum of each item’s mean rate of MI and I-Implicature derivation across Experiments 1a and 1b, with values closer to 100% representing a parity between the rates of marked image selection in each baseline condition.

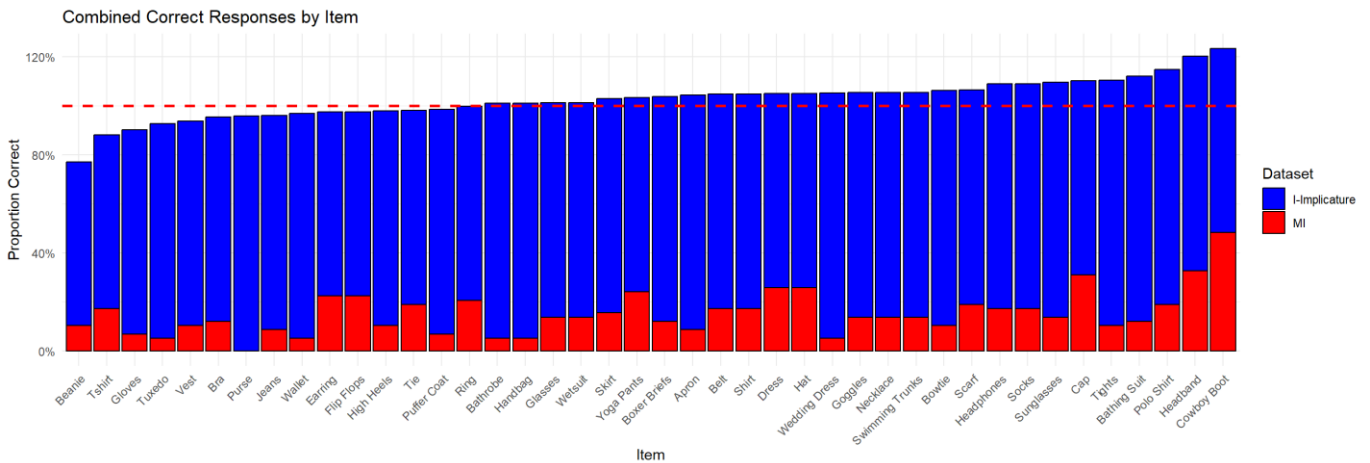


Figure 4-11: Combined rates of respective implicature derivation by item for both Experiment 1a and 1b

As demonstrated, the parity of marked image selection between Experiment 1a and 1b varies across the items. For some items, e.g., ‘beanie’, there is less selection of the marked image in Experiment 1a’s MI baseline (i.e., 10%) than there is in Experiment 1b’s I-Implicature baseline (i.e., 45%), which could suggest that Experiment 1a’s critical trials do not consistently elicit MI. For other items, e.g., ‘cowboy boots’, there is more selection of the marked image in the MI baseline (i.e., 48%) than the I-Implicature baseline (25%), which could be indicative of MI derivation in the former, but not the latter. Of the 42 item trials, 17 had more than a 5% increase in the rate of marked image selection between the I-Implicature baseline and the MI baseline (the smallest increase was in ‘hat’, at 5%). In contrast, five items showed a 5% decrease in the rate of marked image selection between the I-Implicature baseline and the MI baseline.

4.4.3.1. Cause of Inter-Item Variance

The logistic mixed-effects model used to investigate the main effect of the condition included the item type as a random effect. The analysis revealed that the random effect was statistically

significant, indicating a variance in the intercept between items. Specifically, the variance of the random intercepts was $\sigma^2=0.09$ with a standard derivation of $\sigma=0.29$. To better understand the contribution of item type to the overall main effect of the baseline conditions, I ran a likelihood ratio test that compared the original mixed-effects logistic regression model to a logistic regression model with only the baseline condition as a fixed effect. The results suggest that including the random intercept for item type significantly improves the fit in the model, resulting in a significant reduction in deviance ($\chi^2(1)=1667.3, p<.001$). This suggests that the item type is a strong predictor of the likelihood of MI implicature in Experiment 1a.

4.4.3.2. *The Relationship Between The Availability of I-Implicature and MI*

To investigate whether the availability of item-specific I-Implicature influences the availability of the corresponding MI, I ran a linear regression using the by-item average rates of implicature derivation in both baselines. The results showed that rates of I-Implicature derivation in Experiment 1b significantly predicted the rates of MI derivation in Experiment 1a ($\beta = -0.59, t(40) = -4.0, p < .001$) and revealed a moderate negative correlation between the conditions. This indicates that a higher I-Implicature score in Experiment 1b predicted lower rates of MI derivation in Experiment 1a. However, the model indicates that a minority of the variance in MI rates between the items is explained by the corresponding rates of I-Implicature ($R^2=.29$). I conducted a Pearson's product-moment test that confirmed a moderate, negative correlation between the two conditions ($r(40) = -0.54, p < .001$).

4.4.3.3. *The Relationship Between Item Typicality and I-Implicature Availability*

I ran a logistic regression model with the typicality of the marked image and the typicality of the unmarked image as fixed effects and the likelihood of I-Implicature in Experiment 1b as the outcome variable. The analysis revealed that while the typicality of the unmarked image was not a significant predictor of I-Implicature ($\beta=0.03, p=.35$), the typicality of the marked image was a significant negative predictor of I-Implicature ($\beta=-0.04, p=.03$). This suggests that as the marked images become more typical, the likelihood of the participants deriving I-Implicature decreases.

I then updated the logistic regression model to examine whether the interaction between the typicality of the marked image and the typicality of the unmarked image predicted the likelihood of I-Implicature, finding that the interaction between the typicality of the marked image and the typicality of the unmarked image did not predict I-Implicature derivation ($\beta = -0.0001, p = 0.97$). In the updated model, neither the typicality of the marked image ($p = 0.86$) nor the typicality of the unmarked image ($p = 0.93$) was a significant predictor of I-Implicature derivation. The models were compared with the Akaike Information Criterion (AIC) to determine the best model fit for predicting the likelihood of I-Implicature derivation in Experiment 1b but fit the data almost equally. For the regression with the typicality of the images as a main effect the AIC = 768.08, while for the regression examining their interaction the AIC = 770.08.

4.4.3.4. *The Relationship Between Item Typicality and MI Availability*

I ran a logistic regression model with the typicality of the marked image and the typicality of the unmarked image as the main effects and the likelihood of MI as the outcome variable for Experiment 1a. The analysis showed that the typicality of the unmarked image was a significant negative predictor of MI ($\beta = -0.07, p < .001$), and the typicality of the marked image was a significant positive predictor of MI ($\beta = 0.02, p = .02$) in Experiment 1a's critical trials, indicating that as the typicality of the unmarked image increased, the likelihood of MI derivation decreased, but that as the typicality of the unmarked object increased, the likelihood of MI increased. I then adjusted the logistic regression model to account for any interaction between the typicality ratings of both of the images on the likelihood of MI in Experiment 1a.

The updated model indicated that the main effects of both the typicality of the marked image ($\beta = 0.44, p < 0.001$) and the typicality of the unmarked image ($\beta = 0.87, p < 0.001$) were significant positive predictors of MI derivation. As typicality increases in both items, the likelihood of MI increases. However, the interaction between the main effects is a significant negative predictor of MI derivation ($\beta = 0.01, p < .001$), which suggests that the main effects of the typicality of each image are not additive but that when they increase simultaneously, they decrease the likelihood of MI derivation.

4.5. Experiment 1: Combined Discussion

4.5.1. Manner Implicature

The data analysis of Experiments 1a and 1b suggests that there is a higher likelihood of selecting the marked image in Experiment 1a, the MI condition. Considering the only difference between the critical trials of Experiment 1a and Experiment 1b was the presence of the prolix captions, I suggest that the marginal increase in the rate of marked image selection is indicative of an increased rate of MI in Experiment 1a's critical trials (Experiment 1a: 15% vs. Experiment 1b: 13%). Nevertheless, the theoretical significance of a 2% difference in marked image selection requires justification, regardless of its statistical significance.

Considering this, the fundamental difference between the inferences available in the critical trials in Experiment 1a and 1b means that respective rates of marked image selection, 15% vs. 13%, cannot be compared in absolute terms. In Experiment 1b, the participants can either derive MI or can derive I-Implicature based on the semantic content of the prolix caption – this perspective assumes that I-Implicature is the default alternative to other response strategies, i.e. MI. On the flip side, the item labels used in Experiment 1b are highly unlikely to elicit MI. Consequently, the responses to Experiment 1b are likely to only comprise either I-Implicature derivation or random selections, even if the 'incorrect' image, i.e., the marked image, is selected. Therefore, the 13% rate of marked image selection in Experiment 1b is highly unlikely to reflect a rate of MI derivation.

Considering this, the proportion of responses involving MI derivation in the critical trials of Experiment 1a is likely far larger than the rate of 2 % indicated by comparing the absolute 'increase' between the marked image selection rate in Experiment 1a and 1b. Therefore, it is important to clarify that the I-Implicature condition (Experiment 1b) should not be considered a baseline with which to interpret the rate of MI derivation in Experiment 1a after all. Instead, Experiment 1b should be considered a baseline for rates of I-Implicature derivation rather than a 'neutral' condition.

4.5.1.1. *The Availability of MI*

The low rate of MI availability in Experiment 1a is predicted by the Gricean-inspired account, whereby MI is not considered to be derived by default and instead requires ad-hoc,

context-dependent calculation. Consequently, this likely predicts lower baseline rates of implicature derivation compared to other types of implicature that might be derived more readily. In comparison, Marty et al. (2022) established a baseline rate of some QI derivation of 59% but a far low rate of ad-hoc QI derivation of 11%, which is similarly context-dependent enrichment type.

In contrast, Wilson and Katsos's (2016) investigation of MI involved prolixity-based MI alongside generalised, periphrastic-causative MI and found far higher rates of MI (56%) across both than the rate observed in Experiment 1a. However, Wilson and Katsos' paradigm involved presenting participants with a MI-eliciting caption and asking which of two sentences was most similar to it. One of the sentences explained the manner-implicated meaning, while the other was neutral. See below:

"Nick and Dan were watching a history programme, but Nick fell asleep. Afterwards, he asked, 'What happened at the end?' Dan replied, 'The invaders caused the villagers to die'"

'The invaders killed the villagers but indirectly, by introducing disease' **OR** 'The invaders killed the villagers directly.'

Example 4-1

Unlike Wilson and Katsos' paradigm, the critical trials in Experiment 1a did not highlight the possibility of MI, nor did they lay bare the particular MI the 'Alan' caption sought to elicit. As such, the rates of MI derivation observed in Experiment 1b are predictably lower and are perhaps a better reflection of the availability of MI generally. Moreover, the previous attempts of Katsos and Wilson (2016) and Wilson (2017) to elicit MI experimentally did so across far fewer items (6 and 9, respectively). Given that Experiment 1a involves 42 different stimuli, the rate of MI derivation is far more likely to be reflective of a general rate of prolixity-based MI and less likely to be representative of the availability of the particular MIs involved in the study than in the baselines established by Katsos and Wilson (2016) and Wilson (2017). Although Katsos and Wilson (2016) do not report a difference between the availability of generalised vs. particularised MI, the current study involves particularised (i.e., prolixity-based) MI, which likely show lower rates of derivation in contexts that do not directly support their derivation when compared to generalised MI.

4.5.1.2. *Effect of item*

As demonstrated by the comparison of the models with and without item type as a random effect, the increase in the likelihood of marked image selection between Experiment 1a and 1b is influenced by the experimental items. As the effect of the condition varies across the items, this indicates that some of the experimental items may elicit different degrees of MI than others. Although this may appear obvious when Figure 4-7 and Figure 4-10 are considered, the significance of item type as a random effect suggests that the variance in both baselines is not simply a product of the items but also of the interaction between the items and the experimental conditions.

According to the Gricean-inspired perspective, MI derivation is reliant on the computation of an alternative I-Implicature. Considering the salience of the alternative is thought to facilitate implicature derivation, e.g., Bott and Frisson (2022), higher rates of I-Implicature in Experiment 1b should predict higher rates of MI derivation in Experiment 1a between the same items (e.g., ‘dress’ in both the I-Implicature and MI condition). Nevertheless, the linear regression predicts a negative correlation between the by-item rates of implicature in each condition, such that higher rates of I-Implicature derivation in Experiment 1b predict lower rates of MI derivation in Experiment 1a. Although this finding contrasts with the predictions made by the Gricean-inspired perspective, considering the overall rates of MI and I-Implicature within Experiment 1 are almost complementary (15% and 87%, respectively), the negative correlation between the two scores is likely a product of the similar rates at which the marked or unmarked image is selected in each item, regardless of condition. Consequently, it is difficult to ascertain whether more available I-Implicature facilitates the derivation of respective MI.

4.5.1.3. *Item Typicality*

As described, markedness is foundational to the derivation of MI – after all, MI is essentially a marked-to-marked pairing between an utterance and a referent. Consequently, evidence suggesting that the typicality of each image and their interaction predicts the likelihood of MI provides convincing evidence that the critical trials in Experiment 1b indeed elicit MI. On the other hand, if there was no evidence to suggest that the likelihood of MI was influenced by

the typicality of either image, this may indicate that the responses were not guided by a marked-to-marked association.

Broadly speaking, there are three patterns of interaction. Firstly, there appears to be a positive effect of the typicality of the marked image, such that as the typicality of the marked image increases linearly, the rate of MI derivation does, too – this effect is predicted by both regression models. Secondly, there is mixed evidence concerning the direction of the effect of the typicality of the unmarked image. In the regression model without the interaction between the effects, the effect of the marked image was negative ($\beta=-0.07, p < .001$), suggesting that as the typicality of the marked image increased, the rate of MI derivation decreased. In contrast, in the model fitted with the interaction, this effect was positive ($\beta=0.02, p=0.02$), which suggests the opposite. Thirdly, the interaction between the main effects was a significant negative predictor of MI derivation ($\beta=0.01, p < 0.001$). This suggests that as the typicality of both increases, the rate of MI derivation decreases. However, none of the effects predicted a large increase or decrease in the likelihood of MI, in any case.

While the effects are difficult to interpret, evidence that typicality influences the rate of MI derivation is predicted by the Gricean-inspired view that MI is context-dependent and reliant on a marked referent – indeed, the small negative correlation between the typicality of the unmarked image and the rate of MI derivation may suggest that more marked, marked images lead to higher rates of MI derivation. Nevertheless, the majority of the responses to the critical trials in Experiment 1a (~85%) do not involve MI; therefore, assessing the effect of item typicality on the rate of MI derivation is likely obfuscated by the responses in which it does not occur.

4.5.1.4. Conclusion

In sum, I conclude that the critical trials of Experiment 1a do elicit MI but that the availability of MI is low. In all likelihood, the rate of MI derivation in Experiment 1a is around 15% or lower if random selection and other response strategies are accounted for. Moreover, the rate of MI derivation appears to be influenced by the individual experimental items, and part of this influence appears to be tied to the typicality of the images involved in each trial – that said, the direction and size of these effects are difficult to ascertain.

Although the 42 individual critical trials were constructed to elicit MI as uniformly as possible, each individual item varies in its rate of marked image selection ($r=0.05$).

Consequently, it is likely that MI is more available in some items than others. For instance, for more ambiguous items, e.g., ‘cowboy boot’, the bulk of the marked image selections may be driven by I-Implicature rather than MI, while this pattern may be reversed for other items. As a result of the ambiguity concerning which experimental items elicit MI and to what degree, I did not exclude any of the experimental stimuli from use in the succeeding experiments based on their rate of MI derivation in Experiment 1a.

Overall, the findings are compatible with the Gricean-inspired perspective of MI derivation, whereby MI is considered a heterogeneous, context-dependent type of conversational inference.

4.5.2. I-Implicature

The baseline rate of I-Implicature (87%) established in Experiment 1b is far higher than that of the MI baseline (15%). High rate of I-Implicature availability is predicted by Levinson’s (2000) perspective that I-Implicature is generated by default without contextual input. As described, it is unclear what response strategies are used in Experiment 1b’s critical trials when the marked image is selected over the unmarked image. Considering this is highly unlikely to indicate MI, then it is likely the result of either random selection or ‘misguided’ I-Implicature. Given this, the rate of I-Implicature in Experiment 1b is likely as high, or higher, than 87%.

While the Gricean-inspired account is divided in the importance it places on ‘I-Implicature’, the results of Experiment 1b do evidence the psychological reality of the unmarked-to-unmarked association between utterance and referent; the observed association is not a product of a generic preference for the selection of the unmarked image, as evidenced by the Marked-Unmarked-False filler trials that involved marked and unmarked images, but no labels, in which the unmarked image was selected at a rate of 56%. The Marked-Unmarked-False filler trial for the item ‘lamp’ is depicted below:

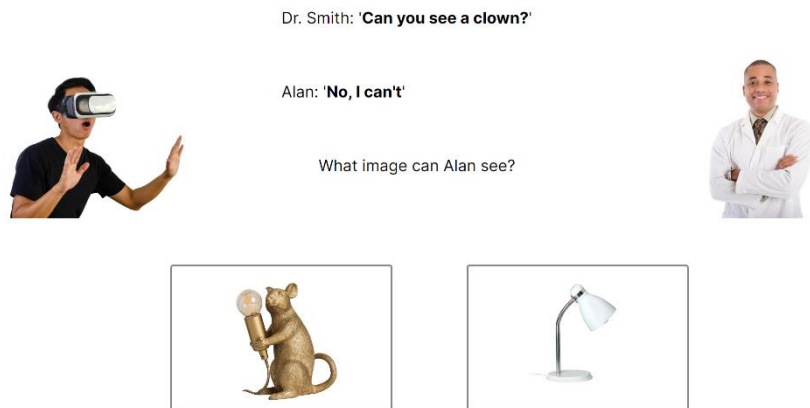


Figure 4-12: the Marked-Unmarked-False filler trial for the item 'lamp'

4.5.2.1. *Effect of Item*

Unlike in the MI critical trials, the derivation of I-Implicature in Experiment 1b's critical trials does not appear to be modulated by the interaction between the typicality of the marked and unmarked images.

A lack of an effect of the interaction between the respective images is predicted by Levinson's (2000) conceptualisation of I-Implicature whereby I-Implicature is derived by default, does not involve the consideration of an alternative way of speaking and is not sensitive to contextual input. Nevertheless, the results of the regression model indicated that there is a negative relationship between the typicality of the marked image and the likelihood of I-Implicature derivation ($\beta = -0.04$, $p = .026$), such that as the typicality of the marked image increases, the likelihood of I-Implicature derivation decreases. However, this observation is easily explained as the result of relatively unmarked marked images being more likely to be selected as referents of I-Implicature than relatively marked, marked images, given that the former confer more ambiguous critical trials.

Consequently, the interplay between the typicality of each image does not appear to affect the likelihood of I-Implicature derivation, which is predicted by Levinson's (2000) account.

4.5.3. Overall Conclusion

To conclude, Experiment 1a and 1b's results suggest that the constructed paradigm is successful in eliciting MI and I-Implicature, with the caveat that the extent of the implicature derivation is unknown. Evidently, the availability of MI is far lower than that of I-Implicature, which is a pattern predicted by the Gricean-inspired perspective. As established, there are fundamental differences between the responses that are available in each of the critical trial types. Namely, in the MI critical trials, participants can respond and derive MI or derive I-Implicature, providing the captions are interpreted without pragmatic enrichment. On the other hand, MI is not available in the I-Implicature critical trials. As such, the experiments should not be considered counterparts to one another in light of this asymmetry.

The observation that the rates of MI derivation vary between the experimental items opens up interesting discussion concerning the nature of MI and how it is elicited by the paradigm. Although investigating the factors that influence the availability of MI is not a primary concern of this thesis, this question will be revisited in 4.9 as it can provide an interesting insight into the nature of MI, more generally. Nevertheless, the findings from Experiment 1a remain inconclusive concerning the source of the variation in MI rates between the experimental items.

4.6. Experiment 2a – Manner-to-Manner Priming

Having established the baseline rates of MI derivation, Experiment 2a investigated whether or not the derivation of MI in preceding trials can prime MI derivation in succeeding trials. Understanding whether MI can be primed is important for multiple reasons. Firstly, if MI can indeed be primed, then this provides evidence that the mechanism responsible for MI derivation, or some level of it, can be directly accessed. In turn, this can provide an insight into how the mechanism that derives MI may be structured. Moreover, if MI can be primed by the paradigm, then the paradigm can be used to broadly investigate the nature of the mechanism(s) responsible for the derivation of implicature, a central goal of this thesis.

Furthermore, understanding whether priming exists between the stimuli used in Experiment 1a sheds light on the extent to which MI is processed as a homogenous phenomenon. If priming exists between the experimental items, it suggests that the experimental stimuli are derived by the same mechanism. Therefore, evidence of a between-category priming effect

would suggest that the mechanism used to derive MI is generic and may be shared between different types of MI or different types of implicature, more generally.

Finally, evidence that MI can be primed by MI suggests that implicature derivation can be manipulated, which provides an insight into the factors that increase the availability of MI beyond those discussed in light of Experiment 1a, e.g., the residual activation of the derivational mechanism, or context adaptation, etc. In turn, if MI can be primed by other MI, then this suggests that MI can be primed by related phenomena, which is important to establish for the between-category priming investigation in Experiments 3 – 6.

4.6.1. Participants

Participants (N = 80) were recruited from Prolific (www.prolific.com) [2024] on the basis that they were adults (≥ 18 years old), self-identified as native English speakers and did not participate in any of the study's previous experiments. Incomplete data was rejected from N = 9 of the participants such that data was collected from N = 71 participants.

4.6.2. Design

The content of Experiment 2a is largely identical to that of Experiment 1a, except for the inclusion of the prime trials. Overall, Experiment 2a involved 42 critical trials that corresponded to the 42 marked-unmarked image pairs used in Experiment 1a and 42 prime trials that were identical to the 42 critical trials. Unlike previous experiments, Experiment 2a had a between-participants structure, such that each participant provided a response for 14 primed critical trials, 28 prime trials and 56 filler trials (non-filler-to-filler ratio, 1:1); this was motivated by a desire to reduce the length of the experiment, both to avoid participants losing focus and to avoid prime trials being repeated (i.e., 42 critical trials requires 84 primes, which would require each experimental item, e.g., 'dress', featuring once as a critical trial and twice as a prime per experiment).

Consequently, participants gave responses to 98 trials. To ensure that the individual critical trial items only appeared as a critical trial in any given experiment and that the individual prime trials were not repeated, participants were randomly assigned to one of three trial blocks. Each trial block comprised 14 unique critical trials, 28 unique prime trials, and 56 filler trials presented in a *filler* → *filler* → *filler* → *prime #1* → *prime #2* → *critical trial* order in which the constituents were completely randomised. Note, due to the limitations of

the experiment builder's randomisation options, the individual experimental items that appeared alongside one another were not completely randomised, however. In each trial block, 14 of the 42 experimental items appeared as critical trials, while the remaining 28 experimental items appeared as prime trials, which were completely randomised. To avoid items repeating within the same trial, any given experimental item was only primed by 28 of the 42 experimental items, i.e., all the experimental items bar those presented as critical trials in the same experiment. The critical trials involved in each trial block are presented in the 9.2

4.6.3. Materials

4.6.3.1. Critical Trials

Experiment 2a's critical trials were identical to those of Experiment 1a, consisting of the following component: 1) a caption reading “**Dr Smith: ‘Can you see a [item label]?’**” which contained one of the 42 item labels presented e.g., ‘t-shirt’, 2) a second caption, reading ‘**Alan: ‘Yes, I [can] see [...]**’ that contained a prolix description 3) the marked and unmarked images representative of the mentioned object and 4) static images of both ‘Alan’ (a man in a black shirt wearing a VR headset) and ‘Dr Smith’ (a man dressed in a lab coat).

4.6.3.2. Prime Trials

The prime trials were identical to the critical trials used in Experiments 1a and 2a. The prime trials were modelled after the prime trials used in Bott and Chemla (2016), wherein participants were given a free choice between selecting an image that corresponded to the enriched meaning of the caption (in Experiment 2a's case, the MI) or an image that matched semantically, but not pragmatically. As with the critical trials in Experiment 1a, the prime trials were designed to elicit MI. However, as the prime trials immediately preceded a MI critical trial, they were designed to prime the derivation of MI in the succeeding critical trial. This was based on the assumption that the derivation of MI in the prime trials would lead to an increase in the activation of some level of the mechanism responsible for MI derivation that would remain residually activated in the succeeding critical trials. In turn, this would increase the rate of MI in the critical trials – this was based on the same reasoning presented in Rees and Bott (2016), Rees et al. (2023) and Meyer and Feiman (2021).

However, as demonstrated by the results of Experiment 1a, MI appears to not be as available as other types of implicature, e.g., *some* QI in Marty et al. (2022). Consequently, the efficacy of using MI primes in this type of paradigm is unclear. Given that MI enrichment seems to occur in ~ 15% of the experimental items, then the likelihood of two successive prime trials involving MI priming the succeeding MI is low. The efficacy of this particular priming methodology will be discussed further in 5.1.9. The prime-to-critical trial progression is depicted below:

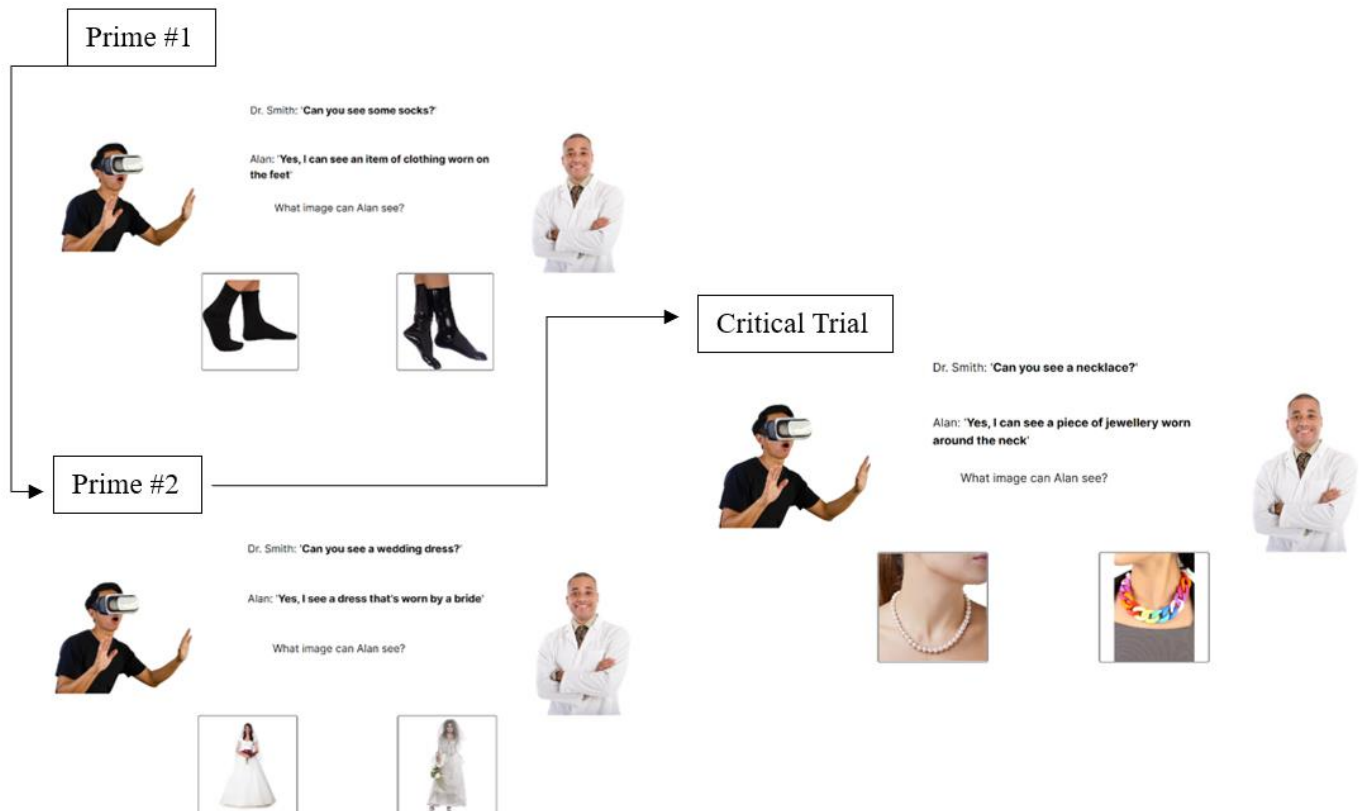


Figure 4-13: the trial progression of Experiment 2a

4.6.3.3. Filler Trials

Experiment 2a used the same fillers as Experiment 1a, the 'Yes', No-Marked-Unmarked and No-False-False filler trials. The *filler* → *filler* → *filler* → *prime #1* → *prime #2* → *critical trial* blocks in Experiment 2a contained an additional filler to those of Experiment 1a. However, owing to the addition of the prime trials, the overall proportion of filler trials to non-filler trials decreased. The additional filler trial was added to ensure that the majority of the utterances in Experiment 2a did not involve potential MI. Nevertheless, I decided not to increase the amount of filler trials beyond the 1:1 ratio in order to avoid longer trial blocks.

4.6.4. Procedure

The task experiment was hosted on Gorilla Experiment Builder (www.gorilla.sc) and accessed by participants via Prolific (www.prolific.com) [2024]. After the collection of consent and demographic data, the participants were shown the following instruction screen (the same as that of Experiment 1).

Instruction Screen 1: *“Welcome to the task! In this task you will be given sets of two images and a caption. You will be tasked with matching the caption to one of the images - you will do this by using by clicking on the image”*

Instruction Screen 2: *“In this experiment, you will be presented with a fictional scenario involving Alan and a researcher. Alan is participating in a visual recognition experiment and has been tasked with telling the researcher what image he can see in the VR headset. The researcher will show Alan two images in the VR headset, one will then disappear. Alan is tasked with communicating which image stays, i.e., which image he can still see. Try not to think too hard about your answer!”*

4.6.5. Research Hypotheses & Predictions

At the most basic level, I expect there to be a priming effect between the MI primes and the MI critical trials, assuming that MI can be primed like other types of EVA (e.g., Bott & Chemla, a.o.) Because Experiment 2a involves within-category priming, the existence of a priming effect is largely dependent on whether or not MI can be primed rather than on a specific theoretical conceptualisation of MI and its relationship to other implicatures. That said, if MI can be primed, then any priming effect is contingent on whether or not the critical trials of Experiment 1a involve MI; because of this, the existence of a priming effect can provide post-hoc evidence that the critical trials in Experiment 1a involve MI and that all MI is derived by a unified derivation mechanism; otherwise there is no theoretical reason why they should prime one another, assuming priming is indicative of mechanistic activation.

Nevertheless, the priming trials in Bott and Chemla (2016) involved primes in which the potential implicatures were more available (e.g., *some* QI). In contrast, there is no reason to expect that the priming trials in Experiment 2a will elicit a higher degree of MI than the critical trials of Experiment 1a. Consequently, I predict that any priming effect will be small,

given that the majority of the responses to the critical trials of Experiment 1a will not be preceded by two prime trials in which MI is derived. In this sense, a significant increase in the rate of MI beyond that of ~15% will likely be driven by a minority of critical trials in which priming has been successful. In this sense, I predict there will be a larger priming effect in critical trials immediately preceded by two ‘correct’ prime trials and no priming effect in those not preceded by incorrect prime trials if spill-over effects are not present.

Considering this, I make the following predictions for Experiment 2a:

- I predict the rate of MI to be higher in Experiment 2a than that of Experiment 1a if priming occurs.
- I expect a priming-based increase between Experiment 1a and Experiment 2a to be small.
- Any priming effect should be localized to the critical trials immediately succeeding the ‘correctly’ answered prime trials.

4.6.6. Coding

The selection of the marked image was interpreted as indicative of MI derivation in Experiment 2a’s critical trials, and the prime trials and responses were coded with the numerical score 1 (i.e., ‘correct’) if they selected the unmarked image. The responses to the filler trials were coded in the same way as those of Experiment 1a.

4.6.7. Results & Analysis

The rate of correct image selection was > 95% for each participant across the ‘Yes’ filler trials (i.e., those with an objectively ‘correct’ answer). Therefore, none of the participants were excluded from the analysis. The rate of marked image selection was 21% across all of the MI critical trials and 21% across the MI prime trials. For the filler trials, the rate of correct image selection was 94% for the ‘Yes’ fillers, 54% for the No-Marked-Unmarked filler trials and 49% for the No-False-False filler trials. There was an increase in the rate of MI derivation in the critical trials of Experiment 2a when compared to the baseline established in Experiment 1a (15% → 21%).

The mean rate of derivation in the primed trials varied between items ($r=0.51$, $sd=0.13$). The lowest rate of MI derivation was observed in ‘wedding dress’ at 0% (baseline = 5%) and the

highest in 'cowboy boot' at 52% (baseline = 48%). The difference between the baseline and primed rates of MI varied by item ($r=0.36$, $sd= 0.06$), with the smallest 'priming' effect (or largest 'negative priming' effect) observed in the item 'goggles' at -10% and the largest in 'skirt' at 26%.

4.6.7.1. *Primed Subset*

Although it is customary to remove critical trials that are not preceded by correctly answered primes (e.g., Rees & Bott, 2018; Bott & Chemla, 2016; Raffray & Pickering, 2010, a.o.), this is not appropriate for the dataset of the current study – as mentioned, the rates of MI derivation are low, such that of Experiment 2a's 1007 critical trial responses, only 83 were immediately preceded by two 'correctly' answered prime trials (i.e., trials in which the marked image was selected). Consequently, an analysis using a subset of the data including only the primed critical trials is statistically underpowered. However, it is perhaps worth reporting that in comparison to the overall rate of MI derivation (21%), the mean rate of MI derivation for the critical trials immediately preceded by two correct prime trials was 54%.

Comparatively, if only critical trials that are preceded by one 'correct' prime are considered, the mean rate of MI derivation is 28%, which accounts for 730 of the 1007 observations. In contrast, in the 'unprimed subset' (i.e., the critical trials immediately preceded by two incorrect primes), the rate of MI derivation was 11%, which suggests that the observed increase in the rate of MI between the critical trials of Experiment 1a and 2a is driven by a subset of critical trials involving 'correctly' answered primes.

A consequence of this decision – or, perhaps, of the paradigm – is that the priming effect observed is likely far smaller than the 'true' magnitude of any priming effect, considering that the vast majority of the trials are not 'primed'.

4.6.7.2. *The effect of the primed condition*

To assess whether there was a significant increase in the likelihood of MI derivation in the primed condition, I fitted a maximal generalized, linear mixed-model regression model with a binomial distribution and a logit link function (i.e., a logistic regression) with the condition (Primed vs. Baseline) as a fixed effect and random intercepts for experimental item and individual participant. The model indicated a significant main effect of the experimental

condition ($\beta = 0.66$, $p = 0.01$), suggesting that the primed condition increases the likelihood of MI derivation (i.e., selection of the marked image).

The described logistic mixed-effects regression model (AIC = 2594.0) provides a better fit to the data than a logistic fixed-effects model without random intercepts (AIC = 3120), indicating that the inclusion of random effects for items and participants improves the model's ability to account for variability in the rate of MI derivation.

I used the same maximal model fitted with random intercepts for the experimental item and individual participants on a subset of data, the unprimed subset, i.e., on the critical trials not immediately preceded by two 'correctly' answered prime trials. The maximal model would not converge (perhaps due to low variability in the random effects). Therefore, separate models with by-item and by-subject random effects were fitted. The model fitted with the experimental item as a random effect was insignificant ($\beta = -0.08$, $p = 0.5$), as was the model fitted with individual participants as a random effect ($\beta = 0.15$, $p = 0.56$), indicating that the significance of the effect observed in the model fitted to the entire data set is likely driven by the primed subset, which comprises 28% of the overall data.

4.6.7.3. *Item Analysis*

As mentioned, the priming effect is not uniform between experimental items. To investigate whether the size of the priming effect varied between items based on the availability of MI derivation, I performed a linear regression to predict the difference between the baseline and primed rate of MI derivation (i.e., the magnitude of any priming effect) based on the baseline rate of MI derivation by item. The model was insignificant ($\beta = 0.08$, $p = 0.6$), suggesting that an increased rate of MI availability in the baseline condition did not confer a more significant priming effect in the primed condition. Below, Figure 4-14 illustrates the variability between the baseline rates of MI derivation and the rates of MI in Experiment 1b's primed condition. This either means that priming is not contingent on whether or not MI is more or less available or that perhaps the rates of MI derivation in the baseline condition are not representative of priming.

4.6.7.4. 'Primeable' Subset

Assuming that an increase in MI derivation in Experiment 2a indicates a priming effect, Figure 4-14 suggests that priming does not occur across all experimental items. Taking a 5% increase in MI derivation between the baseline and primed conditions as an arbitrary threshold for priming, a priming effect is observed in the following 21 of the overall 42 items in Experiment 2a¹³. That said, it is unclear whether these items are 'primeable' because they elicit MI, or because they are primed by particular primes. Note that although the primes are randomised, any priming effect is contingent on the prime trials, and the critical trials elicit MI. Consequently, the degree to which the increase in the below items indicates that they are uniquely 'primeable' is unclear. Nevertheless, there is some indication that the items below are more susceptible to the priming effect than the other items.

Bathing Suit	Vest	Apron	Yoga Pants
Handbag	Glasses	Hat	Wallet
Socks	Cap	Skirt	Bathrobe
Tights	Swimming Trunks	Belt	Headband
T-Shirt	Headphones	Bowtie	Necklace

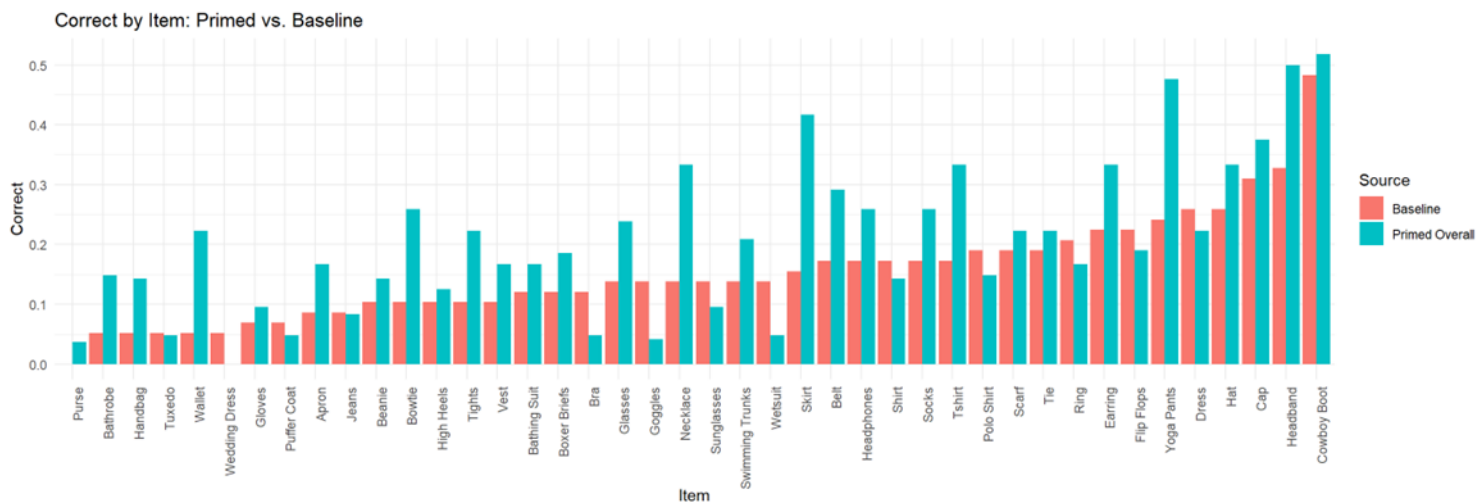


Figure 4-14: The mean rates of MI derivation by item in Experiment 1a and 2a ordered by the baseline

¹³ The threshold is calculated by the degree of increase in marked image selection between the baseline and primed conditions. As such, this calculation disregards the fact that 'priming' should only occur after preceding MI. So, the 'primeable' subset of Experiment 2a's data is a subset comprising of trials involving experimental items that show a susceptibility to priming rather than those that have been primed. Unfortunately, a data subset comprising only trials involving items that show an overall priming effect in those preceded by two correctly answered prime trials would be too small for use in any analysis.

4.6.8. Discussion

4.6.8.1. Overall Effect of The Prime

I interpret the significant increase in the likelihood of MI derivation in Experiment 2a compared to the baseline as suggestive of a priming effect in the MI critical trials. Considering the only difference between the trial blocks in Experiment 1a and Experiment 2a is the inclusion of the prime trials and that other types of conversational implicature can be primed in the same way, e.g., Bott and Chemla (2016), a.o., the existence of a priming effect is the most likely explanation for the increase in the rate of MI derivation between the baseline condition and the primed condition.

While including the prime trials does alter the ratio of trials containing MI, which can be considered another difference between the baseline and the primed condition, it renders the prolix captions relatively more frequent. Consequently, this would likely have the opposite effect on MI derivation, making it less likely, as the prolix captions may be considered less marked. Therefore, this difference is unlikely to account for the priming effect observed.

Given that there is no significant increase or decrease in the likelihood of MI derivation between the baseline and the unprimed subset data, the priming effect seen in Experiment 2a is unlikely to be a product of the paradigm more generally. Moreover, the localisation of the priming effect to the primed trials suggests that it hinges on the preceding derivation of MI.

As with the rate of MI derivation in the baseline, the extent of the priming effect in Experiment 2a is unclear. When considered at the broadest level, the data suggests that the priming effect is small, with the priming effect conferring an approximately 5-to-6 per cent increase in overall MI derivation rates. As discussed, the relative unavailability of MI in the prime trials predicts a small priming effect, given that the majority of the priming trials are unlikely to involve MI. This assertion is supported by the data, which suggests that two ‘correct’ priming trials precede a small minority of the critical trials (< 3%). However, the MI derivation rate in critical trials immediately preceded by either one or two correct primes, 54% and 28%, respectively, is far higher than in the critical trials immediately preceded by two ‘incorrect’ prime trials, 11%. This pattern is illustrated below in Figure 4-15. Moreover, while the increase between the baseline condition and the primed condition was statistically significant for the dataset as a whole, the increase was not significant for the unprimed subset, which suggests that the priming effect observed in Experiment 2a is driven by the

critical trials that are preceded by correct prime trials, i.e., are primed, regardless of the item. As such, the magnitude of the priming effect in Experiment 2a is likely far higher than the raw increase between the baseline rate of MI derivation and that of Experiment 2a, but only for a subset of the data.

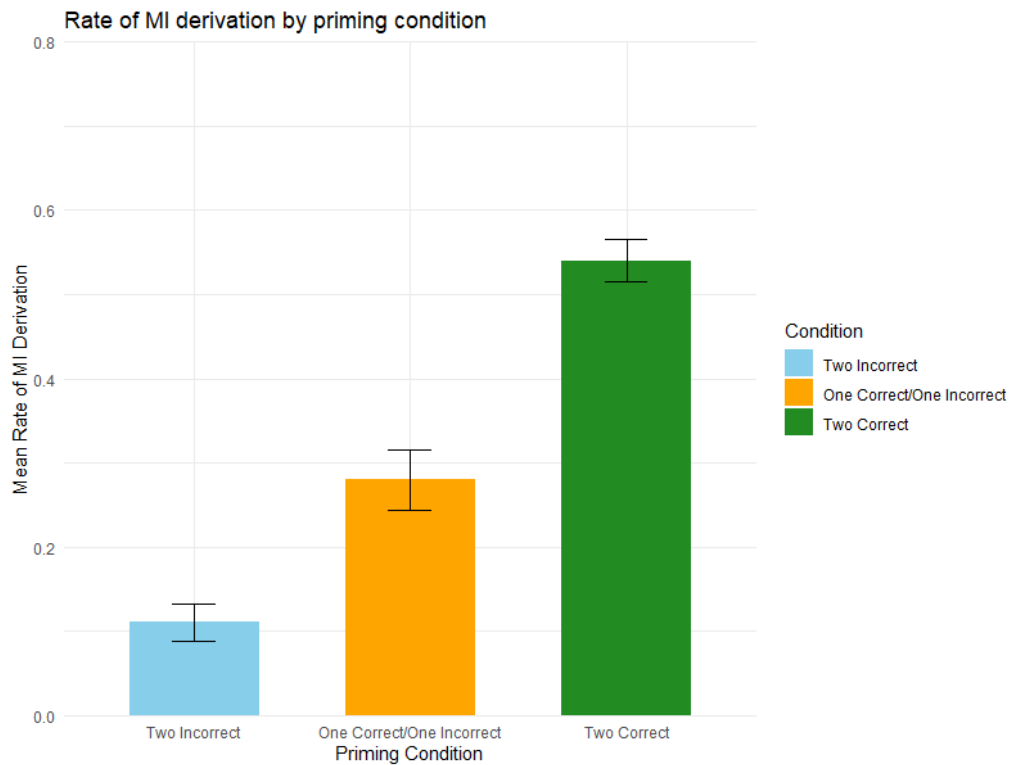


Figure 4-15: bar chart depicting the rate of MI derivation by 'primed' status in Experiment 2b with error bars representing the standard error of the by-item mean rate of MI derivation.

4.6.8.2. Nature of the Priming Effect

The observation that the priming effect is driven by critical trials that are directly primed intuitively supports an activation-based perspective of priming whereby the residual activation of a representation, structure or mechanism facilitates its re-accessing in subsequent trials. Here, I interpret the findings to suggest that a mechanism involved in MI derivation is activated in the prime trials and remains activated in the critical trials, leading to an increase in the rate of MI derivation. Considering the critical trials are primed by randomised and, crucially, different experimental items (i.e., other ad-hoc MIs), the priming effect must act on a generic mechanism responsible for deriving either MI or implicature more generally, rather than acting on a particular MI specifically. This perspective is

supported by the observation that the priming effect increases in line with the critical trials' proximity to MI derivation – Figure 4-15

The localisation of the priming effect observed in Experiment 2a suggests that the effect is not a product of the modulation of the QUD, speaker knowledge or semantic and socio-pragmatic context, among other higher-order influences on implicature availability.

Considering that these factors operate at the discourse level rather than the utterance level, they are unlikely to vary across the individual trials in the experimental paradigm. Moreover, the experimental trials depict almost identical interactions, such that factors like the QUD are likely uniform between them. Consequently, if priming were to occur via the modulation of discourse-level factors, we should observe a more uniform, non-localised priming effect, yet the critical trials not immediately preceded by correct prime trials show no priming effect. Furthermore, the increase in the relative frequency of prolix utterances in Experiment 2a when compared to the baseline condition suggests that, if priming were an effect of the modulation of contextual factors, the rate of MI derivation should decrease in the primed condition, as the prolix captions become relatively less marked and MI relatively less warranted.

4.6.8.3. *Locus of Residual Activation*

Between the two main perspectives of implicature derivation, the salience model and the combination model, there are contrasting points where an activation-based priming effect could occur. While implicature derivation relies on the sufficient activation of a usage mechanism, in the salience model, this mechanism is automatically activated, providing that the alternative is sufficiently salient. In contrast, in the combination model, the usage mechanism has a unique threshold for use and is an independent level of activation (Rees & Bott, 2018). Practically, if the derivation of MI is primed under the salience model, then priming must occur via the alternative. In contrast, the combination model's usage mechanism can be accessed or primed directly.

It is unclear how the prime trials of Experiment 2a would raise the salience of the alternative beyond the baseline level for each critical trial item; in the MI critical trials, the alternative is explicitly mentioned in the 'Dr Smith' captions and is represented by the unmarked image. Moreover, the alternative differs between the prime trials and between the prime trials and the critical trials (e.g., the alternative may be 'dress' in one and 'shirt' in the other). While the

saliency of the alternative can be increased or suppressed by other pragmatic factors in the saliency model (e.g., QUDs, speaker knowledge, semantic or socio-pragmatic context, whether the alternative has been mentioned), it is unclear how these factors would be affected by the inclusion of the prime trials if comparing Experiment 1a to Experiment 2a, such that they increase the saliency of the alternative in the critical trials.

However, the nature of priming-via-the-alternative is not fully defined in the literature that seeks to investigate it, e.g., Rees and Bott (2018), Rees, Carter and Bott (2023), and Bott and Frisson (2022). Subsequently, it is unclear to what degree priming-via-the-alternative is thought to be the product of the residual activation of the alternative as a representation, e.g., a specific lexical or scale-based representation or a component of an abstract meaning-based representation, see Bott and Chema (2016), or the mechanism responsible for the search for an alternative. Each perspective makes different predictions for whether heterogeneous MI can be primed via the alternative. For instance, the residual activation of a specific lexical representation (e.g., 'skirt') is unlikely to prime subsequent MI that involves a different alternative in the way theoretically possible for lexical scalars (e.g., if *some* QI derivation is primed via the representation of 'all'). However, if the representation is abstract and meaning-based, e.g., like Rett's (2014) formal account of MI and MI alternatives, then it may be generic enough to facilitate MI-to-MI priming. Similarly, if priming can occur via the residual activation of the mechanism responsible for finding the alternative, then this, too, could explain MI-to-MI priming.

In contrast, the perspective that priming occurs via the usage mechanism is less problematic. As per the combination model, the usage mechanism has an individual level of activation and its own threshold for use. If the priming effect is the result of residual activation, the prime trials in which MI is derived will activate the usage mechanism such that it remains activated in the succeeding critical trials, leading to an increase in the likelihood of MI derivation. Unlike if priming were to occur via the alternative, priming via the usage mechanism does not require an explanation for how the prime trials facilitate the priming of the critical trials beyond the fact that they elicit MI derivation and, therefore, activate the usage mechanism. Although the experimental data is compatible with either the priming of the alternative or the usage-based mechanism, the latter is my preferred explanation.

4.6.8.4. Conclusion

To conclude, I take the increase in the rate of MI between the baseline and Experiment 1b to indicate that the derivation of MI is primed in the latter. Given that the priming effect increases depending on the proportion of 'correctly' answered primes it succeeds (see Figure 4-15), the priming effect is the result of the prime trials. Considering this pattern, the priming effect is unlikely to be the product of discourse-level priming or the product of the conversational context (e.g., QUD, assumptions of speaker knowledge, etc.)

Tentatively, the priming effect results from the residual activation of the usage mechanism responsible for implicature derivation. However, I do not explicitly rule out other perspectives, like priming via the alternative. Although the priming effect is unlikely to be driven by the activation of a specific representation, priming via an abstract representation or the mechanism responsible for the search for the alternative are both viable explanations. Although the findings of Experiment 2c do not shed much light on the nature of the primed mechanism, the between-category priming studies (Experiments 3 – 6) will help complete this picture.

As discussed, MI is a heterogeneous phenomenon, and this is reflected in the variation in MI availability observed in Experiment 1a. Despite the uncertainty concerning whether MI is derived in the experiments' critical trials, evidence of MI-to-MI priming suggests it is. Moreover, the within-category priming effect suggests that a shared mechanism processes different types of ad-hoc MI. However, it cannot be determined whether this mechanism is MI-specific or generic to all implicature. Nevertheless, the findings provide post-hoc evidence that the critical trials of Experiments 1a and 2c elicit a type of conversational inference governed by a shared mechanism. As Experiment 2a involved only MI-to-MI priming, the findings are compatible with a broad range of accounts of conversational implicature, including the Gricean-inspired perspective. That said, Experiment 2a offers evidence that MI can be structurally primed, a novel observation at the point of writing.

However, a significant issue with the current implementation of MI priming is that MI is not derived in the majority of the prime trials, and therefore, the prime trials are not *primes* in the strictest sense. Nevertheless, the current implementation was a necessary first step in constructing the novel MI priming methodology developed in this thesis. Given that the prior implicature priming literature is established on more uniform phenomena, the direct

application to MI is unsurprisingly problematic. However, Experiments 2b and 2c seek to address these methodological hurdles.

4.7. Experiment 2b – Manner-to-Manner priming (with training block)

While Experiment 2a established that MI-to-MI priming is possible, the extent of the potential priming effect, both generally and within individual experimental items, is unclear. As discussed, the majority of the critical trials in Experiment 2b were not primed, i.e., preceded by ‘correct’ prime trials. Consequently, Experiment 2a sought to increase the rate of ‘correctly’ answered prime trials to establish a more accurate picture of the priming effect elicited in Experiment 2a. To this end, Experiment 2b involved the participants completing a pretrial ‘training block’. The pretrial training block involved a series of critical trials similar to those of the previous experiments in which the participants are directed towards the ‘correct’, i.e., pragmatic, answer.

4.7.1. Motivations

As discussed, the majority of the prime trials in Experiment 2a were not answered ‘correctly’. Consequently, the critical trials were not sufficiently primed, and the extent of any priming effect could not be accurately observed. The training block aims to increase the participants’ metalinguistic awareness and highlight the possibility of deriving MI in the prime trials. In this sense, the intervention does not seek to direct participants towards a particular answer in the prime trials but to raise the overall salience of potential MI in order to increase the efficacy of the prime trials and the degree of MI derivation in the critical trials generally. The design of the training block was based on evidence that QUD and relevance can influence the availability of implicature (e.g., Ronai & Xiang, 2021; Pankratz & van Tiel, 2021; Vicente, 2024, a.o.) and the training block was presented in tandem with an updated set of instructions. Together, both components were designed to increase the salience of potential MI in both the prime and critical trials.

4.7.2. Participants

Participants (N = 150) were recruited from Prolific (www.prolific.com) [2024] on the basis that they were adults (≥ 18 years old) and self-identified as English speakers and had not previously participated in any of the study's previous experiments. Incomplete data was rejected from N = 4 of the participants, therefore data was collected from N = 146 participants.

4.7.3. Design

The content of Experiment 2b was almost identical to that of Experiment 2a. Experiment 2b involved 42 different critical trials and 42 different prime trials, corresponding to the 42 marked-unmarked image pairs and 42 different filler trials. Like Experiment 2a, Experiment 2b had a between-participants structure; therefore, each participant provided a response for 14 primed critical trials, 28 prime trials and 56 filler trials (non-filler-to-filler ratio, 1:1) and therefore gave responses to a total of 98 experimental trials. Participants were randomly assigned to one of three trial blocks, each comprising 14 unique critical trials, 28 unique prime trials, and 56 filler trials presented in a filler \rightarrow filler \rightarrow filler \rightarrow prime #1 \rightarrow prime #2 \rightarrow critical trial order in which the constituents were completely randomized, like Experiment 2a.

Unlike the previous trials, Experiment 2b involved a training block that participants were required to complete before completing the main experiment, which was identical to that of Experiment 2a. The training block consisted of 4 guided critical trials in which the correct answer (i.e., the marked image) was highlighted in green in order to direct participants to this choice. The training block was designed to increase the salience of the possibility of MI derivation in the critical trials of the main experimental block, which was unguided. The four items used were the same for all participants and were presented in a random order. Moreover, the experimental instructions were adapted slightly to accommodate for the differences between Experiment 2a and Experiment 2b.

4.7.4. Materials

4.7.4.1. Critical Trials

Experiment 2a's critical trials were identical to those of Experiment 2a and Experiment 1a, consisting of the following component: 1) a caption reading “**Dr Smith: ‘Can you see a [item label]?’**” which contained one of the 42 item labels presented e.g., ‘t-shirt’, 2) a second caption, reading ‘**Alan: ‘Yes, I [/can] see [...]**’ that contained a prolix description 3) the marked and unmarked images representative of the mentioned object and 4) static images of both ‘Alan’ (a man in a black shirt wearing a VR headset) and ‘Dr Smith’ (a man dressed in a lab coat). Crucially, the critical trials allow the free choice between the marked and the unmarked image when responding to the ‘Dr Smith’ caption. See the critical trial for the item ‘tuxedo’ below:

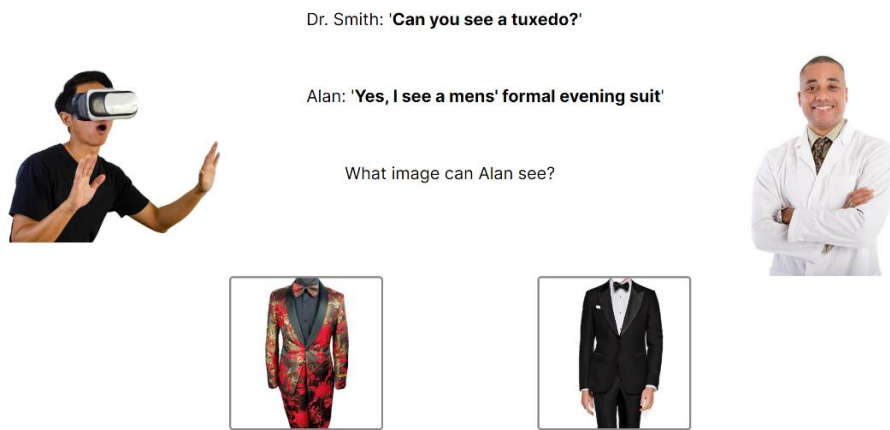


Figure 4-16: MI critical trial for the item ‘tuxedo’

4.7.4.2. Prime Trials

The prime trials were identical to the critical trials and were constructed in the same way as those of Experiment 2a.

4.7.4.3. Fillers

Experiment 2b used the same fillers as Experiment 2a, the ‘Yes’, No-Marked-Unmarked and No-False-False filler trials.

4.7.4.4. Training block

The configuration of the four critical trials in the training block was the same as that of the critical trials used in the main body of the experiment, except that the critical trials did not involve articles of clothing. Instead, the items ‘chair’, ‘dog’, ‘cat’ and ‘spaghetti’ were used. The items were adaptations of the Marked-Unmarked-False filler trials but instead paired an affirmative response from the ‘Alan’ character and a prolix caption based on the English-language Google dictionary definition of the item (Oxford Languages, 2023) with the marked and unmarked representations of the item label (N.B., the items were not screened for typicality like the main critical trials and were instead selected using intuition). The training block critical trial for the item ‘chair’ is shown in Figure 4-17.

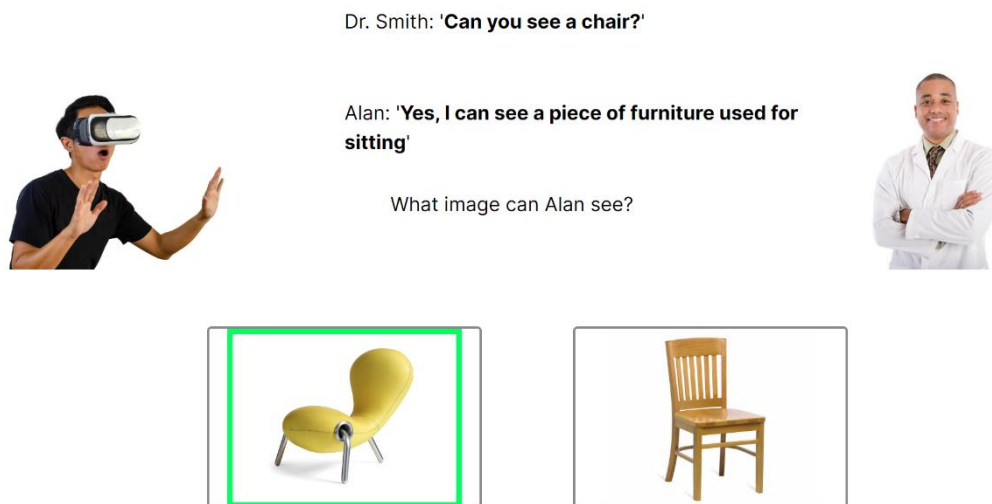


Figure 4-17: the critical trial for the item ‘chair’ in the training block

4.7.4.5. Procedure

The task was hosted on Gorilla Experiment Builder (www.gorilla.sc) and accessed by participants through Prolific (www.prolific.com) [2024]. After obtaining consent and

collecting demographic information, participants were presented with the instructions (1 and 2) listed below, an adapted version of that presented in previous experiments. After reading the instruction screens, participants completed four critical trials in the training block, after which they were presented with a third instruction screen and then the main experiment.

Instruction Screen 1: *“In this experiment, you will be presented with a fictional scenario involving Alan and a researcher. Alan is participating in a visual recognition experiment and has been tasked with telling the researcher what image he can see in the VR headset. The researcher will show Alan two images in the VR headset, one will then disappear. Alan is tasked with communicating which image stays, i.e., which image he can still see. You will be tasked with selecting the image that Alan can see. N.B **In some trials, the answer will be given for you indicated by a green outline around the object that Alan can see. In others, the correct answer is not explicitly indicated. Try not to think too hard about your answer!**”*

Instruction Screen 2: *“The next few trials will get you used to the way in which Alan is communicating. Here, you will see how Alan varies his use of language to indicated which image he sees. The image that Alan sees will be highlighted by a green square. Note: it is **very important** that you select the correct image in the trials in which the answer is given to you **as indicated by the green border**. Try to think about the way in which Alan is using language to communicate his answer”*

Instruction Screen 3: *“Welcome to the main task! In this task you will be given sets of two images and a caption. You will be tasked with matching an image to the caption. Try not to think too hard about your answer!”*

Here, the instructions sought to encourage the participants to engage ToM and to interpret the communicative motivations behind the ‘Alan’ character’s use of language to increase the salience of potential MI derivation.

4.7.4.6. Predictions

Although there is more precedence with which to estimate the effect of the pre-experimental training block, I make the following tentative predictions:

- If the rate of MI derivation increases in the prime trials, an increase in the rate of implicature derivation in the critical trials is expected.

- The trial block may prime the critical trials independently rather than via a higher proportion of ‘correctly’ answered prime trials.
- If the low rate of MI in the experimental trials is not caused by a lack of metalinguistic awareness but rather by other factors, then no increase is expected.

4.7.4.7. *Coding*

The selection of the marked image was interpreted as indicative of MI derivation in Experiment 2b’s critical and prime trials, and responses were coded with the numerical score 1 (i.e., ‘correct’) if they selected the marked image and filler trials were coded in the same way as previous experiments.

4.7.5. *Results & Analysis*

The rate of correct image selection was > 95% for each participant across the ‘Yes’ filler trials. Therefore, no participants were excluded from the analysis. Across all MI critical trials, the rate of marked image selection was 23%, while the MI prime trials showed a selection rate of 23%. The correct image selection rate for filler trials was 95% for ‘Yes’ fillers, 61% for No-Marked-Unmarked fillers, and 48% for No-False fillers. There was a notable increase in the rate of MI derivation in the critical trials of Experiment 2b, rising from the baseline established in Experiment 1a (15% → 23%) and a smaller increase when compared to the primed trials in Experiment 2a (21% → 23%).

The mean MI derivation rate in the primed trials varied between items ($r=0.43$, $sd=0.1$). The lowest rate of MI derivation was observed in ‘wedding dress’ at 6% (baseline = 5%) and the highest in ‘cowboy boot’ at 52% (baseline = 49%). The difference between the baseline and primed rates of MI varied by item ($r=0.3$, $sd=0.07$), with the smallest ‘priming’ effect (or largest ‘negative priming’ effect) observed in the item ‘wetsuit’ at -7% and the largest in ‘wallet’ at 22%.

4.7.5.1. *Primed Subset*

As with Experiment 2a, the rates of MI in the prime trials are relatively low, and as a result, only 287 of the 2044 critical trial responses in Experiment 2b were immediately preceded by

two ‘correctly’ answered prime trials. In comparison to the overall rate of MI derivation (23%), the mean rate of MI derivation for the critical trials immediately preceded by two correct prime trials was 78%; for those preceded by one, 30% and for those preceded by two ‘incorrect’ primes, 10%. Therefore, the observed increase in the rate of MI between the critical trials of Experiment 1a and 2a and 2b is likely driven by a subset of critical trials involving ‘correctly’ answered primes, as is the case with Experiment 2a. However, the rate of MI in the primed subset for Experiment 2b (78%) is numerically larger than that of the primed subset for Experiment 2a (54%), which suggests that any effect of the training block may be concentrated in the critical trials that are primed, rather than shared across all of the critical trials. Nevertheless, the sample sizes of the primed subsets are too small in both Experiments 2a and 2b to make statistically sound conclusions. The effect of the primes is visualised in Figure 4-18:

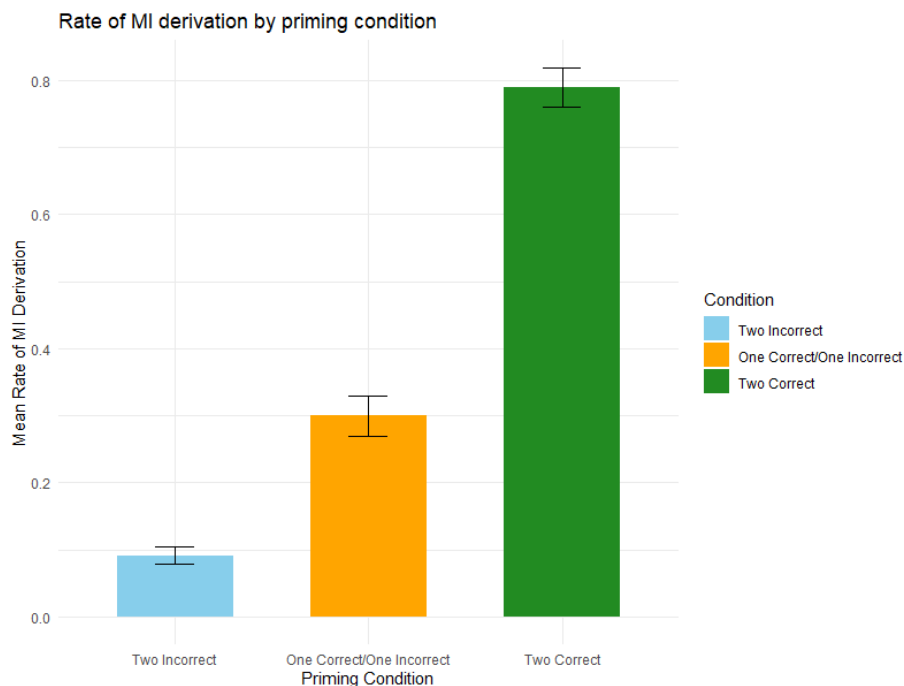


Figure 4-18: bar chart depicting the rate of MI derivation by ‘primed’ status in Experiment 2b with error bars representing the standard error of the by-item mean rate of MI derivation.

4.7.5.2. The effect of the primed condition

I fitted a generalized, linear mixed-model using a binomial distribution and a logit link function (i.e., logistic regression) to the data with the condition (Primed: *Training Block* vs.

Baseline) as a fixed effect and a random intercept and slope for the experimental item and intercept for the individual participant. However, the model failed to converge. Instead, I refitted the model with an intercept and slope for the experimental item. The overall model was significant ($\beta = -1.83$, $p = <.001$), as was the main effect of the experimental condition ($\beta = 0.57$, $p = <.001$).

4.7.5.3. *Comparison to the Experiment 2a*

To compare the effect of Experiment 2a's prime condition to that of Experiment 2b (+ *Training Block*), I fitted the same logistic regression model with an intercept and slope for the experimental item and an intercept for the participant to a dataset containing just the critical trials from Experiment 2a and 2b, without the baseline. While the overall model was significant ($\beta = -1.96$, $p = <.001$), the main effect of Experiment (i.e., 2a or 2b) was not ($\beta = 0.03$, $p = 0.9$), indicating that there is not a significant difference between the degree to which the prime conditions in Experiment 2a and 2b increase the rate of MI derivation beyond the baseline condition.

4.7.5.4. *Item Analysis*

Given the lack of significant difference between Experiment 2a and 2b, I opted not to analyse the items independently.

4.7.6. *Discussion*

4.7.6.1. *Overall priming effect*

Like in Experiment 2a, there is a significant increase ($p = <.001$) between the baseline rate of marked image selection and the rate of marked image selection in the primed condition in Experiment 2b (15% → 23%). I interpret this as an increase in the rate of MI derivation in the critical trials of Experiment 2b for the same reasons discussed in light of Experiment 2a.

Comparison to Experiment 2a

The results of the regression analysis on the critical trial data from both Experiment 2a and 2b suggest that the inclusion of the training block does not significantly increase the overall likelihood of MI derivation beyond the level of the increase afforded by the prime condition

alone. In other words, it is unlikely that the inclusion of the training block increases the magnitude of the priming effect. That said, there is a numerical increase between the rate of MI derivation in prime trials (21% → 23%) and the primed critical trials between Experiment 2a and Experiment 2b (54% → 78%). These increases may indicate an effect of the training block limited to the critical trials that are immediately preceded by two ‘correct’ prime trials. However, given that the successfully primed trials comprise a small minority of the overall critical trials, there is not enough data to explore the significance of this increase.

4.7.6.2. *Nature of the training block ‘prime’*

Theoretically, the training block was designed to increase the rate of MI derivation by increasing the metalinguistic awareness of possible MI derivation in both the prime and the critical trials on Experiment 2b based on the assumption that MI was not sufficiently available across the prime and critical trials of Experiment 2a. Nevertheless, despite the semi-explicit direction, the rate of MI derivation did not increase in Experiment 2a. This lack of increase could suggest that participants are already sufficiently aware of the general availability of MI as a possible mode of pragmatic communication within the paradigm but that the hurdle to increased rates of priming is the lack of item-specific MI availability. That is, participants may be ‘willing and able’ to derive MI, but MI is not sufficiently available in the trials for non-metalinguistic reasons, e.g., there is insufficient activation of the usage mechanism.

4.7.6.3. *Conclusion*

The inclusion of the training block does not seem to increase the metalinguistic awareness of potential MI to the extent that it significantly increases the rate of MI derivation in the critical trials of Experiment 2b beyond that of Experiment 2a. Nevertheless, the priming effect is not dampened by the inclusion of the training block, and there is some evidence of a positive effect in the numerical increase between the rates of MI derivation between Experiment 2a and 2b in both the prime trials (21%→ 23%), the critical trials (21% → 23%) and, particularly, in the ‘primed’ critical trials (54% → 78%). Although the increase in the rate of MI in the ‘primed’ trials is anecdotal, it may suggest that the training block has some type of effect, but this effect is obfuscated by the nature of the experimental paradigm.

4.8. Experiment 2c – Manner-to-Manner priming (Training Block & Guided Primes)

The results of Experiment 2b suggest that including the training block does not significantly increase the rate of ‘correctly’ answered prime trials. Consequently, the critical trials of Experiment 2b are ‘primed’ at similar rates to those in Experiment 2a, and there is no significant increase in the rate of MI derivation in the critical trials of Experiment 2b. In addition to this, the training block does not appear to have an independent effect on the critical trials. That said, the superficial increase in the rate of MI derivation in the primed trials of Experiment 2b beyond those of Experiment 2a suggests that including the training block may benefit sufficiently primed trials.

To increase the rate of ‘correctly’ answered primes, Experiment 2c investigated whether including guided primes, i.e., those with the same green border as the training block (see Figure 4-20), increases the rate of ‘correctly’ answered primes. The guided primes were used with the training block and the updated instructions.

4.8.1. Motivations

The findings of Experiment 2b indicate that increasing the metalinguistic awareness of MI does not substantially increase the relatively low rates of MI derivation observed across the prime and critical trials of Experiments 2a and 2b. As noted in Experiment 2a’s discussion, if the priming effect is the result of the residual activation of a mechanism used in implicature derivation, then increasing the activation of this mechanism in the prime trials should lead to an increase in MI derivation in the succeeding trials. Given that the level of activation appears resistant to external modulation (i.e., by increasing metalinguistic awareness), I decided to attempt to raise the activation of the usage mechanism in the prime trials directly by ‘guiding’ participants to select the image compatible with MI in the prime trials. If MI is derived at a higher frequency in the prime trials, then there should be a higher degree of residual activation in the critical trials and, therefore, a larger degree of MI derivation.

4.8.2. *Participants*

Participants (N = 150) were recruited from Prolific (www.prolific.com) [2024] based on their being adults (≥ 18 years old), self-identifying as English speakers, and having not previously participated in any of the study's previous experiments. Incomplete data was rejected from N = 5 of the participants, so data was collected from N = 145 participants.

4.8.3. *Design*

The content of Experiment 2b was almost identical to that of Experiment 2a. Experiment 2b involved 42 critical trials, corresponding to the 42 marked-unmarked image pairs, and 42 different prime trials, slightly adapted versions of the critical trials. Like Experiment 2a, Experiment 2b had a between-participants structure; therefore, each participant responded to 14 primed critical trials, 28 prime trials and 56 filler trials (non-filler-to-filler ratio, 1:1) and therefore gave responses to a total of 98 experimental trials. Participants were randomly assigned to one of three trial blocks, each comprising 14 unique critical trials, 28 unique prime trials, and 56 filler trials presented in a *filler* \rightarrow *filler* \rightarrow *filler* \rightarrow *prime #1* \rightarrow *prime #2* \rightarrow *critical trial* order in which the constituents were completely randomized, like Experiment 2a.

4.8.4. *Materials*

4.8.4.1. *Critical Trials*

Experiment 2a's critical trials were identical to those of Experiment 2a and Experiment 1a, consisting of the following component: 1) a caption reading “**Dr Smith: ‘Can you see a [item label]?’**” which contained one of the 42 item labels presented e.g., ‘t-shirt’, 2) a second caption, reading ‘Alan: **Yes, I [can] see [...]**’ and involving a prolix description 3) the marked and unmarked images representative of the mentioned object and 4) static images of both ‘Alan’ (a man in a black shirt wearing a VR headset) and ‘Dr Smith’ (a man dressed in a lab coat). Crucially, the critical trials allow the free choice between the marked and the unmarked image when responding to the ‘Dr Smith’ caption. See the critical trial for the item ‘tuxedo’ below:

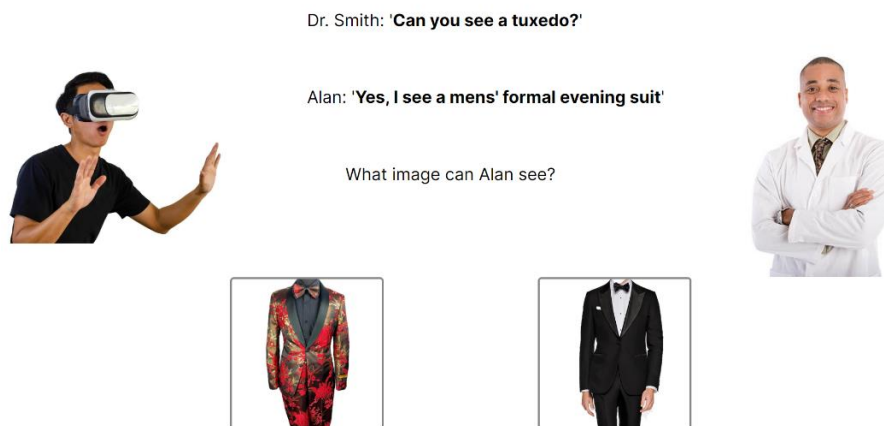


Figure 4-19: Critical trial for the item 'tuxedo' in the MI condition.

4.8.4.2. Prime Trials

The prime trials were identical to the critical trials, except that the marked image was presented with a green border. The participants were instructed that the correct answer to the 'Dr Smith' caption would be highlighted in some trials – see 5.8.5. below. The green border was designed to direct participants to the selection of the response compatible with MI derivation, with the assumption that directing participants to a response compatible with a particular reading would lead to them processing this particular reading, e.g., as was the motivation behind Bott & Chemla's (2016) *weak* prime. The prime trial for the item 'polo shirt' is depicted below:

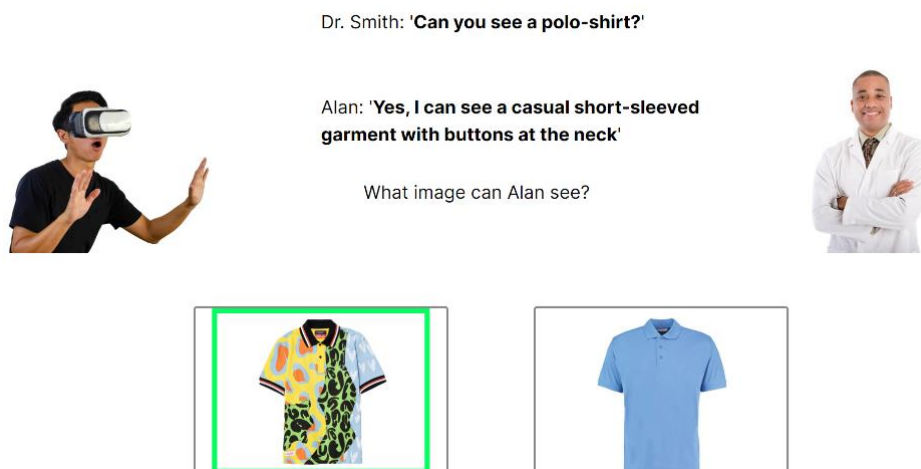


Figure 4-20: 'Guided prime' critical trial for the item 'polo shirt' in Experiment 2c.

Participants could select either image in all prime trials despite the green border. While there are other methods with which to direct the participants' choice (e.g., obscuring the unmarked image or rendering it unclickable), I wanted to ensure that the direction could be interpreted as evidence of the 'Alan' character's intentions rather than as a limitation or condition of the paradigm. As such, the green border served as a visual representation of the 'Alan' character's perspective, motivating his use of the prolix caption (i.e., to reference the marked object by eliciting MI).

4.8.4.3. *Filler Trials*

Experiment 2a used the same fillers as Experiment 2a, the 'Yes', No-Marked-Unmarked and No-False-False filler trials.

4.8.5. *Procedure*

The task was hosted on Gorilla Experiment Builder (www.gorilla.sc) and accessed by participants through Prolific (www.prolific.com) [2024]. After obtaining consent and collecting demographic information, participants were presented with the following instruction screen, which was an adapted version of that presented in previous experiments:

Instruction Screen 1: "Welcome to the task! In this task you will be given sets of two images and a caption. You will be tasked with matching the caption to one of the images - you will do this by using by clicking on the image"

*Instruction Screen 2: "In this experiment, you will be presented with a fictional scenario involving Alan and a researcher. Alan is participating in a visual recognition experiment and has been tasked with telling the researcher what image he can see in the VR headset. The researcher will show Alan two images in the VR headset, one will then disappear. Alan is tasked with communicating which image stays, i.e., which image he can still see. You will be tasked with selecting the image that Alan can see. **N.B In some trials, the answer will be given for you indicated by a green outline around the object that Alan can see.** In others, the correct answer is not explicitly indicated. Try not to think too hard about your answer!"*

Here, the instructions sought to highlight that the green outline represented Alan's perspective rather than the 'correct' choice on a paradigmatic level (although the two are functionally identical).

4.8.6. Predictions

In light of the points addressed above, I make the following predictions for the findings of Experiment 2c:

- If the guided primes are followed, I predict an increase in the rate of MI in the critical trials.
- If the guided primes are followed but do not entail the computation of MI, then an increase in the rate of MI is not expected.
- I expect the rate of MI to vary between items, as seen in Experiments 2a and 2b. However, the variation may be less extreme if the items are more uniformly primed.
- If the priming effect is a function of the exact mechanism as that of Experiments 2b and 2c, I expect a similar, localized priming effect.
- Priming via a different mechanism, such as a higher-order mechanism, likely predicts a different pattern of priming, such as a more general priming effect.

4.8.7. Coding

The selection of the marked image was interpreted to indicate that MI had been derived in Experiment 2c's critical trials, and responses were coded with the numerical score 1 (i.e., 'correct') if they selected the unmarked image and filler trials were coded in the same way as previous experiments.

4.8.8. Results & Analysis

The rate of correct image selection did not exceed 95% for 14 of the participants in the 'Yes' filler trials, and data from these participants was excluded from the overall analysis. Across all of the MI critical trials, the rate of marked image selection was 50%, while the MI prime trials showed a selection rate of 85%. The correct image selection rate for filler trials was 95% for 'Yes' fillers, 56% for No-Marked-Unmarked fillers, and 51% for No-False-False

fillers. There was a notable increase in the rate of MI derivation in the critical trials of Experiment 2b, rising from the baseline established in Experiment 1a (15% → 50%) and when compared to the primed trials in Experiment 2a (21% → 50%).

The mean rate of derivation in the primed trials varied between items ($r=0.57$, $sd=0.12$), although all items increased from their baseline levels. The lowest rate of MI derivation was observed in 'bra' at 24% (baseline = 12%) and the highest in 'cowboy boot' at 81% (baseline = 49%). The difference between the baseline and primed rates of MI varied by item ($r=0.38$, $sd=0.1$), with the smallest priming effect observed in the item 'bra' at 12% and the largest in 'wallet' at 50%.

4.8.8.1. *Primed Subset*

Unlike Experiment 2a and 2b, the rate of MI derivation in the prime trials was relatively high (85%); therefore, of the 2030 critical trial responses in Experiment 2c, 1614 were immediately preceded by two 'correctly' answered prime trials. In this subset, the rate of MI derivation was somewhat higher than that of the overall data, at 59%. The primed critical trials were distributed across participants, such that 143 participants were represented in the subset data (i.e., for 2 of the participants, all of the prime responses were 'incorrect'). On the flip side, in the excluded data (i.e., critical trials not immediately preceded by two correct primes), 140 participants were represented. In the subset of data consisting of critical trials

preceded by two incorrect primes, the rate of MI was 11%, while in those preceded by one correct prime, it was 29%.

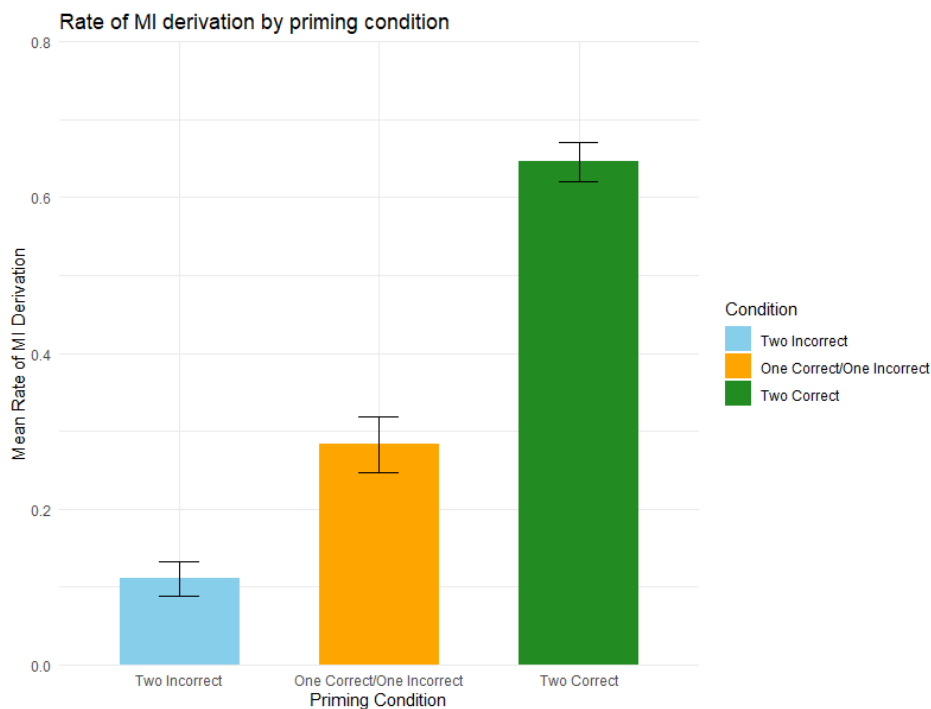


Figure 4-21: bar chart depicting the rate of MI derivation by 'primed' status, with error bars representing the mean rates of by-item MI derivation in each condition for Experiment 2c.

4.8.8.2. Overall Priming Effect (Training Block & Guided Primes)

I fitted a maximal generalized, linear mixed-model regression model with a binomial distribution and a logit link function to the dataset (i.e., logistical regression), with the condition (Primed: *Training Block & Guided primes* vs. Baseline) as a fixed effect and with a random intercept and slope for the experimental item and an intercept for the individual participant to the overall data.

The model indicates a significant main effect of the experimental condition ($\beta = 2.6$, $p = <.001$), indicating that the combination of the training block and the green border increases the likelihood of MI derivation (i.e., selection of the marked image) in the critical trials. The logistic mixed-effects regression model ($AIC = 3459.3$) with random intercepts provides a better fit to the data than that without ($\beta = 2.6$, $p = <.001$, $AIC = 4431.1$).

4.8.8.3. Compared to Experiment 2a & 2b

To compare the effect of Experiment 2a's & 2b's (+*Training Block*) prime conditions to that of Experiment 2c (+ *Training Block & Guided primes*), I fitted the same linear mixed

regression model with an intercept and slope for the experimental item and an intercept for the participants to a dataset containing just the critical trials from Experiment 2a & 2b & 2c, without inclusion of baseline data. The overall model was significant ($\beta = -1.99$, $p = <.001$), as was the effect of Experiment 2c's condition ($\beta = -2.01$, $p = <.001$), indicating that there is an increased likelihood of MI derivation in Experiment 2c's condition (+ *Training Block & Guided primes*) than in Experiment 2a's or 2b's.

4.8.9. Discussion

4.8.9.1. Overall Priming Effect

There is a large and, unsurprisingly, statistically significant increase in the rate of marked image selection in the critical trials of Experiment 2c when compared to Experiment 1a's baseline (15% → 50%). I suggest that this difference indicates an increase in the rate of MI derivation in the critical trials of Experiment 2c and that the magnitude of this increase is driven by the increased efficacy of the guided prime trials compared to those of the previous studies.

4.8.9.2. Comparison to Experiment 2a & 2b

The priming effect elicited by the inclusion of both the training block and Experiment 2c's guided prime trials is far larger than the effect elicited by the unguided prime trials (i.e., Experiment 2a) and the inclusion of the training block alone (e.g., Experiment 2b), as evidenced by the observed numerical increase in the rate of MI derivation in Experiment 2b (21% → 50%) and the significance of the logistical regression that modelled the difference in the likelihood of MI derivation between Experiment 2a, 2b and 2c. As predicted, the data suggests that Experiment 2c's condition increased the rate at which participants 'correctly' answered the prime trials, which in turn increased the overall priming effect, i.e., 85% of the prime trials were answered 'correctly'. Unfortunately, due to scant data available concerning 'primed' trials in Experiments 2a and 2b, it is unclear whether the guided or non-guided primes were more effective primes when answered 'correctly'. The mean rates of MI derivation across the baseline (Experiment 1a), Experiment 2a, 2b and 2c are presented in **Error! Reference source not found.:**

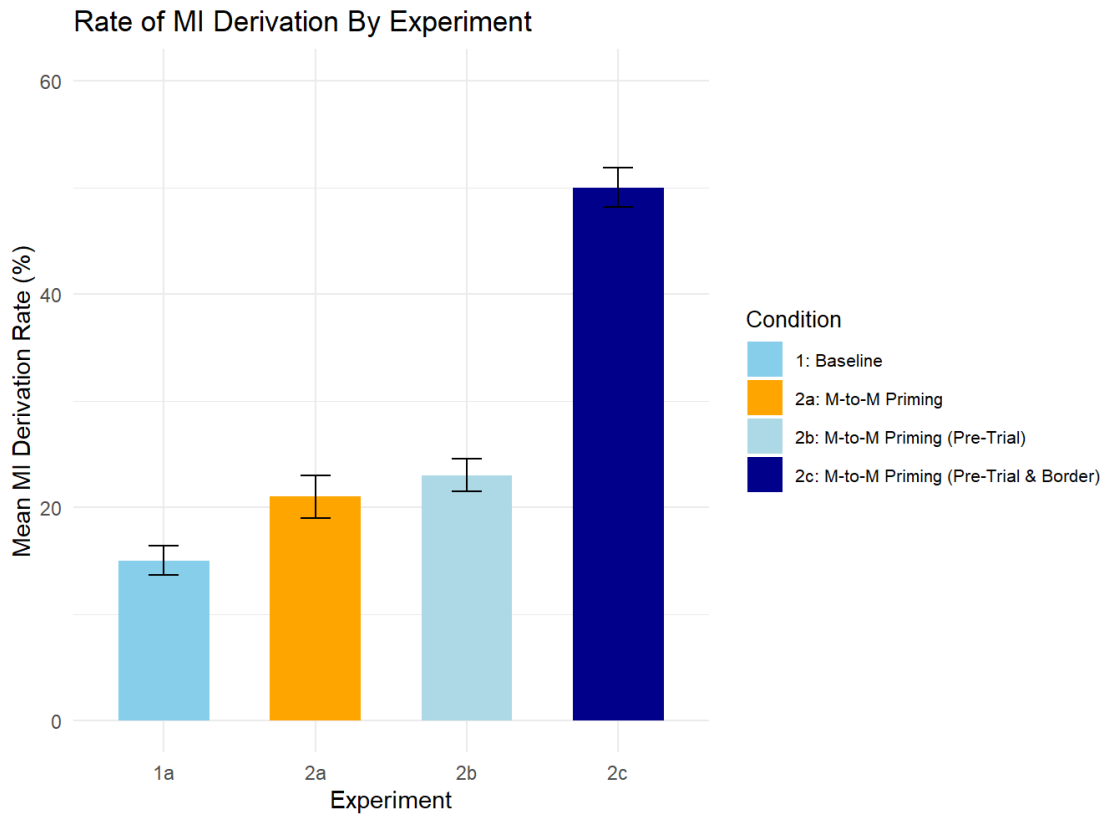


Figure 4-22: The mean rate of MI derivation across Experiment 1a and Experiments 2a, 2b, 2c and 3, and 5 with error bars representing the standard error of the by-item mean rate of MI derivation.

4.8.9.3. Nature Of The Priming Effect

Although I conclude that the increase in the rate of MI derivation in Experiment 2c is the result of priming, the nature of the priming effect may be different to that of Experiments 2a and 2b. As discussed, the most plausible explanation for the priming effect observed in Experiment 2a (and 2c) is that of residual activation, given that the priming effect is driven by the critical trials that immediately succeed two ‘correct’ primes. Moreover, there are no major contextual differences between Experiment 1a’s baseline and Experiment 2a, aside from including the prime trials, which suggests that the effect is not the result of a change in context.

In contrast, the prime trials of Experiment 2c are different to the critical trials; namely, the responses (i.e., image selection) in the prime trials are not necessarily indicative of the participant’s natural response to the trial as the responses are directed by the inclusion on the green border. Consequently, the degree to which MI is elicited in the ‘correct’ responses to

Experiment 2c's critical trials is unclear, as participants are free to simply follow the direction of the paradigm without processing the potential MI at hand. However, the significant increase in marked image selection in critical trials succeeding one or two 'correct' prime trials suggests that the participants' processing of the prime trials is more than superficial.

Moreover, the results of the linear regression that modelled the relationship between the by-item priming effects in Experiment 2a and Experiment 2c show a positive correlation between the two. This suggests that either the nature of the priming effect is the same in both experiments or that the experimental items respond similarly to different types of priming. Given this, it appears likely that participants are primed by the processing of the MI implicature in the prime trials. However, the findings are also compatible with participants' following a more basic 'select the weird image' heuristic when responding to the critical trials.

However, if participants have simply learnt that the 'correct' response is to select the marked image in the trials, then it is unclear why this strategy is arbitrarily abandoned in some prime trials and why the suspension of the heuristic bleeds into the succeeding critical trials. Even if the direction is ignored in some of the prime trials, e.g., because the potential MI is unfeasible, say, it seems unlikely that an explicitly learnt answering strategy would be applied so dynamically and locally, i.e., in certain *prime* → *prime* → *critical trial* blocks, but not others. Moreover, the unprimed critical trial responses are present in data gathered across almost all of the participants in Experiment 2c. Intuitively, if the effect were driven by implicit or explicit learning, as per Bock and Griffin's (2000) perspective of priming, we should expect the effect to be more enduring across the responses of an individual participant (Hartsuiker et al., 2008). Furthermore, in critical trials preceded by 'correct' primes, the rate of MI derivation is only 59% and varies between items, a rate far higher and far more uniform than this seems to be predicted by an answering strategy that simply involves selecting the more marked image.

It should be noted that the priming effect observed in Experiment 2c is far larger than the size of the between-category priming effects observed in Bott and Chemla (2016), Rees and Bott (2018) and Marty et al. (2022) for *some* QI, e.g., Baseline: 59% → *Strong* Prime: 68% (Marty et al., 2022). Interestingly, the size of the priming effect was strikingly similar to that of Marty et al.'s *ad-hoc* QI, 11% → 42%. The large MI-to-MI priming effect may be the result of an inverse preference effect, whereby the activation of relatively uncommon

structures, i.e., MI, confer larger priming effects due to their unexpected nature. Taken together with Marty et al.'s findings, the results suggest that ad-hoc implicature subtypes have lower baseline rates of implicature derivation but are more susceptible to priming effects – perhaps as a result of the former.

Residual Activation

Taken together, the findings of Experiments 2a, 2b and 2c are compatible with an activation-based account of priming. From this perspective, the ‘correctly’ answered prime trials can be considered to raise the level of activation of a mechanism involved in the derivation of MI, such that the mechanism remains residually activated in succeeding trials, leading to an increase in subsequent MI derivation. The localisation of Experiment 2c’s priming effect supports this explanation, and Figure 4-23 shows a linear increase in the rate of MI in the number of ‘correctly’ answered preceding primes¹⁴. Assuming that the ‘correctly’ answered primes lead to higher levels of residual activation, this pattern is predicted by an activation-based account. In contrast, in the prime trials that are answered ‘incorrectly’, activation of the mechanism may be suppressed, leading to lower-than-baseline rates of MI derivation.

As discussed, the nature of the primed mechanism is unclear. The findings of Experiment 2 are equally compatible with an account whereby priming is the result of the residual activation of the usage mechanism or of a mechanism responsible for finding an appropriate alternative. Moreover, the priming effect may not be related to a domain-specific mechanism and may be related to a generic mechanism of implicature derivation. Consequently, Experiment 2’s findings are compatible with both the *salience* and *combination* models of implicature derivation.

¹⁴ Figure 4-23 should be interpreted with caution as the amount of critical trials used to calculate each plotted mean vary drastically, such that they are not equally reliable, as described in the results of each experiment.

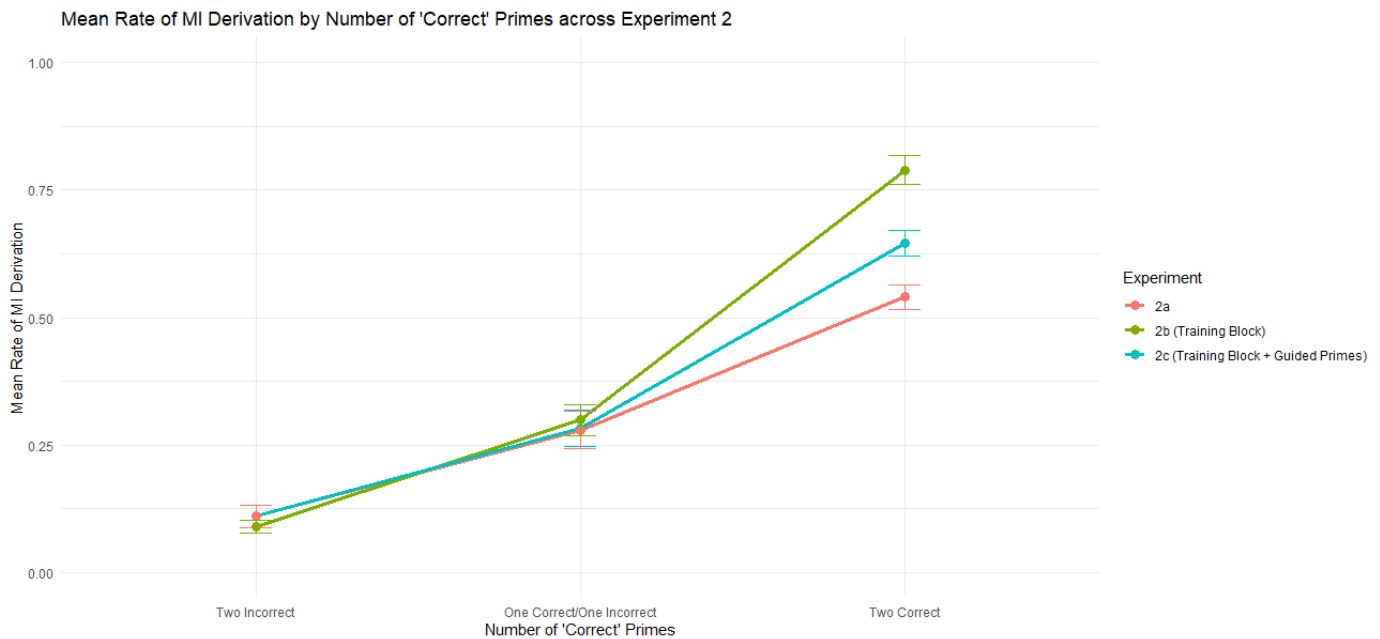


Figure 4-23: Mean rate of MI derivation by the number of correct primes across Experiments 2a, 2b, 2c, with error bars representing the standard error of the by-item mean rate of MI derivation.

Context Adaptation

While I favour an activation-based account of Experiment 2’s findings, Marty et al. (2022) account for the implicature priming in their study as a product of a mechanism of *context adaptation*. To recap, an account of *context adaptation* considers priming to be the product of a quick-cycling, context-sensitive adaptive mechanism that calculates the likelihood of implicature— among other linguistic phenomena. Consequently, priming is explained by a change in the perceived likelihood of implicature elicitation in the conversational context when implicature is derived in the prime trials. Therefore, MI, in this case, is more likely to be derived if it has been in succeeding prime trials.

Although the findings of Experiment 2 are ostensibly compatible with Marty et al.’s (2022) account, this explanation is largely dependent on the nature of the ‘Better Picture?’ critical trials. In the ‘Better Picture?’ trials used by Bott and Chemla (2016) et al., an image compatible with the semantic interpretation of the caption is shown alongside a card that reads ‘Better Picture?’ and the ‘Better Picture?’ card is selected if a participant believes there to be a pragmatically compatible, alternative image. Consequently, in a baseline condition (i.e., a block of unprimed ‘Better Picture?’ critical trials), the possibility of implicature derivation is not made explicit by the context, and participants must, therefore, rely on

context-independent prior assumptions of the likelihood of implicature derivation (e.g., whether ‘there is a star’ is an appropriate description of an image involving a star and a circle). In contrast, the prime trials depict both a semantically compatible image and a pragmatically compatible image, which activates the mechanism of *context adaptation* as the likelihood that the captions contain implicature is increased, as per Marty et al.’s (2022) account.

However, the explanation of context adaptation does not fit with the current study’s paradigm. In Experiment 1 – 2, the pragmatic image, i.e., the marked image, is depicted in the prime trials and the critical trials of both the baseline and primed conditions. In contrast, the pragmatic image is only present in the prime trials of Marty et al. (2022) and is replaced by the ‘Better Picture?’ image in the critical trials of both the baseline and primed conditions. Consequently, the difference between the perceived likelihood of MI in the baseline and primed conditions of Experiment 2 – particularly, Experiment 2a and 2b – is likely minimal. As such, *context adaptation* is unlikely to account for the priming effect observed in Experiment 2, although it is not explicitly incompatible with the findings.

Questions Under Discussion & Referential Pacts

Evidence suggests that QUD and relevance are taken into account when deriving conversational implicature (e.g., Yang, Minai & Fiorentino, 2018; Kursat & Degen, 2020; Ronai & Xiang, 2021; Pankratz & van Tiel, 2021). Moreover, Wilson and Katsos (2023) point out that MI and *referential pacts* have some degree of conceptual overlap. Considering this, it is possible that discourse-level phenomena like the QUD, relevance and referential pacts play a role in the findings of Experiment 2.

However, comparing the results of Experiment 2a and Experiment 2b suggests that modulating the salience of the QUD and the relevance of the ‘Alan’ character’s perspective does not significantly increase the rate of MI in the critical trials – although there is anecdotal evidence to suggest that it may increase the rate of MI in the ‘primed’ critical trials. The lack of a significant difference between the rate of MI derivation in Experiment 2a and 2b indicates that the main priming effect is not the product of discourse-level factors like the QUD or relevance but does not rule out the possibility that they play a complementary role in facilitating the derivation of available MI.

While the idea that Experiment 2c establishes a referential pact whereby marked captions are matched to marked images is an ostensibly attractive explanation for the observed increase in

MI derivation, this account does not explain why the priming effect is localised to trials that succeed ‘correctly’ answered primes. Intuitively, discourse-level phenomena like referential pacts are not subject to such quick-cycling effects and are unlikely to be applied locally. The localised effect observed in Experiment 2 suggests that priming may occur via a more automatic and potentially subconscious mechanism than the type of sociopragmatic ‘mechanism’ responsible for establishing referential pacts.

4.8.10. Experiment 2c - Post-Hoc Item Analysis

The heterogeneity of MI observed in Experiments 1a and Experiment 2 poses interesting questions. Although investigating the factors that affect MI availability is not a primary concern of the current thesis, post-hoc analysis of the data gathered in Experiment 2 can shed some light on the nature of MI. The post-hoc analysis of Experiment 1a’s by-item rates of MI provided anecdotal evidence to suggest that the typicality of the marked and unmarked images and their interaction may play a role in the availability of MI. Nevertheless, this effect was obfuscated by the fact that the majority of the critical trials did not involve MI derivation. However, data from Experiment 2c contains a far higher proportion of critical trials that involve MI and, therefore, may be able to clarify the effect of item typicality of the availability of MI.

To investigate this, I fitted a linear regression with polynomial terms, a type of regression used to model non-linear relationships. The model used the size of the by-item priming effect over the overall by-item rate of MI derivation on the basis that it is a better reflection of MI derivation and is less likely to reflect other factors, e.g., ambiguity between the images. To investigate the extent to which the typicality of the marked and unmarked images and their interaction predicted the magnitude of the priming effect in Experiment 2c. The model was fitted to the entirety of the data (i.e., all of the critical trial responses). The overall model was significant ($F(8, 4121) = 92.5, p < .001, R^2 = .015$). Figure 4-24 summarises the linear and quadratic terms and the interactions between them for each variable.

	Estimate	SE	z	p
Intercept	0.33	0.001	209.33	<0.001
Poly(Typicality, Unmarked Image,2)1	0.15	0.09	1.67	0.1
Poly(Typicality, Unmarked Image,2)2	-1.39	0.10	-14.20	<0.001
Poly(Typicality, Marked Image,2)1	0.66	0.12	5.32	<0.001
Poly(Typicality, Marked Image,2)2	-2.13	0.12	-17.00	<0.001
Poly(Typicality, Unmarked Image,2)1: Poly(Typicality, Marked Image,2)1	158.89	11.32	14.04	<0.001
Poly(Typicality, Unmarked Image,2)2: Poly(Typicality, Marked Image,2)1	-12.16	12.16	-1.00	0.31
Poly(Typicality, Unmarked Image,2)1: Poly(Typicality, Marked Image,2)2	64.77	15.07	4.30	<0.001
Poly(Typicality, Unmarked Image,2)2: Poly(Typicality, Unmarked Image,2)2	-94.28	13.31	-7.08	0.02

Figure 4-24: the results of the polynomial regression: `model_poly <- lm(Magnitude of Priming Effect in Exp 2c ~ poly(Typicality of Unmarked Image, 2) * poly(Typicality of Marked Image))`

Notably, there is a significant interaction between the linear effects of the typicality of the marked and unmarked image ($\beta = 158.886$, $p = <.001$), suggesting when both values are high in tandem, a much stronger increase in the magnitude of the priming effect is predicted. There is also a significant interaction between the linear term of the typicality of the unmarked image and the quadratic term of the typicality of the marked image ($\beta = 64.17$, $p = <.001$) and a significant interaction between the quadratic terms of both ($\beta = -94.28$, $p = 0.02$). Figure 37 provides a visualisation of the magnitude of the priming effect predicted by the model across different combinations of the respective typicality of the marked and unmarked images.

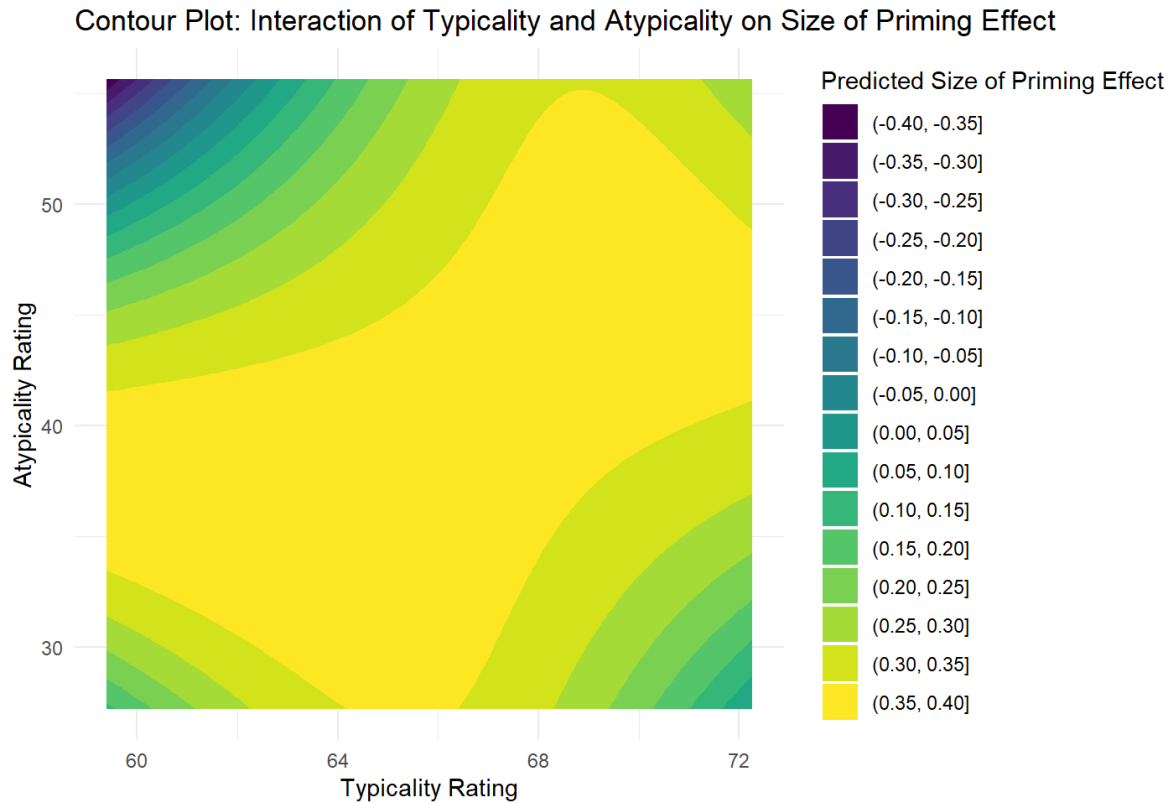


Figure 4-25: contour plot based on the predictions of Figure 5-24's polynomial regression. Here, the magnitude of the priming effect is predicted across continuous combinations of typicality scores for the marked and unmarked images.

4.8.11. *Effect of item typicality*

The predictions of the polynomial regression model suggest that the respective typicality of the marked and unmarked images and the interaction between them has an effect on the magnitude of the priming effect in the experimental items in Experiment 2c – although, this accounts for a small amount of the variation in the size of the priming effect ($R^2 = .015$).

In the broadest terms, the effect of both the typicality of the marked image and of the unmarked follows an inverse 'U'-shape, such that they increase the size of the priming effect when it is at moderate levels but has a diminished, or negative, effect when it is either relatively low or relatively high (i.e., the quadratic effect). The interaction of these effects (i.e., the interaction of the quadratic terms) suggests that the magnitude of the priming effect is 'boosted' when the typicality of both the marked and unmarked images increases to moderate levels, but that that at extreme levels of either, the terms individual terms positive effects on the magnitude of any priming decrease. As demonstrated in Figure 4-25, the largest degree of priming is predicted when levels of both are moderate. I interpret these findings to

suggest that less available MI, e.g., those with less stark contrast between the marked and unmarked image, benefit the most from the priming condition, providing the typicality of neither image is too extreme, e.g., ‘skirt’, pictured below.¹⁵



Marked	Unmarked
	
<p>(marked image typicality: 34, unmarked typicality: 62.55, difference: 28.55, baseline: 15%, magnitude of priming effect in 2c: 26%)</p>	

Figure 4-26: marked and unmarked images for the item ‘skirt’ and data from Item Screening Task 2 and Experiments 1a and 2c.

In contrast, the interaction between the linear term of the typicality of the unmarked image and the quadratic term of the typicality of the unmarked image suggests that high typicality levels of the marked image contribute a negative effect to the magnitude of the priming effect (i.e., make it smaller) and that this negative impact is not offset by the effect of the typicality of the marked image, when it is at the extreme of its typicality levels, e.g., relatively low. Consequently, the ‘full force’ of the effect on the typicality of the marked image is unmodulated when the typicality of the unmarked image is low. Here, ambiguous MI that likely incurs a high degree of I-Implicature is predicted low (or negative) priming effects, e.g., ‘cowboy boot’ in . This finding chimes with my earlier predictions – if particular experimental items do not elicit genuine MI, a priming effect should not be observed in them.

¹⁵By definition, the model predicts the magnitude of the priming effect for combinations of the variables that do not exist, such that the examples used in this section are approximate fits to the patterns predicted by the model.



Marked	Unmarked
	
<p>(marked image typicality: 44.18, unmarked typicality: 59.52, difference: 15.34, mean rate of marked image selection: 48%, magnitude of priming effect in 2c: 4%).</p>	

Figure 4-27: marked and unmarked images for the item 'cowboy boot' and data from Item Screening Task 2 and Experiments 1a and 2c.

The quadratic term of the typicality of the unmarked image suggests when the typicality of the marked image is high, its effect has already peaked (i.e., at moderate levels) and has diminished. Similarly, the quadratic term of the marked image suggests that its negative effect is not triggered when the typicality of the marked image is low. Moreover, the interaction between the two quadratic terms peaks when both are elevated in tandem. Consequently, low rates of priming are predicted for MI with relatively marked, marked images and relatively typical unmarked images. Although this configuration is intuitively ideal for eliciting MI, such MI may be less susceptible to priming because MI is already sufficiently available, and this availability may not be improved beyond its baseline levels by the experimental paradigm. While this conclusion may appear to suggest that such MI should confer the highest rates of MI derivation, it should be noted that item typicality paints just part of the picture and the ultimate rate of MI derivation is likely affected by factors not considered in the analysis, too, e.g., objective utterance length, perceived utterance markedness, etc.

Overall, it is difficult to understand what motivates the selection of the marked image in the MI critical trials, how this differs between experimental items and how these motivations are modulated by the priming condition. Although the priming effect should apply to all of the experimental items in a uniform manner, there are manifold factors at hand that could limit, prevent or augment the priming effect itself and whether the priming effect leads to sufficient activation of the implicature usage mechanism. That said, the results of the current analysis

suggest that the relative markedness of an MI's referent influences the magnitude of any priming effect and that 'good-but-not-great' MIs seem to be the best candidates for priming – as predicted.

4.9. Experiment 2 – Overall Conclusions

As discussed, Experiments 2a, 2b and 2c provide convincing evidence to suggest that MI can prime and be primed by other MI. This observation follows studies that show that other types of conversational implicature can be primed (e.g., Bott & Chemla, 2016; Rees & Bott, 2018; Meyer & Feiman, 2021). MI-to-MI priming is a novel observation that contributes to the scant body of experimental research involving MI. Moreover, Experiment 2 provides a novel framework for the experimental operationalisation of MI from a methodological perspective.

Across all experiments, the priming effect was primarily concentrated in critical trials that were immediately preceded by 'correctly' answered prime trials, indicating that the increase in MI derivation is caused by a direct priming effect rather than an effect that is applied uniformly across the critical trials in the prime conditions. As discussed, the immediacy of the priming effect and the fact that the priming effect is not the product of explicit instruction or learning – at least in Experiment 2a, suggests that the increase in the rate of MI derivation in the 'primed' critical trials is driven by a short-lived, mechanistic priming that functions on each experimental item individually. This is most compatible with an activation-based account of priming, although the mechanism by which the priming occurs is unclear. As discussed, the findings are equally compatible with priming occurring via a usage mechanism, a mechanism involved in the search for an alternative, or a domain-generic but implicature-specific mechanism.

Moreover, the priming effect may not be homogenous, and it is possible that 'priming' can occur across different domains in tandem; for instance, while the direct priming of the usage mechanism may be the simplest explanation for the results of Experiment 2, this may occur alongside the priming of alternative-specific mechanism. Moreover, it is possible that discourse-level mechanisms like *context adaptation* or phenomena like QUD, relevance and referential pacts may modulate the priming effect – although they are unlikely to contribute to the main effect observed.

Moreover, while Wilson (2017) and Katsos and Wilson (2016), a.o. Do provide evidence that adults and children are sensitive to Manner-based contrasts; the current paradigm's specific targeting of the derivational mechanism provides novel evidence that such contrasts are the product of a cognitive, inferential mechanism. Moreover, the priming effect occurs between different types of ad-hoc MI (i.e., across the 42 items). Consequently, Experiment 2 provides evidence for the psychological reality of a distinct class of Manner-based inferences. Nevertheless, the extent to which MI occurs across the experimental items and the factors that govern MI availability are unknown, although the markedness of the particular referents may play a role.

Experiment 2 does not deal with between-category priming effects; therefore, the findings are compatible with multiple accounts of implicature derivation, e.g., the Gricean-inspired perspective, Relevance Theory, Probabilistic Accounts, a.o.

5. Priming MI with I-Implicature

5.1. Experiment 3 – I-Implicature-to-MI Priming

Experiment 3 investigates the extent to which MI can be primed by I-Implicature using an adapted version of the paradigm presented in Experiment 2a that involves I-Implicature primes modelled after the I-Implicature critical trials in Experiment 1b.

5.1.1. Motivations

Investigating whether I-Implicature-to-MI priming occurs in the same way as MI-to-MI priming is interesting from multiple perspectives.

5.1.1.1. The nature of I-Implicature

The existence of I-Implicature as a type of conversational implicature is unclear. While It features prominently in Levinson's (2000) framework of conversational implicature, other accounts give it less focus; for Horn (1984), the same unmarked-to-unmarked is referred to as *R-Association*, while it is unclear whether I-Implicature should be considered an example of

conventional implicature or *conversational implicature*, or something in between, for Grice (1975).

According to Horn, while *R-Association* is the product of *the division of pragmatic labour* (i.e., the conflict between the Q- and the R-Principle), it is not strictly an inference, but rather an antithetical component to *R-Inference* (or MI). Similarly, for Grice (1975), I-Implicature does not arise in the same way as other types of implicature (e.g., via the violation of a conversational maxim¹⁶). Instead, I-Implicature surfaces as the alternative in MI derivation, although it is not afforded any deal of focus. Discussion concerning the nature of I-Implicature is complicated by the observation that implicature can be defined in terms of its primary extension (denoted by the symbol ‘+>’) or its secondary extension (denoted by ‘++>’), which usually entails the integration of some degree of contextual information. Therefore, although neither Grice nor Horn affords I-Implicature any status as an actual type of conversational implicature, both would likely consider inferences like ‘++> I walked our dog across Wandsworth Common’ from the utterance ‘*I walked the dog*’ as a type of conversational implicature.

For Levinson (2000), however, the nature of I-Implicature is comparatively fleshed-out. According to Levinson’s perspective, although I-Implicature is a subtype of conversational implicature, its derivation differs from that of QI and MI. I-Implicature is derived by default and does not involve the processing of an alternative and is consequently not considered to be metalinguistic in nature (i.e., the listener does not need to consider the way in which the speaker has or has not spoken in order to derive it). As described, this perspective is intuitively incompatible with some of the reported examples of I-Implicature cited by Levinson, which seem to require a level of contextual input that suggests that some I-Implicature cannot be generated by default (i.e., the types of secondary extensions denoted by ‘++>’). Here, it appears that I-Implicature falls on both sides of Levinson’s *generalised versus particularised* distinction of conversational implicature.

Considering MI is a type of particularised conversational implicature that involves the processing of an alternative, the observation of a priming effect, or lack thereof, between I-

¹⁶ Although, Huang (2003; 2017) notes that conversational implicature can arise both via the adherence to and violation or flouting of conversational maxims in the Gricean perspective.

Implicature and MI may shed light on their similarities or differences, offering insight into the nature of I-Implicature.

5.1.1.2. *The nature of MI's derivation mechanism*

Understanding whether I-Implicature can prime MI can help us understand the nature of the mechanism responsible for the derivation of MI and, in turn, the nature of MI itself. According to the Gricean-inspired perspective, the derivation of MI is ad-hoc, particularised and involves the processing of an alternative (i.e., an alternative, unspoken I-Implicature). Consequently, MI does not show a high degree of availability, unless primed by other MI (see Experiment 1a and Experiment 2). In contrast, I-Implicature – to some extent, at least – is said to be generated by default, irrespective of context and without the consideration of alternative – this perhaps explains the high degree of I-Implicature availability observed in Experiment 1b and in Katsos and Wilson (2016). Given the fundamental differences between I-Implicature derivation and MI, their derivation is likely not governed by the same cognitive mechanism – as predicted by Levinson's (2000) perspective. Moreover, there are fundamental differences between the derivation of MI and QI – as discussed in 2.3.6. The disparity between MI and I-Implicature raises interesting questions concerning the degree to which the mechanism responsible for their derivation is shared or distinct.

5.1.1.3. *Shared and distinct mechanisms in implicature derivation*

Bott and Chemla (2016) pay particular attention to the nature of the mechanism responsible for deriving enrichments-via-alternatives (EVAs), i.e., implicatures that involve the nature of the alternative, e.g., MI and QI, but not I-Implicature. While the between-category priming effect observed in Bott and Chemla (2016) is taken as evidence that the derivational mechanism is shared between ad-hoc QI, some QI and numeral enrichment, alternate theoretical perspectives are acknowledged, e.g., that the derivational mechanisms for each EVA are distinct, or that they are partially-shared, but not fully shared, between subtypes. Similarly, Meyer and Feiman (2021) consider the perspective that implicature derivation involves a series of sub-computations, i.e., as outlined in Figure 3-5, that can be primed independently. From this perspective, there exists multiple possibilities for which sub-computations may be shared or distinct between different implicature subtypes. Meyer and

Feiman (2021) conclude that the generation of the alternative and the negation of the alternative can be considered independent mechanisms that can be shared, or not, between different implicatures. Rees and Bott's (2018) finding that EVA's can be primed with 'alternative' primes supports this.

As noted in 2.3.5, I-Implicature is not thought to be an EVA and its derivation is unlikely to involve an alternative, in any manner. In contrast, while MI and QI are both considered examples of EVA, there are fundamental differences in their nature. For instance, in MI derivation, the alternative is semantically equivalent to the 'what is said', while in QI it is not. Moreover, MI is particularised, ad-hoc and globally-enriched, but QI is not. On the flip side, I-Implicature shares qualities with MI that QI does not, for instance, I-Implicature is context-based. Further to that, depending on the nature of the I-Implicature elicited in the paradigm (i.e., primary or secondary implicature), more qualities may be shared between it and MI e.g, context-dependency.

If Bott and Chemla's (2016) paradigm accesses an EVA-specific derivational mechanism, then it is unlikely to be shared with I-Implicature. In addition, it is unclear whether MI is derived by this mechanism – after all, the EVAs tested by Bott and Chemla (2016) share more than just that they are enriched via an alternative. Nevertheless, the findings of Bott and Chemla (2016), a.o., indicate that partly-generic mechanisms that govern the derivation of types of enrichment exist and can be primed. Considering this, the study aims to investigate whether there exists a shared or partly-shared mechanism responsible for the derivation of I-Implicature and MI given the theoretical ramifications the existence of such a mechanism has for the Gricean-inspired perspective's account of both phenomena.

5.1.2. Participants

Participants (N = 80) were recruited from Prolific (www.prolific.com) [2024] on the basis that they were adults (≥ 18 years old) and self-identified English speakers and had not previously participated in any of the study's experiments. Incomplete data was rejected from N = 3 of the participants, therefore data was collected from N = 77 participants.

5.1.3. Design

The design of Experiment 3 was almost identical to that of Experiment 2. Experiment 3 involved the same 42 critical trials as Experiment 2 and included 42 different prime trials. The prime trials were identical to the 42 I-Implicature critical trials used in Experiment 1a.

Like Experiment 2, Experiment 3 had a between-participants structure. Therefore, each participant provided a response for 14 primed critical trials, 28 prime trials, and 56 filler trials (non-filler-to-filler ratio, 1:1) and gave responses to a total of 98 experimental trials. Participants were randomly assigned to one of three trial blocks, each comprising 14 unique critical trials, 28 unique prime trials, and 56 filler trials presented in a *filler* → *filler* → *filler* → *prime #1* → *prime #2* → *critical trial* order in which the constituents were completely randomized. Experiment 3 did not involve a training block before the main body of the experiment.

5.1.4. Materials

5.1.4.1. Critical Trials

Experiment 3’s critical trials were identical to those of Experiment 2 and consisted of the following components: 1) a caption reading “**Dr Smith: ‘Can you see a [item label]?’**” which contained one of the 42 item labels presented, e.g., ‘tuxedo’, 2) a second caption, reading ‘**Alan: ‘Yes, I [can] see [...]**’ that contained a prolix description 3) the marked and unmarked images representative of the mentioned object and 4) static images of both ‘Alan’ (a man in a black shirt wearing a VR headset) and ‘Dr Smith’ (a man dressed in a lab coat). See the critical trial for the item ‘tuxedo’ below:

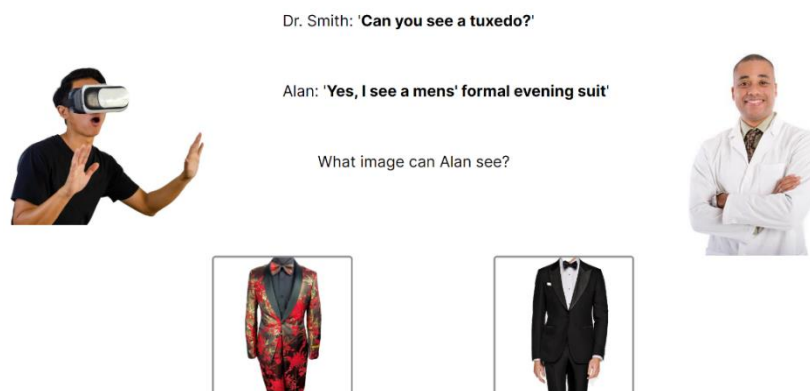


Figure 5-1: MI critical trial for the item 'tuxedo'

5.1.4.2. Prime Trials

The prime trials were identical to the critical trials of Experiment 1a. Unlike Experiment 2c, the critical trials were not guided – this rationale was two-fold. Firstly, the availability of I-

Implicature in the I-Implicature prime trials is likely sufficiently high, as evidenced by Experiment 1b's baseline. Consequently, the proportion of 'correctly' answered primes will likely be far higher than those of Experiments 2a and 2b. Secondly, although the participants in Experiment 2c did appear to engage with the primes pragmatically, I wanted to limit the extent to which priming could be considered the product of either implicit or explicit leaning or changes in a perceived referential pact. Resultingly, I decided to include unguided I-Implicature prime trials in Experiment 3. Pictured below is the I-Implicature prime trial for the item 'vest':

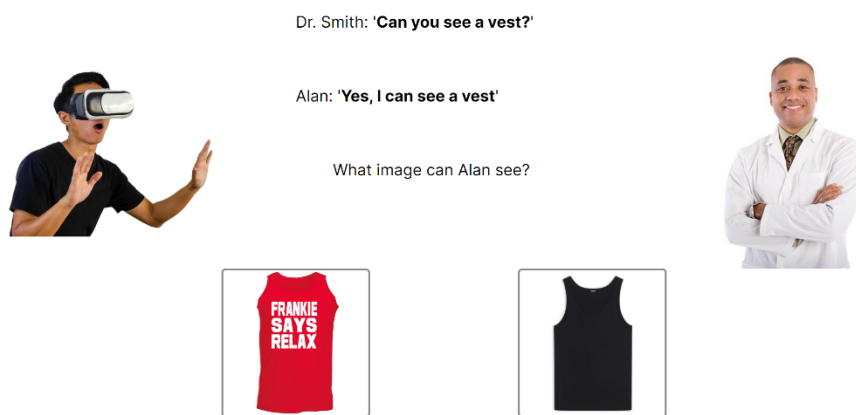


Figure 5-2: I-Implicature prime trial for the item 'vest'

5.1.4.3. Filler Trials

Experiment 3 used the same fillers as Experiment 2a, the 'Yes', No-Marked-Unmarked and No-False-False filler trials.

5.1.5. Procedure

The task experiment was hosted on Gorilla Experiment Builder (www.gorilla.sc) and completed by participants recruited via Prolific (www.prolific.com) [2024]. After the collection of consent and demographic data, the participants were shown the following instruction screen, which is the same as that of Experiment 2a. Note that the elements added to the instructions in Experiments 2b and 2c are removed from the instructions in Experiment 3. Given that Experiments 3-to-6 involve neither guided primes nor a training block, the data from Experiment 2a will be the point of comparison. As such, I chose to include the same instructions as Experiment 2a for the sake of uniformity – see below:

Instruction Screen 1: *“Welcome to the task! In this task you will be given sets of two images and a caption. You will be tasked with matching the caption to one of the images - you will do this by using by clicking on the image”*

Instruction Screen 2: *“In this experiment, you will be presented with a fictional scenario involving Alan and a researcher. Alan is participating in a visual recognition experiment and has been tasked with telling the researcher what image he can see in the VR headset. The researcher will show Alan two images in the VR headset, one will then disappear. Alan is tasked with communicating which image stays, i.e., which image he can still see. Try not to think too hard about your answer!”*

5.1.6. Predictions

Based on the Gricean-inspired perspective’s account of I-Implicature and MI and my own intuitions, I make the following tentative predictions:

- The I-Implicature primes will lead to lower rates of MI derivation in succeeding critical trials compared to MI primes (i.e., Experiment 2a).
- If ‘I-Implicature’ primes elicit secondary inferences (i.e., those denoted by the symbol ‘++>’), there may be a small priming effect.
- If priming occurs, the priming effect will be concentrated in critical trials immediately preceded by two correct primes, compared to those that are not.
- Any priming effect is likely of a different nature to the MI-to-MI priming effect.

5.1.7. Coding

The selection of the marked image was considered indicative of MI derivation in Experiment 3’s critical trials, while the selection of the unmarked image in prime trials was interpreted to indicate I-Implicature, and responses were coded with the numerical score 1 (i.e., ‘correct’) if they derived implicature while filler trials were coded in the same way as previous experiments.

5.1.8. Results & Analysis

The rate of correct image selection was > 95% for each participant across the ‘Yes’ filler trials. Therefore, no participants were excluded from the analysis. The rate of marked image selection across the critical trials of Experiment 3 was 26%, while the rate of unmarked image selection (i.e., ‘correct’ responses) was 86% in the I-Implicature prime trials. In the filler trials, the correct image selection was 94% for Yes’ fillers, 60% for No-Marked-Unmarked fillers, and 46% for No-False-False fillers. There was an increase in the rate of marked image selection between the baseline condition (Experiment 1a) and Experiment 3 (15% → 26%) and a small increase between Experiment 2a, the MI-to-MI priming condition, and Experiment 3 (23% → 26%). Figure 5-3 demonstrates the difference between the rate of MI derivation across Experiments 1a – 3.

The mean rate of MI derivation in the critical trials varied between experimental items ($r=0.63$, $sd=0.13$). The item ‘purse’ had the lowest rate of MI derivation at 0% (baseline =0%), while the item ‘cowboy boot’ had the highest at 63.15% (baseline = 49%). The difference between the baseline and primed rates of MI varied by item ($r=0.36$, $sd= 0.09$), with the smallest ‘priming’ effect (or largest ‘negative priming’ effect) observed in the item ‘necklace’ at -9% and the largest in ‘beanie’ at 27%.

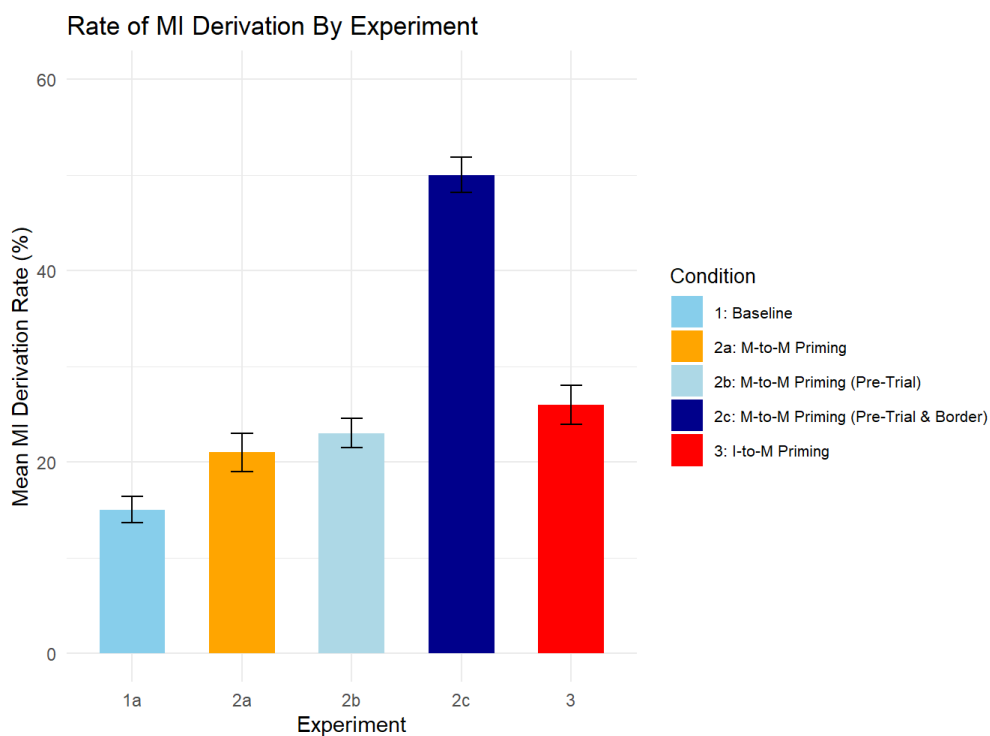


Figure 5-3: The mean rate of MI derivation across Experiment 1a and Experiments 2a, 2b, 2c and 3, and 5 with error bars representing the standard error of the by-item mean rate of MI derivation.

5.1.8.1. *Primed Subset*

Unlike Experiment 2a, the rates of ‘correct’ responses (i.e., rates of I-Implicature) in the unguided I-Implicature prime trials were relatively high – of 1022 critical trial responses, 794 were immediately preceded by two ‘correctly’ answered prime trials. The ‘primed’ subset, i.e., critical trials immediately preceded by two correct trials, had a rate of MI derivation of 22%. In contrast, those that were ‘unprimed’, i.e., immediately preceded by two incorrect primes, i.e., marked image selections, had a rate of MI derivation of 39%. Data from all of the participants was present in each subset, indicating that all of the participants answered the primes incorrectly at some point.

5.1.8.2. *The effect of the primed condition*

I fitted a generalized, linear mixed-model using a binomial distribution and a logit link function (i.e., logistic regression) to the overall data with the condition (Primed: *I-Implicature* vs. Baseline) as a fixed effect and a random intercept and slope for the experimental item and intercept for the individual participants. The overall model was significant ($\beta = -2.58$, $p = <.001$), as was the main effect of the experimental condition ($\beta = 0.99$, $p = .002$). I then fitted the same model to the ‘primed’ subset data. The overall model was significant ($\beta = -2.58$, $p = <.001$), too, as was the main effect of the experimental condition ($\beta = 0.9$, $p = 0.006$).

5.1.8.3. *Comparison to the Experiment 2a*

I fitted the same logistic regression model but with just an intercept for the experimental item and an intercept for the participants to a dataset that contained the critical trials from Experiments 3 and 2a, without the baseline – the model failed to converge with a slope for the experimental item. The overall model was significant ($\beta = -1.6$, $p = <.001$), while the main effect of Experiment (i.e., 2a or 3) was not ($\beta = -0.16$, $p = 0.57$), indicating that there is not a significant difference between the likelihood of MI derivation in either of the experiments.

5.1.8.4. *Item Analysis*

As described, the priming effect is not uniform between the experimental items. To investigate if the size of the priming effect varied between the items in a similar fashion to that of Experiment 2a, I performed a linear regression to predict the by-item difference between the baseline and primed rate of MI derivation (e.g., the magnitude of any priming effect) based on the same value in Experiment 2a – in other words, to see if the size of the by-item priming effect in Experiment 3 is predicted by that of Experiment 2a. The size of the priming effect in Experiment 2a was not a significant predictor of the size of the priming effect in Experiment 3 ($\beta = 0.08$, $p = 0.59$), suggesting the priming effect elicited in each experiment primed the experimental items differently.

I ran a second linear regression to predict the difference between the by-item baseline and primed rate of MI derivation (i.e., the magnitude of any priming effect) based on the baseline rate of MI derivation by item. The by-item baseline rate of MI derivation was not a significant predictor of the size of the priming effect in Experiment 3 ($\beta = -0.15$, $p = 0.37$).

Given the relatively small size of the priming effect, I did not run any item specific analyses (i.e., those involving item typicality, etc.)

5.1.9. *Discussion*

5.1.9.1. *Overall priming effect*

Like the MI-to-MI condition of Experiment 2a, the I-Implicature-to-MI condition significantly increased the likelihood of MI derivation in the MI critical trials and the overall rate of MI derivation in Experiment 3 increased to 26%. Unexpectedly, unlike Experiment 2, the priming effect was not primarily driven by the critical trials immediately preceded by two ‘correct’ primes, but rather the priming effect was spread across all of Experiment 3’s critical trials. In the ‘primed’ trials, i.e., those following two correct primes, the rate of MI derivation was 22%, while the rate of MI derivation in the ‘unprimed’ condition, i.e., those preceded by two ‘incorrect’ primes was 39%. As noted, in both ‘primed’ critical trial subset and the overall data, the increase from the baseline was significant.

I conclude that the significant increase in the rate of MI derivation in critical trials preceded by ‘correctly’ answered primes indicates that the I-Implicature prime trials have successfully

primed the derivation of MI in the succeeding MI critical trials. Although the results of the logistical regression suggest that the likelihood of MI derivation is not significantly different between Experiment 2a and 3, it must be noted that there is a crucial difference between the two experiments, as the majority of the critical trials in the former are ‘unprimed’. As such, the lack of a significant difference between the two is either coincidental or may indicate a type of priming that is not the result of preceding MI, that occurs across both experiments. From another perspective, the lack of a significant difference between the likelihood of MI in each could be indicative of a ceiling effect for the particular type of priming elicited in Experiment 2a & 3, perhaps.

5.1.9.2. *Locus of the priming effect*

At a basic level, Experiment 3 shows that the likelihood of MI derivation increases after the derivation of I-Implicature compared to the baseline. Notably, this increase is also present in ‘incorrectly’ answered primes – perhaps to a significantly larger degree. There are a few possible explanations for this:

There may be two independent priming effects happening simultaneously in Experiment 3. On the one hand, the ‘correctly’ answered I-Implicature primes may increase the likelihood of MI derivation in the succeeding critical trials (e.g., as in Experiment 2). On the other hand, the ‘incorrectly’ answered I-Implicature primes may elicit a different type of ‘priming’ effect, e.g., a basic ‘*pick the weird image*’ heuristic, that increases the rate of marked image selection in the succeeding trials.

Interestingly, the increase in MI derivation in ‘unprimed’ trials may be an inverse-preference effect, i.e., whereby less frequent interpretations (e.g., the selection of the marked image in critical trials) confer stronger priming effects than more frequent interpretations (e.g., the selection of the unmarked image) – see Ran et al., (2022) for an overview. This may explain the higher rate of MI derivation in the unprimed trials – although it is unclear whether this increase is statistically significant. Nevertheless, the majority of the MI critical trials do not involve genuine MI derivation. Considering this, it remains a possibility that alternatives to MI, e.g., I-Implicature, may be primed by the paradigm – this perspective is supported by Marty et al.’s (2022) finding that ‘weak’ primes (i.e., non-pragmatic primes) inhibit implicature derivation in succeeding primes.

An alternate perspective is that the nature of the between-category priming effect in Experiment 3 is fundamentally different from the within-category priming effect of Experiment 2 and does not operate on a local level. While previous discussion conceives of the priming effect as the result of the residual activation of a usage-mechanism or mechanism involved in the search of an alternative, the I-Implicature primes may operate via a different type of higher-order implicature usage mechanism – this perspective requires the reconceptualization of the priming effects of Experiment 2, as discussed below.

That said, carving a distinction between the ‘primed’ and ‘unprimed’ critical trials in Experiment 3 may be fallacious. Clearly, the ‘incorrectly’ answered primes in Experiment 2a are unlikely to involve MI; MI is relatively unavailable (see Experiment 1a) and requires a sensitivity to context (e.g., the identification of a marked referent) that makes the selection of the unmarked image unlikely if MI is derived. As mentioned, this justifies the ‘primed’ vs. ‘unprimed’ distinction. In contrast, I-Implicature is much more available (see Experiment 1b), and it is improbable that it can be ‘opted out of’ given that it occurs automatically and does not involve a large degree of computation – at least, according to Levinson (2000). In addition, because I-Implicature is not dependent on context, the relative markedness of any potential referent is less important, meaning that the selection of a particular image is less indicative of implicature derivation than it is for the MI trials. Consequently, it is possible that the ‘incorrectly’ answered I-Implicature primes may still involve I-Implicature to some degree.

Evidently, the priming effect in Experiment 3 is not limited to the ‘primed’ trials. My preferred interpretation is that the priming effect is the same across all trials, given the unlikelihood of competing, quick cycling answering strategies. Moreover, I interpret the priming effect in Experiment 3 to be higher-order and more generic than that of Experiment 2c – as outlined below – which is compatible with the effect applied more evenly across the critical trials. Finally, I consider the ‘correctness’ of Experiment 3’s prime to be of less theoretical importance than that of Experiment 2, which explains the unexpected behavior of the ‘primed’ and ‘unprimed’ critical trials.

5.1.9.3. *Shared and distinct mechanisms*

Until now, the nature of the priming effect elicited by the paradigm has been discussed in vague terms. While I have committed to the perspective that the effect in Experiment 2 is the

result of the activation of some type of implicature-deriving mechanism, the nature of this mechanism is not well defined. Although I conclude that the primed mechanism is more likely a ‘usage mechanism’ than a mechanism involved in the search for, or integration of, an alternative, the scope of this mechanism is yet undiscussed. Experiment 3’s priming effect adds some detail to this picture.

Bott and Chemla (2016) conclude that there exists a subtype-specific (i.e., EVA-specific) mechanism that is, at least partially, shared between different types of EVAs and not with other phenomena (e.g., plurals). In addition, Meyer and Feiman (2021) conclude that different types of implicature may share different levels of sub-computation. As mentioned, an EVA-specific mechanism is unlikely to derive I-Implicature but may be involved in the derivation of MI. Interestingly, evidence of a priming effect in Experiment 3 suggests a different level or type of derivational mechanism than that in Bott and Chemla (2016) and Meyer & Feiman (2021) – and other studies, e.g., Marty et al. (2022). Given that I-Implicature and MI are unlikely to share specific levels of sub-computation, I interpret Experiment 3’s results to be a function of a shared, higher-order level of implicature computation that is primed by the paradigm – that said, this interpretation is tentative and subject to caveats, as discussed below.

5.1.9.4. *Within- and between-category priming effects*

Turning our attention back to Experiment 2c, the advantage of the within-category priming effect over the between-category priming effect of Experiment 3a suggests that the nature of the priming effect is different between the two.

Interestingly, this perspective is supported by the fact that the by-item priming effects in Experiment 3 are not predicted by those of Experiment 2a, nor by the item’s baseline rates of MI derivation. In contrast, the by-item priming effects of Experiment 2a do predict those of Experiment 2c. This suggests that the priming effect in Experiment 3 is not modulated by MI-specific effects, as it appears to be in Experiment 2. In addition to this, the priming effect in Experiment 2 is local, whereas the priming effect in Experiment 3 is not.

As concluded, the priming effect in Experiment 3 is likely the result of a general mechanism of implicature derivation. Moreover, this effect appears to be different to that of the effect in Experiment 2. I interpret these findings to suggest that the priming effect – or part of it – is the product of a priming effect that acts on a level of MI sub-computation that is not shared

with I-Implicature. Without input from evidence or a lack thereof, of a QI-to-MI priming effect, it is unclear whether the priming effect in Experiment 2 involves MI-specific derivation mechanisms or whether the priming effect acts on a mechanism that plays a role in other types of implicature derivation.

5.1.9.5. *Contextual adaptation*

The aforementioned conclusions rely on an activation-based account of the priming effect in Experiments 2 and 3. Nevertheless, I also considered the possibility that the results could be explained by an account of context adaptation, as postulated by Marty et al. (2022). However, adaptation-based views of priming (e.g., Fine, Jaeger, Farmer & Qian, 2013) typically consider the adaptive mechanism to be sensitive to *unexpected* constructions, which, in turn, leads to the rapid online adaptation of probabilistic expectations. Specifically, Marty et al.'s (2022) model of context adaptation to more surprising information results in larger learning effects

Clearly, utterances are less likely to involve implicature than not. From this perspective, the existence of implicature in the prime trials may increase the probabilistic expectation of implicature in the critical trials, as discussed. Nevertheless, I-Implicature is likely an exception to this. Compared to other types of implicature, I-Implicature is far more frequent and processed in a manner that is likely far less cognitively taxing. Considering this, the presence of I-Implicature in the prime trials seems unlikely to lead to radical shifts in the probabilistic expectation of subsequent implicature – particularly types of implicature that require more involved calculation. Consequently, an adaptation-based perspective does not appear to predict the priming of MI by I-Implicature primes.

5.1.9.6. *Priming-via-the-alternative & contextual modulation*

There are other possible explanations for the priming effect in Experiment 3 that deserve attention. Firstly, the results in Experiment 3 could be explained as *priming-via-the-alternative*– as described in Rees and Bott (2018). According to the Gricean perspective, the derivation of MI involves the derivation of a countervailing I-Implicature. For instance, to derive the implicature ‘+> he danced in an unusual way, e.g., awkwardly’ from the utterance ‘*he moved side to side in time with the music*’ a listener must first derive the alternative I-

Implicature ‘+> he danced in a usual way’, so the stereotypical associations with the word ‘danced’ can be negated.

Given this, Experiment 3 could prime the derivation of MI via the salience of its alternative – note, this explanation involves the priming of a representation rather than of an alternative-specific mechanism. That said, there are multiple reasons why this hypothesis is unattractive. Firstly, the MI critical trials already mention the specific alternative (i.e., in the Dr Smith caption). Moreover, a visual representation of the alternative is depicted in each trial (i.e., the unmarked image). Secondly, the alternative in MI is contextually defined and ad-hoc in nature. Furthermore, the I-Implicature prime trials do not involve the same alternatives as the MI critical trials (e.g., the I-Implicature trial item ‘dress’ may precede the MI critical trial item ‘flipflops’).

Consequently, although the alternatives in each trial may share a vague conceptual representation (e.g., of markedness), they do not share a linguistic representation, e.g., as the alternatives for different *some* QI do. Considering these factors, there appears to be no reason why the I-Implicature prime trials would increase the salience of the alternative beyond the levels intrinsic to the MI themselves unless the I-Implicature primes activate a meaning-based representation of I-Implicature as an alternative – however, they are not computed in the same way as MI, so are unlikely to do so.

An alternative explanation for the priming effect in Experiment 3 is that the I-Implicature prime trials increase the likelihood of MI derivation in the critical trials by modulating socio-pragmatic factors that may increase the availability of MI. For instance, the fact that the I-Implicature critical trials increase the proportion of regular, non-prolix language in the experimental trials may serve to further highlight the markedness of the prolix language used in the MI critical trials. Moreover, the I-Implicature trials may simply contrast with the MI trials in a manner that leads to the polarisation of the answering strategies used in each trial type. Common to the vast majority of models of implicature derivation is the existence of a mechanism that allows for the integration of contextual information (e.g., QUDs, knowledge of referential pacts, etc.) From this perspective, the priming effect in Experiment 3 may be one that operates via the modulating effect the I-Implicature trials have on the conversational context.

Although there is no direct evidence to suggest that Experiment 3’s priming effect is not the result of the modulation of the conversational context, this perspective likely predicts that I-

Implicature will have a larger priming effect than other types of prime trials (i.e., those that are not diametric to MI) and will therefore be revisited in light of Experiment 3.

5.1.9.6.1. *The Nature of I-Implicature*

The priming effect observed in Experiment 3 has ramifications for the way I-Implicature is conceived, which in turn has ramifications for how the priming effect in Experiment 3 should be understood.

As described, I interpret the I-Implicature-to-MI priming effect to result from the residual activation of a shared, generic, higher-order mechanism. However, this assumption is based on Levinson's (2000) account of I-Implicature; from this perspective, I-Implicature is generalized, generated by default, does not involve an alternative, is not metalinguistic and is positive in character. Evidently, I-Implicature and MI are very different phenomena and are unlikely to involve shared sub-computational levels – especially considering that the computations levels of the generally accepted model of implicature derivation is stratified by the role of the alternative (Meyer & Feiman, 2021).

Moreover, given that I-Implicature is the result of an adherence to conversational maxims (e.g., Grice's Maxim of Manner or Levinson's 'I-Heuristic') rather than the violation of them, it is unclear why a mechanism specific to conversational implicature derivation¹⁷, no matter how generic or high-order, would be responsible for the generation of I-Implicature.

Nevertheless, Given this, the type of mechanism primed may be at the level of Grice's *Cooperative Principle* or Relevance Theory's *Principle of Relevance* (Grice, 1975; Sperber & Wilson, 1995, respectively) and activate some mechanism responsible for an engagement with pragmatics, generally. Nevertheless, this explanation is dissatisfying, given that a mechanism of general pragmatic adherence likely juggles input from a caddy of different cognitive functions, e.g., memory, auditory perception, Theory of Mind, etc., and therefore, the processing of the I-Implicature-containing utterances in the experimental paradigm is unlikely to substantially activate it. While I have no evidence that broad-scope pragmatic

¹⁷As described by Huang (2017; 2003), the Gricean account of conversational implicature distinguishes between conversational implicature generated by an adherence to maxims (denoted as *conversational implicature₀*) and those that are generated by the flouting or violation of a maxim (denoted as *conversational implicature₁*). Here, I reference the latter.

engagement cannot be primed, it is not intuitive that such a mechanism would be reactive enough to be modulated by the simple processing of I-Implicature.

Assuming the effect in Experiment 3 is the result of an activation-based, mechanistic priming, there are two resolutions to the described conflict. Firstly, the nature of I-Implicature may be different to that described by Levinson (2000). As detailed in 2.3.5 some instances of I-Implicature do appear to share qualities with other types of implicature in a way incompatible with Levinson's conceptualisation of I-Implicature. If the Gricean-inspired perspective of I-Implicature is inaccurate and I-Implicature and MI have more in common than predicted, then the priming effect in Experiment 3 may be a function of a shared level of sub-computation or a function of a generic mechanism responsible for implicature derivation. While Levinson's I-Implicature is context-independent and non-calculable, this may not be the case for all I-Implicature (see Example 2-44). For instance, the type of I-Implicature derived in Experiment 1b and Experiment 3 – i.e., *implicature-to-stereotype* – shows qualities of context-dependency and calculability. Considering this, certain contexts may give rise to certain types of 'I-Implicature' that may be calculated using an alternative – much in the same way as other types of implicature. 7.4.1.3 discusses this perspective in more detail.

5.1.9.7. *Relevance?*

The idea that I-Implicature is a heterogeneous phenomenon is not controversial. Indeed Levinson (2000) does not suggest that all types of I-Inference are derived by the same mechanisms: "The claim therefore is not that the same mental algorithms compute this range of inferences but rather that there is a single overarching principle that licenses them" (Levinson, 2000, p. 118). Considering this, the computation of I-Implicature may be different between primary and secondary inferences (those denoted by '+>' and '++>', respectively). This stance is attractive given that each type of inference appears to require drastically different degrees of contextual input and calculation to be derived. Moreover, types of MI may be similarly divided into those that appear to have only primary inferences and those that carry both primary and secondary inferences. Therefore, the priming effect in Experiment 3 may be a function of a mechanism responsible for the derivation of secondary inference instead.

Take the I-Implicature and MI's associated with the experimental item 'beanie':

- i. *'I see a beanie'*.
 I +> I see a stereotypical beanie.
 I ++> [...], i.e., one that is not colourful and does not have a distinctive pattern.

- ii. *'I see a close-fitting hat'*.
 M +> I do not see a stereotypical beanie.
 M ++> [...], i.e., I see a beanie that is unusually colourful and has a distinctive pattern.

Example 5-1

While the I-Implicature's primary inference may be neither context-dependent nor calculable, the secondary inference appears to be. In contrast, the MI's primary and secondary inferences appear to be both context-dependent and calculable, although perhaps to different degrees. Given this, Experiment 3's preceding I-Implicature trials may activate the mechanism responsible for the derivation of secondary inferences, and this mechanism may also be involved in the general computation of MI or a level of its sub-computation. While the nature of this mechanism is unknown, Levinson (2000) suggests that the Maxim of Relevance may have a role in licencing secondary inferences.

5.1.9.8. *The nature of MI*

Above, I have discussed the unexpected priming effect of Experiment 3 in light of I-Implicature's nature. On the flip side, MI may be closer to I-Implicature than accounted for by the Gricean-inspired perspective, instead. While MI is considered to be an ad-hoc, particularised phenomenon, this perspective often accounts for the type of inferences I have referred to as 'secondary inferences', e.g., Grice's (1975) seminal 'Home Sweet Home' example. In contrast, other proposed types of MI, e.g., *litotes* (Horn, 1991), *evaluatives* (Rett, 2021), *periphrasis* (McCawley, 1978; Katz, 1972), *non-zero-morphemes* (Levinson, 2000), a.o., do not appear to be comparably dependent on context for their derivation, nor do they appear to require the same level of calculation.

From this perspective, the priming effect observed in Experiment 3 may indicate that MI and I-Implicature do share a mechanism of sub-computation but that the existence of such a mechanism need not require a radical reconceptualization of I-Implicature. On the contrary, I-

Implicature and MI may share or partially share a mechanism involved in the computation of primary inference.

Let us revisit the following example:

‘She caused the car to crash’
+> She crashed the car in an unusual way.

Example 5-2

Here, the associated MI does not necessarily require the integration of contextual information to be understood. Moreover, the Gricean-inspired perspective does not give a comprehensive account of which stage of MI derivation relies on contextual input (See 2.3.3.4). Considering this, the generation of the generic, primary MI ‘+> [...] in an unusual way’ may not be radically different from the generation of the generic, primary I-Implicature ‘+> in a stereotypical way’ and both may involve contextually independent alternatives without a large degree of ad-hoc, on-line calculation. Consequently, the priming effect observed in Experiment 3 may be a function of the mechanism of sub-computation responsible for the generation of primary inference, e.g., an EVA- or EVA-like usage mechanism or a mechanism responsible for the search for context-independent alternatives. This account is discussed in 7.4.1.3.

5.1.10. Conclusion

Overall, I interpret the results of Experiment 3 to indicate that the succeeding I-Implicature trials prime the derivation of MI in the succeeding critical trials. Comparing the ‘primed’ vs ‘unprimed’ trials suggests that this effect is not localised to the critical trials immediately preceding ‘correctly’ answered primes. I conclude that the pattern of priming is likely the result of two factors. Firstly, the nature of the priming effect is likely higher-order than that of Experiment 2. Secondly, the ‘correct’ or ‘incorrect’ answering of the I-Implicature primes appears to be less theoretically important than the MI primes. Consequently, the dichotomy between the ‘primed’ and ‘unprimed’ trials is likely less important.

I interpret the priming effect to be fundamentally different to that of Experiment 2 and conclude that the MI-to-MI priming effect is a function of a level of sub-computation specific

to MI derivation. On the flip side, I conclude that the I-Implicature-to-MI priming effect is the result of a more generic, higher-order mechanism that is shared or partially shared between I-Implicature and MI – although this mechanism is unlikely to be a generic ‘pragmatic’ mechanism. However, as discussed, this perspective is superficially incompatible with the Gricean-inspired perspective’s account of I-Implicature and MI. I discuss possible ways the nature of both I-Implicature and MI can be reconciled with the observed priming effect.

As discussed, I-Implicature and MI appear to generate both primary and secondary inferences that vary in the degree to which they are context-dependent and calculated and similar to one another. On the one hand, the derivation of certain types of I-Implicature (e.g., *implicature-to-stereotypes*) or certain levels of I-Implicature derivation (e.g., secondary inferences) may be theoretically closer to the nature of MI than the Gricean-inspired perspective accounts for. On the other hand, the nature of MI’s primary inference may not be as context-dependent, ad-hoc, and calculated as MI is generally considered to be. Both explanations are compatible with the existence of a shared or partially shared mechanism involved in the derivation of I-Implicature and MI – although I make no strong claims about the nature of the mechanism at this point.

While I consider other possible explanations for the effect observed in Experiment 3, e.g., accounts of context adaptation, priming-via-the-alternative and context modulation, I tentatively concluded that effect is the result of an activation-based priming of a level of implicature computation that is either shared or partially-shared between I-Implicature and MI.

6. Priming MI with QI

Having established that MI can be primed by I-Implicature. Experiments 4 and 5 investigate the extent to which MI can be primed by QI – specifically, using two different types of ad-hoc QI. As discussed in 2.3.6, there are fundamental differences in the processing of MI and QI. That said, both are considered types of EVA and therefore, the nature of MI and QI appears to be more similar than that of I-Implicature and MI – at least, according to the Gricean-inspired perspective.

To understand the boundaries of the mechanism governing QI and MI, Experiment 4 presents an adapted version of the paradigm used in Experiments 2 and 3 that involves ad-hoc QI primes, modelled after those used in Bott and Chemla's (2016) original study. Experiment 5 presents a similar experiment that involves adjectival scales, a type of ad-hoc QI implicature that is generated using a contextually-defined, lexical Horn Scale. Experiment 5's prime trials are modelled after the adjectival scales used in van Tiel and Pankratz (2021).

Given that some perspectives consider ad-hoc QI (i.e., '+> there is one cat' from 'There's a cat!') and lexical scalars (e.g., '+> it is not raining torrentially' from 'It is drizzly') to be derived in different ways – as discussed below – understanding whether they differ in their ability to prime MI derivation will shed light on the mechanism by which the priming effect occurs. Moreover, the nature of any priming effect between ad-hoc QI and MI helps contextualise the results of Experiment 3, in turn. Together, Experiments 2, 3, and 4 shed light on the nature of MI and the mechanism by which it is derived.

6.1. Motivations

6.1.1. The nature of QI

For the purposes of this thesis, I have used the term 'Gricean-inspired' to reference a caddy of interrelated perspectives that take inspiration from Grice's (1975) original conceptualization of conversational implicature. While the Gricean-inspired perspective provides a relatively uniform account of MI, this cannot be said for that of QI, and the Gricean-inspired perspective is divided in the way it considers QI.

To recap, QI is considered an ad-hoc, calculated and context-dependent phenomena by some perspectives, e.g., Grice (1975), Horn (1989) and contemporary Gricean theorists (see., Russell, 2006; Geurts, 2009). The Pragmatic Account of QI maintains that all instances of QI derivation are the result of a made-to-order process of Gricean implicature computation. Here, the implicature is the result of a purely pragmatic inferential process that is applied to the utterance by the listener after it has been constructed – in this sense, QI is post-compositional, i.e., occurs after the syntactic and semantic construction of an utterance, and is metalinguistic, i.e., relies on a listener considering whether an utterance adheres to the Cooperative Principle. According to the Gricean perspective, QI may vary in the degree to

which it is warranted by the conversational context; therefore, some types of QI, e.g., *some* QI, are considered generalized, while others, e.g., ad-hoc QI, are considered particularized. Nevertheless, all types of QI are considered to be derived by the same mechanism. This mechanism is also responsible for the derivation of MI.

On the other hand, Lexical Accounts (e.g., Chierchia, 2004; Levinson, 2000) maintain that certain types of QIs are not calculated on-line but are rather the result of a prior pragmatic calculation that is semantically incorporated with the lexical item that elicits QI.

Consequently, the enriched meaning of a scalar term is considered to be the default inference. For instance, high-frequency Horn Scales like $\langle \textit{some}, \textit{all} \rangle$ or $\langle [x], [x] \textit{ and } [y] \rangle$ may become lexicalized, and their constituents automatically interpreted with the pragmatic enrichment. From this perspective, such implicatures are derived by default, not reliant on conversational context¹⁸ and do not require ad-hoc calculation. Moreover, because the enrichment occurs lexically, they are derived locally. However, like the Pragmatic Account (or ‘Gricean’ account), the Lexical Account does make a distinction between generalized and particularized QI. While particularized QI is derived on an ad-hoc – or partially ad-hoc – basis using the standard Gricean-inspired derivation process, generalized implicature is not as described.

On the flip side, Grammatical Accounts of QI (e.g., Fox, 2007; Chierchia, 2006) consider QI to arise as the function of a syntactic exhaustivity operator. Consequently, QI derivation occurs as part of the semantic processing of an utterance and is, therefore, not calculated by a pragmatic mechanism. Moreover, while specific contexts may override the QI (e.g., presupposition, modals, conditionals, a.o.), QI derivation does not require contextual input to be generated. To this end, there is no major distinction between generalised and particularised QI and ad-hoc scalars are also said to be generated by using the same exhaustivity operator – although the exhaustivity operator is not sensitive to the alternatives it excludes, e.g., ‘I’ll have a coffee’ is interpreted as ‘I’ll have *Exh*[a coffee]’ rather than ‘+> I’ll have a coffee but no cake’.

Relevance Theory (Carston, 2005; Sperber & Wilson, 1995) considers all types of conversational implicature to be derived by the same mechanism and, therefore, does not consider generalised and particularised types of QI to be derived in substantially different

¹⁸ N.B. Levinson (2000) allows for the context-dependent defeasance of QI derived by default, but this occurs post-hoc.

ways, although the perspective acknowledges a distinction between the two phenomena. Similarly, the probabilistic mechanisms posited by Probabilistic Accounts of implicature derivation (e.g., Franke, 2009; Frank & Goodman, 2012; Bergen et al., 2016; Lassiter and Goodman, 2015, a.o.) are neutral to typological distinctions between conversational implicature as inferences are modelled distributionally.

There are numerous perspectives on QI, each proposing different combinations of the various qualities and characteristics discussed.

6.1.1.1. *Ad-hoc QI vs. adjectival scalars*

The linguistic forms that generate QI have been widely documented, both theoretically and experimentally. Typically, the lexical triggers of QI are considered part of a lexical scale ordered by informativeness, in which each item is equally lexicalised (van Tiel et al., 2014). Conventional scales have received significant theoretical (e.g., Horn, 1972) and experimental focus (e.g., Papafragou & Musolino, 2003; Geurts & Pouscoulous, 2009; Noveck et al., 2002; Breheny et al., 2006; Chemla & Spector, 2011). However, less conventional scales, such as *<adequate, horrific >* and *<funny, hilarious >*, have also been studied. For instance, Levinson (1983) and Hirschberg (1985) take stock of a range of scales, from those that are highly conventional to those that are more domain-specific (Potts, 2012; van Tiel et al., 2014). Moreover, experimental investigation has demonstrated the breadth of scalar diversity, documenting cardinal scales, quantified and modal scales, ranked orderings, and a plethora of gradable adjectives as instances of QI (e.g., Baker et al., 2009; van Tiel et al., 2014; Pankratz & Van Tiel, 2021).

Some accounts consider conventionalised and non-conventionalised scalars to be processed in different ways, as discussed. For the purposes of Experiments 4 and 5, I will consider two types of QI: *ad-hoc QI* and *adjectival scalars*. Broadly, ‘ad-hoc’ scalars are those that are enriched using a contextually-defined alternative (Hirschberg, 1991). For example, the utterance ‘I’ll have a coffee’ may implicate ‘+> [...] but I won’t be staying the night’ if uttered in response to a post-date invitation home. In contrast, the same utterance may imply ‘+> I’ll have a coffee but no cake’ if uttered as part of an order in a café. Here, two different scales, *<have a coffee, have a coffee and stay the night>* and *<a coffee, a coffee and cake >*, are generated depending on the context.

As detailed in 3.4.3.1, Bott and Chemla's (2016) study involves ad-hoc QI. Here, the experimental stimuli are designed to generate ad-hoc QI that are computed using scales like < *a circle, a circle and a square* >. Van Tiel and Pankratz (2021) present a study that investigates the processing of scalar inferences and stimuli that also elicit ad-hoc QI – in this case, adjectival scalars. Adjectival scalars are a type of ad-hoc QI implicature that is derived using a Horn Scale consisting of gradable adjectives, e.g., the utterance 'Opportunities are scarce' will likely trigger the QI '+> opportunities are not absent' using the scale < *scarce, absent* >. While the inferences elicited in both studies are assumed to generate some form of ad-hoc QI, they appear to differ in the degree to which context plays a role in their calculation. I will use the generic term 'ad-hoc QI' to refer to the type of implicature derived in Bott and Chemla's trials and 'adjectival scalar' to refer to the specific type derived in van Tiel and Pankratz's (2021) study.

Take the following examples:

- i. 'There is a circle'.
+> there is only a circle
++> there is only a circle and not a circle and an arrow.
- ii. 'The cake was mediocre'.
+> the cake was just mediocre.
++> the cake was not good.
++> the cake was not bad.
- iii. 'It was an okay effort'.
+> it was just an okay effort.
++> it was not a good effort.
++> it was not an outstanding effort.
- iv. 'My husband is funny'.
+> My husband is just funny.
++> My husband is not charismatic.
++> My husband is not hilarious.

++> My husband is not witty.

Example 6-1

As demonstrated in Example 6-1(ii), adjectival scalars involving intermediate terms (e.g., ‘mediocre’ or ‘content’) are bidirectional, i.e., they can imply both a lower and upper bound and the direction of the implicature is thought to be influenced by context (e.g., Horn, 1972, 1989; Hirschberg, 1991; Potts, 2012; Chierchia; 2004). Similarly, adjectival scalars are gradable, as shown in Example 6-1 (iii). Here, the QI derived shows variable grades of exclusion, such that ‘okay’ can imply ‘+> not good’ or ‘+> not outstanding’, etc. Similarly, the scope of this exclusion is thought to be dependent on context (Horn, 1989; Hirschberg, 1991; Potts, 2005; Chierchia, 2004). Moreover, the particular alternative accessed by the implicature can vary depending on context – see Example 6-1 (iv) and this is thought to be influenced by QUDs and relevance (Pankratz & van Tiel, 2021).

In contrast, the ad-hoc QI in Example 6-1 (i) is not context-dependent in the same way. While the exact constituents of the scale vary, the structure of the scale, i.e., < ‘x’, ‘x’ and ‘y’ >, is uniform, and the implicatures it generates are unidirectional and are not gradable.

Consequently, while such implicatures may be ‘ad-hoc’ in the sense that context dictates whether or not they are derived, their calculation does not require the construction of an ad-hoc, context-dependent scale in the same way adjectival scalars do. As such, some perspectives (e.g., Horn, 1972, 1989; Hirschberg, 1991) consider these scales to be lexicalised or semi-lexicalised (Levinson, 2000).

Below is a table summarising the qualities of ad-hoc QI and adjectival scalars according to the discussed accounts of QI derivation:

Table 1: Qualities of ad-hoc QI and adjectival scalars.

Qualities: ¹⁹	Default? ²⁰	Calculated?	Particularized?
--------------------------	------------------------	-------------	-----------------

¹⁹ Other debate exists concerning the nature of QI derivation, e.g., localist vs. globalist contentions (referenced above). Moreover, different discussion concerns the truth-conditional content QI contributes utterance (e.g, Recanati, 2003; Carston, 1998; Horn, 1992). Such debate has been omitted because it is less pertinent to the nature of the mechanism responsible for the derivation of QI.

²⁰ While default accounts are often considered diametric to accounts that require calculation, it must be noted that some perspectives, e.g., Horn (1989; 1984), consider QI derivation to occur automatically, i.e., without appeal to context, but also require contextually-dependent calculation after this. Hence, I have presented them separately.

	Ad-hoc QI	Adjectival Scalar	Ad-hoc QI	Adjectival Scalar	Ad-hoc QI	Adjectival Scalar
Grice (1975)	✗	✗	✓	✓	✓	✓
Horn (1972, 1989)	✓	✓	✓	✓	✗	✓
Levinson (2000)	✓	✓	✗	✓	✗	✓
Grammatical Account	✓	✓	✗	✗	✗	✗
Relevance Theory	✗	✗	✓	✓	✓	✓

6.1.2. Investigating a shared QI & MI derivational mechanism

As discussed, it remains contested whether or not QI and MI are derived via the same mechanism. Moreover, within theories that do consider them to be governed by the same overarching mechanism, the extent to which QI and MI are likely to share levels of sub-computation varies. For instance, if MI is a contextually dependent phenomenon that requires ad-hoc calculation, it is unlikely to share many of the sub-computations responsible for the derivation of default, partially generalized QI that are not calculated in a similar fashion, e.g., ad-hoc QI, as per Levinson’s (2000) perspective. In contrast, levels of sub-computation may be shared between MI derivation and QI derivation for QI more reliant on context, e.g., adjectival scalars, that are not shared between the derivation of MI and less context-dependent QI.

On the flip side, the nature of MI is equally contested. Although MI is clearly context-dependent, calculated and particularised, the extent to which these qualities are reflective of primary or secondary inference is unclear. Evidently, the extent to which the mechanism responsible for MI and QI is shared also depends on the nature of MI. For instance, if Relevance is responsible for licensing particularized secondary inferences (e.g., Levinson, 2000), then the derivation of MI and QI may be more similar than superficially apparent.

6.1.2.1. Priming a general mechanism?

As discussed, Experiment 3 suggests that the derivation of I-Implicature and MI may involve a shared or partially shared mechanism. Given that the nature of I-Implicature is unclear, it is difficult to conclude whether or not the primed mechanism is higher-order and generic (e.g., at the level of implicature derivation) or whether levels of sub-computation are shared between the phenomena, too.

Although I discuss alternative ways of conceptualising I-Implicature derivation in 7.4.1.3, I-Implicature is not generally considered to be a type of EVA. In contrast, both MI and QI are thought to involve the computation of the alternative. According to the ‘core account’ of implicature derivation, the processing of an alternative is a distinct level of EVA sub-computation (Meyer & Feiman, 2021) and previous investigation has sought to determine the extent to which the alternative plays a role in the priming of EVA (e.g., Rees & Bott, 2018; Waldon & Degen, 2020).

Considering this level of sub-computation is shared between MI and QI, but not I-Implicature, Experiments 4 and 5 may offer context to Experiment 3’s findings – QI-to-MI priming may benefit from the phenomena sharing a level of sub-computation specific to the processing of the alternative, which may confer a larger priming effect. A larger priming effect in the QI conditions than in the I-Implicature condition may indicate that the priming effect in Experiment 3 is of a higher order, e.g., at an implicature-generic level. Moreover, given that the nature of the alternative is different between ad-hoc QI and adjectival scalars, e.g., the latter has a greater reliance on context, the nature of the alternative-specific mechanism of sub-computation may be more or less shared between MI and the QI-subtypes – this may confer a difference in priming effect, too.

6.1.3. Predictions of the competing accounts

The different ways QI is accounted for give rise to different predictions of a priming effect between QI subtypes and MI. I have summarised these predictions in Table 2 below. Note that these predictions assume that the priming effect varies between phenomena with more or fewer shared levels of sub-computation, i.e., that it is not a higher-order implicature-general priming effect.

Table 2: Predictions of QI-to-MI priming

Qualities:	<i>Is a priming effect predicted?</i>		<i>Is the predicted to be effect equal?</i>
	Ad-hoc QI	Adjectival Scalar	
Grice (1975)	✓	✓	Yes, both types of QI are particularised and likely involve the same mechanisms.

Horn (1972, 1989)	✓	✓	No. As ad-hoc QI is generalised, there may be a larger priming effect for adjectival scalars considering they are more similar to MI.
Levinson (2000)	✗	✓	No. Ad-hoc QI is partially-lexicalised and therefore likely to be derived by default which likely not prime MI.
Grammatical Account	✗	✗	N/A. Priming is not predicted.
Relevance Theory	✓	✓	Yes, the same mechanism is responsible for the derivation of both phenomena and MI, therefore any priming effect should be equal.

6.2. Experiment 4 – Ad-hoc QI-to-MI Priming

Experiment 4 investigated whether MI could be primed with ad-hoc QI using primes similar to those of Bott and Chemla (2016).

6.2.1. Participants

Participants (N = 75) were recruited from Prolific (www.prolific.com) [2024] on the basis that they were adults (≥ 18 years old) and self-identified English speakers and had not previously participated in any of the study’s experiments. Incomplete data was rejected from N = 2 of the participants, therefore data was collected from N = 73 participants.

6.2.2. Design

Experiment 4’s design closely mirrored that of Experiments 2 and 3. The experiment involved 42 different critical trials and 28 different prime trials. While the 42 critical trials were the same as previous experiments, the 28 prime trials were not. Instead, the prime trials were modelled after Bott and Chemla’s (2018) ad-hoc QI trials and involved cards with different shapes on them – as described below. Experiment 4 utilized a between-participants design, and each participant responded to 14 primed critical trials, 28 prime trials, and 56 filler trials.

Overall, each participant completed a total of 98 trials. Participants were randomly assigned to one of three trial blocks, with each block containing 14 unique critical trials, 28 prime

trials, and 56 filler trials – the primes were identical across all of the trial blocks. The trials were presented in the following sequence: *filler* → *filler* → *filler* → *prime #1* → *prime #2* → *critical trial*, and the specific experimental items were completely randomized, as was the left-right configuration of the images.

6.2.3. Materials

6.2.3.1. Critical Trials

Experiment 4 involved MI critical trials that were identical to those used in the previous experiments. They consisted of the following components: 1) a caption reading “**Dr Smith: ‘Can you see a [item label]?’**” which contained one of the 42 item labels presented e.g., ‘t-shirt’, 2) a second caption, reading ‘**Alan: ‘Yes, I [/can] see [...]**’ and contained a prolix description 3) the marked and unmarked images representative of the mentioned object and 4) static images of both ‘Alan’ (a man in a black shirt wearing a VR headset) and ‘Dr Smith’ (a man dressed in a lab coat). See the critical trial for the item ‘tuxedo’ below:

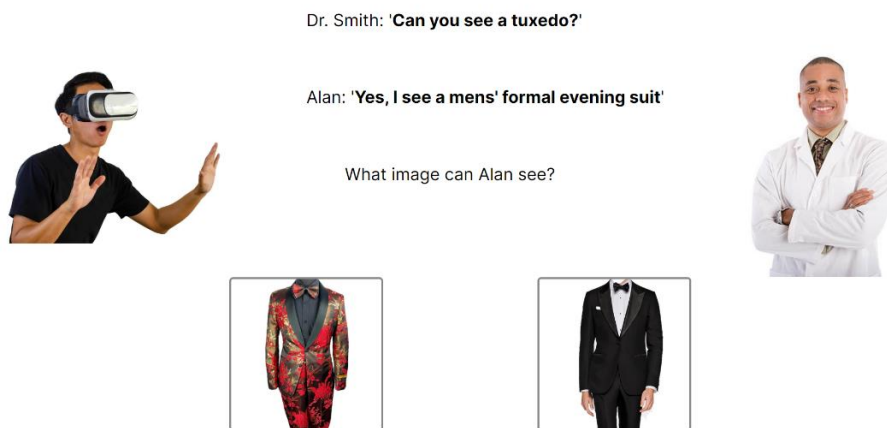


Figure 6-1: MI critical trial for the item 'tuxedo'

6.2.3.2. Prime Trials

Unlike the prime trials of the previous experiments, Experiment 4’s prime trials did not involve a choice between articles of clothing. Instead, the primes were modelled after Bott and Chemla’s (2016) ad-hoc QI trials. They consisted of the following components: 1) a caption reading “Dr Smith: ‘Can you see a card with a [shape]?’” 2) a second caption,

reading ‘Alan: ‘Yes, I [/can] see a card with a [shape]’ 3) two images, one compatible with the pragmatic interpretation of the caption (i.e., +> I can see a card with only a [shape]) and one compatible with a semantic interpretation of the caption, but not the pragmatic interpretation and 4) static images of both ‘Alan’ (a man in a black shirt wearing a VR headset) and ‘Dr Smith’ (a man dressed in a lab coat).

The shapes mentioned in the captions were ‘arrow’ (N = 4), ‘cross’ (N = 4), ‘circle’ (N = 6), ‘star’ (N = 4), ‘triangle’ (N = 6) and ‘heart’ (N = 4). The pragmatically-compatible image (or *strong* image) involved a picture of a playing card with a single shape on it, e.g., a circle. The shape was the same as the shape mentioned in the captions. The semantically-compatible image (or *weak* image) involved a picture of a playing card with two different shapes on it, e.g., a circle and a star, one of which was compatible with the caption and one that was not – in addition to those mentioned above, the incompatible shapes were either ‘square’ or ‘diamond’ shapes. Therefore, the caption sought to elicit the following type of implicature: ‘There is a circle +> there is only a circle and not a circle and a star’.

As discussed above, although the derivation of ad-hoc QI is reliant on context to some extent, the nature of the implicature is relatively uniform (e.g., ‘+> only [x] not [x & y]’). Moreover, the primes involved almost identical implicatures and alternatives (i.e., a single shape vs. a shape and a different shape). Consequently, I did not expect the individual ad-hoc primes to vary much in efficacy – unlike the MI and I-Implicature primes, which elicited diverse rates of enrichment. Resultingly, I used the same 28 prime trials across all of the trial blocks. This meant that each participant was exposed to the same prime trials, unlike in Experiments 2 and 3. Nevertheless, each of the 28 trials was unique and involved a different configuration of three-shaped pictures in the two images, e.g., the single ‘circle’ card is pictured. The images and captions used in each prime trial are represented in 9.5 and 9.6

Like Experiment 3, the primes were unguided. This decision was motivated by the fact that the availability of ad-hoc implicature in the prime trials is likely sufficiently high ²¹ and the concern that guided prime trials may disengage the type of implicature derivation on which the priming effect is reliant. Unlike the MI and I-Implicature trials, I did not pre-screen the

²¹ While Marty et al. (2022) do establish a baseline rate of ad-hoc QI derivation of just 11% in a similarly designed trial, it must be noted that the study’s trials used a ‘Better Picture?’ paradigm. Consequently, Marty et al.’s baseline represents the proportion of participants that would reject a semantic interpretation of an ad-hoc QI implicature in favour of an unseen and, therefore, potentially pragmatic interpretation. In contrast, my paradigm involves a choice between the two interpretations when both are visible, which intuitively increases the likelihood of the participant selecting the pragmatic interpretation.

ad-hoc QI prime trials given the prior testing of almost identical stimuli in Bott and Chemla (2016), Rees and Bott (2018) and Marty et al. (2021; 2022). The ad-hoc QI prime for the shape 'circle' is depicted in Figure 6-2: prime trial for the item 'circle' Figure 6-2 below:

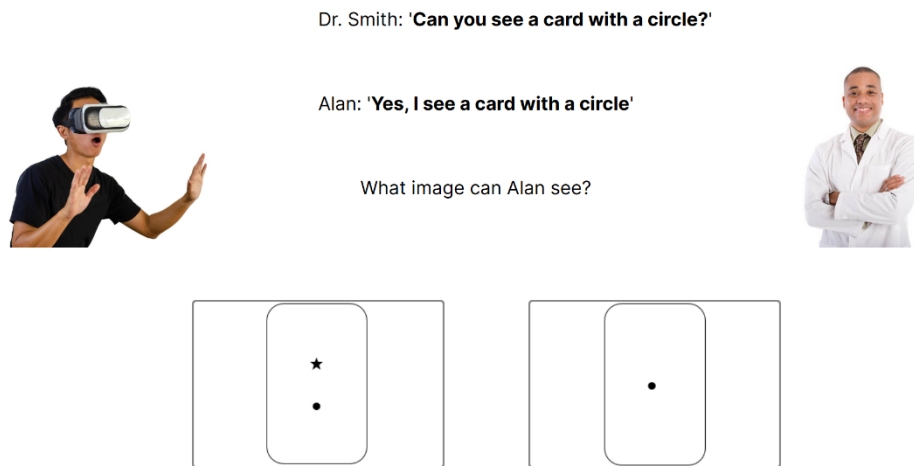


Figure 6-2: prime trial for the item 'circle'

6.2.3.3. Filler Trials

Experiment 3 used the same fillers as Experiment 2a, the 'Yes', No-Marked-Unmarked and No-False-False filler trials.

6.2.4. Procedure

The task experiment was hosted on Gorilla Experiment Builder (www.gorilla.sc) and accessed by participants recruited via Prolific (www.prolific.com) [2024]. After the collection of consent and demographic data, the participants were shown the following instruction screens, which were identical to those of Experiment 3.

Instruction Screen 1: "Welcome to the task! In this task you will be given sets of two images and a caption. You will be tasked with matching the caption to one of the images - you will do this by using by clicking on the image"

Instruction Screen 2: "In this experiment, you will be presented with a fictional scenario involving Alan and a researcher. Alan is participating in a visual recognition experiment and has been tasked with telling the researcher what image he can see in the VR headset. The researcher will show Alan two images in the VR headset, one will then disappear. Alan is

tasked with communicating which image stays, i.e., which image he can still see. Try not to think too hard about your answer!”

6.2.5. Predictions

Based on the Gricean-inspired perspective’s account and my own intuitions, I predict the following outcomes for Experiment 4:

- The ad-hoc QI primes will lead to an increase in the rate of MI derivation in succeeding critical trials.
- The priming effect will be of a different nature to Experiment 2’s MI-to-MI priming effect and will confer a smaller increase in the rate of MI in comparison compared to 2c.
- The priming effect in Experiment 4 will likely be larger than that of Experiment 3, if QI and MI are derived by a more similar mechanism than that of I-Implicature and MI.
- Like Experiment 3, the priming effect will not be localized to ‘primed’ trials.

6.2.6. Coding

The selection of the marked image was considered to indicate that MI derivation had occurred in Experiment 4’s critical trials, and the selection of the strong image, i.e., the image with the single shape, in the prime trials was taken to indicate ad-hoc QI derivation.

Responses were coded with the numerical score 1 (i.e., ‘correct’) if they derived implicature.

The filler trials were coded in the same way as previous experiments.

6.2.7. Results & Analysis

The rate of correct image selection was > 95% for each participant across the ‘Yes’ filler trials. Therefore, no participants were excluded from the analysis. The rate of marked image selection across the critical trials of Experiment 4 was 22%. In the ad-hoc QI primes, the rate of strong image selection was 96%. In the filler trials, the correct image selection was 94% for ‘Yes’ fillers, 65% for No-Marked-Unmarked fillers, and 45% for No-False-False fillers. The rate of marked image selection increased between the baseline condition (Experiment 1a) and

Experiment 4 (~ 15% → 22%), while there was a small numerical decrease from the rates of Experiment 2a and Experiment 3 (26% → 22%). The rate of MI derivation between Experiments 1a, 2a, 3 and 4 are visualised below:

The mean rate of MI derivation in the critical trials varied between experimental items ($r=0.47, sd=0.08$). The item ‘purse’ had the lowest rate of MI derivation at 5% (baseline =

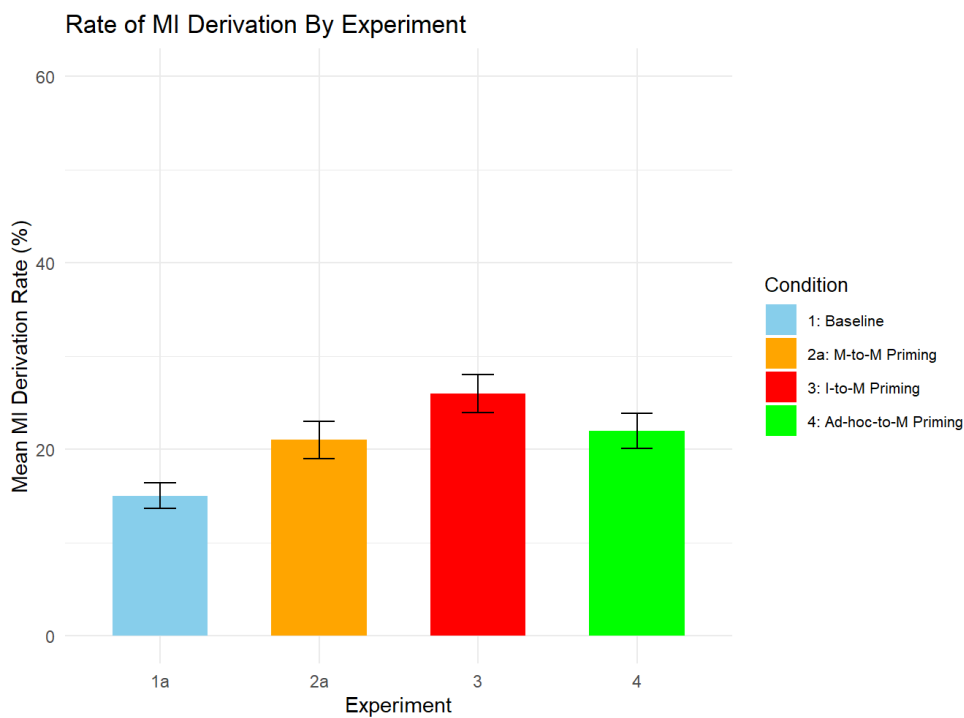


Figure 6-3: The mean rate of MI derivation across Experiment 1a and Experiments 2a, 3, and 4 with error bars representing the standard error of the by-item mean rate of MI derivation. .

0%), while the item ‘headband’ had the highest at 71% (baseline = 33%). The difference between the baseline and primed rates of MI varied by item ($r=0.66, sd=0.13$), with the smallest ‘priming’ effect (or largest ‘negative priming’ effect) observed in the item ‘necklace’ at -8% and the largest in ‘headband’ at 39%.

6.2.7.1. Primed Subset

The rate of ‘correct’ responses to the ad-hoc QI trials was far higher than of the prime trials in previous experiments, at 96% - of the 1022 critical trial responses, 955 were immediately preceded by two correct primes. Removing ‘unprimed’ critical trials that were ‘unprimed’ from the overall dataset led to a small numerical decrease in the mean rate of MI derivation (22% → 21%).

As discussed, the between-category priming effect may not apply locally or may apply both locally and at a higher-order level. Moreover, for some experiments (e.g., Experiment 2a), removing all of the ‘unprimed’ data from the critical trials is not feasible. Consequently, the analyses use the entirety of the dataset gathered from Experiment.

Whether or the data should be screened (i.e., contain just ‘primed’ responses) before being analyzed is an open question. While Bott & Chemla (2016), Rees & Bott (2018) and Marty et al. (2022) all present screened data, they all involve prime trials that elicit implicatures with high rates of availability – in this sense, the analyses can afford to lose the ‘unprimed’ data without compromising statistical power. On the flipside, many of the experiments presented in the current study involve implicatures with less availability. While it is tempting to remove the ‘unprimed’ trials from the data, there is the risk that the remaining data is either cherry-picked or too scant for hypothesis testing, or both. Moreover, if priming does not occur local to ‘primed’ trials, then screening the data loses the contribution of any higher-order priming effects. While I conclude that all of the data should remain in the general analyses, this perspective should be reevaluated in any future research.

6.2.7.2. *The effect of the primed condition*

A generalized linear mixed model using a binomial distribution and logit link function (logistic regression) was fitted to the full dataset, with condition (Primed: *Ad-Hoc QI* vs. Baseline) as a fixed effect. The model included a random intercept and slope for the experimental item and a random intercept for the participants. The overall model was significant ($\beta = -2.53, p < .001$), with a significant main effect of condition ($\beta = 0.65, p = .03$).

6.2.7.3. *Comparison to previous experiments*

I conducted a pairwise comparison to assess the difference between Experiment 4’s priming effects and those of Experiment 3 and 2c using R’s *emmeans* package (Lenth et al., 2024), as presented in Table 3 – note, results are reported with a Tukey adjustment for multiple comparisons – see below:

<i>Contrast</i>	Estimate	SE	z	p
Ad-Hoc QI vs. I-Implicature (Exp.3)	-0.22	0.30	-0.75	0.88
Ad-Hoc vs. MI (Exp. 2a)	-0.01	0.29	-0.04	1.00

Table 3: results of pairwise comparison between the logistical regressions used to model the main effect of condition in Experiments 4 and Experiments 3 and 2a.

The results indicate that there were no significant differences in the priming effects (i.e., the likelihood of MI derivation) between Experiment 4 and Experiments 3 and 2c.

6.2.7.4. Item Analysis

To investigate if the variation in the size of the priming effect in Experiment 4 varied between experimental items in a similar pattern to that of Experiment 3 and Experiment 2c, I fitted a linear regression to see whether the magnitude of the by-item priming effects in each

	Estimate	SE	z	p
By-time magnitude of priming effect (Exp.3)	0.29	0.15	1.9	0.06
By-time magnitude of priming effect (Exp. 2a)	0.15	0.18	0.82	0.42

predicted the magnitude of the priming effect in Experiment 4.

Neither the magnitude of the by-item priming effect in Experiment 2a nor Experiment 3 was a significant predictor of the magnitude of the by-item priming effect in Experiment 4, as

Table 4: results of linear regressions modelling whether the by-item magnitude of the priming effect in Experiment 2a & 3 predicts that in Experiment 4.

presented in Table 4 – although the effect of the former did approach significance. These findings suggest that the priming effect elicited in Experiment 4 effected the items in a different way to that of Experiment 2c and Experiment 3.

I conducted a second linear regression to examine the relationship between the by-item baseline rates of MI derivation and the by-item priming effect magnitude in Experiment 4. While the overall model was significant ($\beta = 0.14, p = < 0.001$), the by-item baseline rate of MI derivation was not a significant predictor of the magnitude of the priming effect by experimental item ($\beta = 0.17, p = < 0.32$). Given the relatively small size of the priming effect, I did not run any item-specific analyses.

6.2.8. Discussion

6.2.8.1. Overall priming effect

Although the bulk of the discussion concerning QI-to-MI priming is presented below (6.3.8), alongside that of Experiment 5, I make the following broad-scope conclusions.

Firstly, I interpret the increase in the rate of MI derivation in the critical trials succeeding ad-hoc QI to indicate that the ad-hoc QI trials indeed prime subsequent MI. Interestingly, while there was no significant difference in the overall likelihood of MI derivation between Experiments 2c, 3 and 4, there was also no correlation between the size of the priming effects in the conditions per experimental item. Moreover, the baseline rates of MI derivation (Experiment 1a) did not predict the size of the priming effect in Experiment 4. Given that the vast majority of the critical trials were ‘primed’, the degree to which the priming effect is localised is difficult to ascertain. Nevertheless, the difference between the rate of MI derivation in the ‘primed’ and ‘unprimed’ trials is small, at least superficially. Considering this, there is no evidence to suggest that the priming effect in Experiment 4 is limited to the ‘primed’ critical trials.

Given this, it seems that the priming effect in Experiment 4 applies to the individual experimental items in a different manner than the previous experiments. On one hand, this may indicate that the nature of the priming effect is different. On the other hand, this may be evidence that the size of the priming effect elicited between the experiments is highly variable and is influenced by other factors that differ between the experiments (e.g., participants, prime type, experiment item as prime) but is not of a fundamentally different nature.

6.3. Experiment 5: Adjectival Scalar QI-to-MI Priming

Having established that MI can be primed by ad-hoc QI, Experiment 5 investigates the extent to which adjectival scalars can prime MI to address the questions outlined in 6.1. The paradigm remains essentially unchanged from previous experiments but involves the adjectival scales modelled after those used in van Tiel and Pankratz (2021).

6.3.1. Participants

Participants (N = 337) were recruited via Prolific (www.prolific.com) [2024] that they were adults (≥ 18 years old) and self-identified English speakers and had not previously participated in any of the study’s experiments. Incomplete data was rejected from N = 40 of the participants, therefore data was collected from N = 297 participants.

6.3.2. Design

The design of Experiment 5 was essentially the same as previous experiments. However, while Experiment 4 used the same 42 critical trials as previous experiments, only 14 prime trials were used. The prime trials were created to elicit adjectival scalars using stimuli from van Tiel and Pankratz (2021) – outlined below.

Like previous experiments, Experiment 5 used a between-participants design. However, there were not enough pretested lexical scalar stimuli from van Tiel and Pankratz's (2021) study to create 28 individual prime trials – the minimum number required to avoid repetition if the three trial blocks are used, i.e., as in Experiment 2 and 3. As such, each participant responded to 7 primed critical trials, 14 prime trials and 28 filler trials, completing 49 trials. Participants were randomly assigned to one of six trial blocks, each containing 7 different critical trials, 14 prime trials and 56 filler trials. – the primes were identical across all of the trial blocks. The trials were presented in the following order: *filler* → *filler* → *filler* → *prime #1* → *prime #2* → *critical trial* and the constituents of this sequence were completely randomised.

6.3.3. Materials

6.3.3.1. Critical Trials

The critical trials were identical to those of previous experiments. They consisted of the following components: 1) a caption reading “**Dr Smith: ‘Can you see a [item label]?’**” which contained one of the 42 item labels presented e.g., ‘t-shirt’, 2) a second caption, reading ‘Alan: **Yes, I [/can] see [...]**’ and involving a prolix description 3) the marked and unmarked images representative of the mentioned object and 4) static images of both ‘Alan’ (a man in a black shirt wearing a VR headset) and ‘Dr Smith’ (a man dressed in a lab coat). See the critical trial for the item ‘tuxedo’ below:

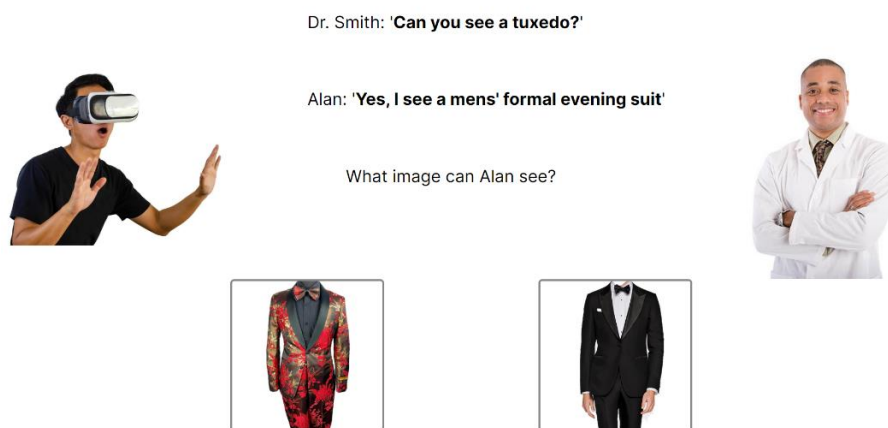


Figure 6-4: MI critical trial for the item 'tuxedo'

6.3.3.2. Prime Trials

Experiment 5's prime trials involved the lexical scalars used in van Tiel and Pankratz (2021). They consisted of the following components: 1) a caption reading "**Dr Smith: 'What can you see?'**" 2) a second caption, reading '**Alan: 'Yes, I [can] see a [adjectival scalar]'**' 3) two images, one compatible with the pragmatic interpretation of the caption (e.g., +> I can see a content but not happy lady) and one compatible with the alternative, or non-enriched interpretation of the caption (e.g., 'I see can a happy lady') and 4) static images of both 'Alan' (a man in a black shirt wearing a VR headset) and 'Dr Smith' (a man dressed in a lab coat).

Van Tiel & Pankratz (2021)

As mentioned, the items were constructed using stimuli from van Tiel and Pankratz's (2021) study, which investigated the influence of adjectival polarity on the processing of scalar inference.

The study presented participants with a sentence containing an adjectival scalar and subsequent image (i.e., a 'true', 'false' or 'target' image. See below). Participants were tasked with indicating whether or not the sentence was a 'good' description for the image. While the study aims of the study were unrelated to that of the current study, the experimental results indicate that each of the sentences was considered a 'good' description of both the 'true' (i.e., enriched or *strong*) and 'target' (i.e., unenriched, or *weak*) images, but that the adjectival scalars were considered better descriptors for the strong images than the weak images for

each item set. In other words, the results indicate that the sentence and image combinations successfully elicited adjectival scalar implicature.




Sentence	<i>True</i>	<i>False</i>	<i>Target</i>
The door is ajar.			

Figure 6-5: Sentence and images for the item 'ajar' take from Pankratz & van Tiel (2021)

Consequently, I adapted van Tiel and Pankratz’s experimental items for use in the current study. I used 14 of the 16 tested items in van Tiel and Pankratz (2021) and adapted the target sentences to make them compatible with the syntactical structure of the ‘Alan’ captions. The 14 items used in Experiment 5, their respective captions, and the Horn Scales established by the images are presented below.

Item	Caption	Horn Scale
Fair	<i>I see a target with fair accuracy</i>	< fair, complete >
Chubby	<i>I can see a chubby man</i>	< chubby, obese >
Content	<i>I can see a content lady</i>	< content, happy >
Cool	<i>I can see a glass of cool water</i>	< cool, cold >
Ajar	<i>I can see an ajar door</i>	< ajar, open >
Drizzly	<i>I can see some drizzly weather</i>	< drizzly, rainy >
Breezy	<i>I can see some breezy weather</i>	< breezy, windy >
Low	<i>I see a bath in which the water is low</i>	< empty, low >
Mediocre	<i>I can see a test with a mediocre grade</i>	< mediocre, good >
Passable	<i>I can see a test with a passable score</i>	< passable, bad >
Ripe	<i>I can see a ripe banana</i>	< ripe, overripe >
Scarce	<i>I can see a field in which red flowers are scarce</i>	< scarce, absent >
Sleepy	<i>I can see a sleepy woman</i>	< sleepy, asleep >
Unlikely	<i>I can see a spinner that’s unlikely to land on blue</i>	< unlikely, impossible >

Figure 6-6: list of Items, Captions and Horn Scales for the items used in Experiment 5

As discussed, each item had two corresponding images: one representing the pragmatically enriched interpretation of the sentence, i.e., in which the adjectival scalar implicature is

derived, and the other representing the alternative, or the unenriched interpretation of the image, i.e., the other constituent of the contextually established Horn Scale.

As discussed, the small number of pre-tested adjectival scalar implicatures used in van Tiel and Pankratz (2021) meant that only 14 unique prime trials were used across the entirety of Experiment 5 compared to the 42 used in Experiment 2c and Experiment 3. While motivated by necessity, the pre-testing of the adjectival scalars shows that they have relatively high and relatively uniform rates of enrichment (see van Tiel and Pankratz, 2021). Consequently, including fewer unique experimental items is less problematic than in the case of more variable implicature subtypes, e.g., MI. Due to the high availability of the adjectival scalars, the prime trials were unguided.

The adjectival scalar QI prime for the item ‘content’ is depicted in Figure 6-7 below, while a complete set of experimental items is presented in 9.7 and 9.8 :

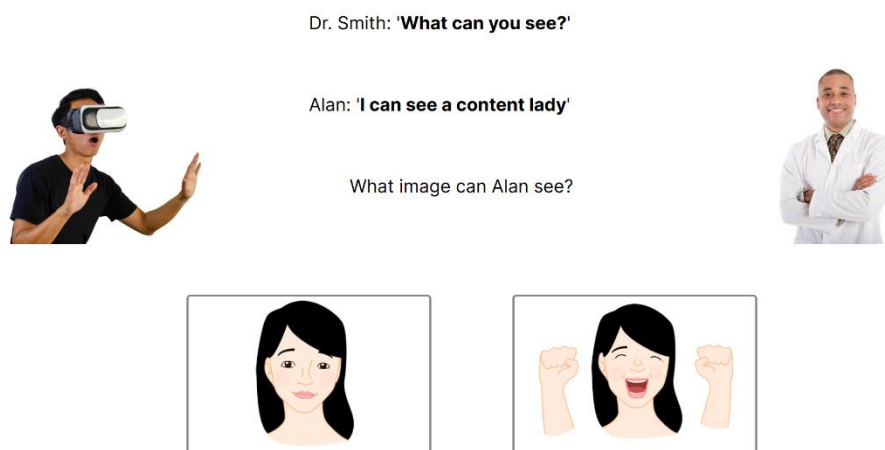


Figure 6-7: Prime trial for the item 'content' in Experiment 5.

6.3.3.3. Filler Trials

Experiment 3 used the same fillers as Experiment 2a, the ‘Yes’, No-Marked-Unmarked and No-False-False filler trials.

6.3.4. Procedure

The task experiment was hosted on Gorilla-TM (Gorilla, 2024) Gorilla Experiment Builder (www.gorilla.sc) and accessed by participants recruited via Prolific (www.prolific.com)

[2024]. After the collection of consent and demographic data, the participants were shown the following instruction screens, which were identical to those of Experiment 3.

Instruction Screen 1: *“Welcome to the task! In this task you will be given sets of two images and a caption. You will be tasked with matching the caption to one of the images - you will do this by using by clicking on the image”*

Instruction Screen 2: *“In this experiment, you will be presented with a fictional scenario involving Alan and a researcher. Alan is participating in a visual recognition experiment and has been tasked with telling the researcher what image he can see in the VR headset. The researcher will show Alan two images in the VR headset, one will then disappear. Alan is tasked with communicating which image stays, i.e., which image he can still see. Try not to think too hard about your answer!”*

6.3.5. Predictions

In light of the discussion presented in 6.1, I predict the following outcomes for Experiment 5:

- The adjectival scalar QI primes will increase the rate of MI derivation in succeeding critical trials.
- The priming effect will be of a different nature to Experiment 2’s MI-to-MI priming effect and will give rise to a smaller increase in the rate of MI in comparison compared to Experiment 2c.
- The priming effect in Experiment 5 will likely be larger than that of Experiment 3, too.
- The priming effect will be larger than that of Experiment 4.
- The priming effect will not be localized.

6.3.6. Coding

The selection of the strong image in Experiment 5’s critical trials was interpreted as evidence of MI derivation, while the selection of *strong* image, i.e., that compatible with the enriched utterance, was taken as indicative of adjectival scalar QI derivation in the prime trials.

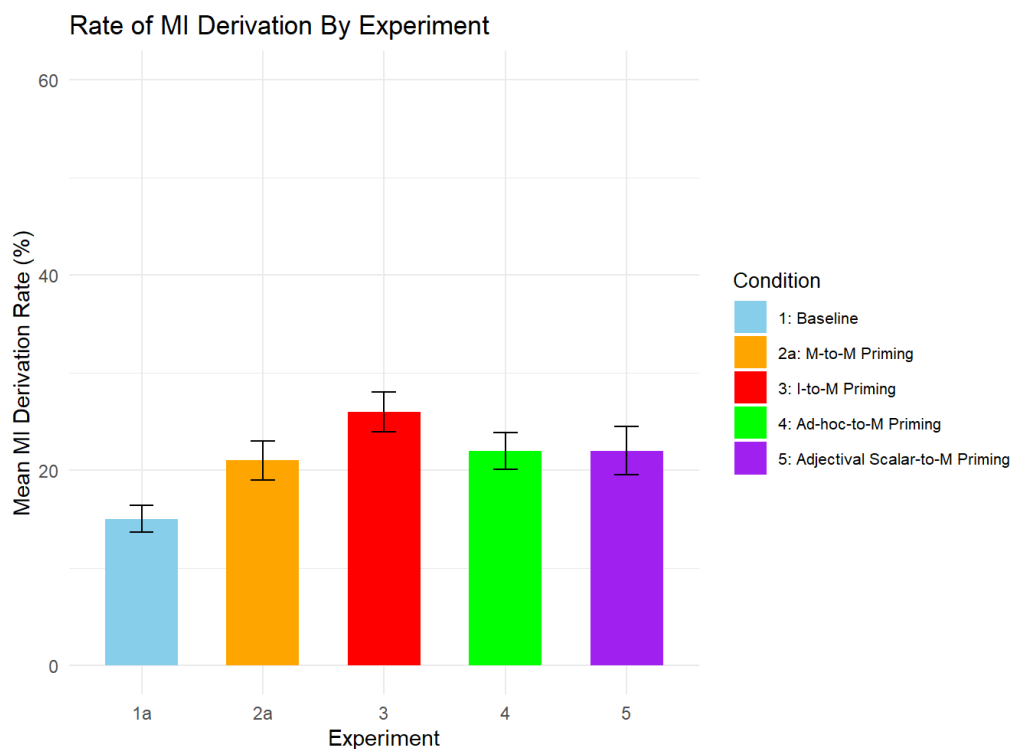
Responses were coded with a numerical score of 1 (i.e., ‘correct’) if participants selected the

image compatible with implicature derivation. Filler trials were coded in the same way as in previous experiments.

6.3.7. Results & Analysis

The rate of correct image selection was > 95% for each participant across the ‘Yes’ filler trials. Therefore, no participants were excluded from the analysis. The mean rate of marked image selection in the critical trials of Experiment 5 was 22%. In the lexical scalar QI primes, the rate of strong image selection was 84%. The rate of correct image selection was 96% for the ‘Yes’ trials, 62% for No-Marked-Unmarked fillers, and 50% for the No-False-False fillers. The rate of marked image selection increased from the baseline (15% → 21%) and was comparable to the rate of marked image selection across other experiments (Exp. 2c: 21%, Exp. 3: 26%, Exp. 4: 22%). The mean rates of MI derivation across Experiments 1a, 2a, 3, 4 and 5 are presented in Figure 6-8 below.

The mean rate of MI derivation in the critical trials varied between experimental items ($r=0.68$, $sd=0.16$). The item ‘purse’ had the lowest rate of MI derivation at 0% (baseline = 0%), while the item ‘cowboy boot’ had the highest at 68% (baseline = 48%). The difference between the baseline and primed rates of MI varied by item ($r=0.38$, $sd=0.11$), with the



smallest ‘priming’ effect (or largest ‘negative priming’ effect) observed in the item ‘dress’ at -10% and the largest in ‘hat’ at 28%.

6.3.7.1. *Primed Subset*

The rate of ‘correct’ responses to the lexical scalar primes was relatively high, with the strong image selected in 84% of the prime trials. Of 2079 critical trial observations, 1512 were immediately preceded by two ‘correct’ filler trials. Removing the critical trials that were ‘unprimed’ from the overall dataset did not lead to a significant numerical increase in the mean rate of MI across the trials (22% → 22%), suggesting that the priming effect is spread across both ‘primed’ critical trials and those immediately preceded by either one or two ‘incorrect’ responses.

6.3.7.2. *The effect of the primed condition*

A generalized linear mixed model using a binomial distribution and logit link function (logistic regression) was fitted to the full dataset, with condition (Primed: *Lexical Scalar* vs. Baseline) as a fixed effect. The model included a random intercept and slope for the experimental item and a random intercept for the participants. The overall model was significant ($\beta = -2.43, p < .001$), with a significant main effect of condition ($\beta = 0.59, p = .005$).

6.3.7.3. *Comparison to previous experiments*

I conducted a pairwise comparison using the logistical regression models used to analyse the main effects of each experiment to assess the difference between the likelihood of MI derivation in Experiment 5 compared to Experiment 4, 3 and 2c, as presented in below – the results are reported with a Tukey adjustment for multiple comparisons.

<i>Contrast</i>	Estimate	SE	z	p
Lexical Scalar vs. Ad-Hoc QI (Exp. 4)	-0.07	0.21	-0.3	1.0
Lexical Scalar vs. I-Implicature (Exp.3)	0.15	0.22	-0.69	0.96
Lexical Scalar vs. MI (Exp. 2a)	0.06	0.2	0.30	1.0

As presented above, the results indicate that there was no significant difference between the

Table 5: results of pairwise comparison between the logistical regressions used to model the main effect of condition in Experiments 5 and Experiments 4, 3 and 2a

likelihood of priming between Experiment 5 and Experiments 4, 3 and 2c.

6.3.7.4. Item Analysis

To investigate if the variation in the size of the priming effect in Experiment 4 varied between experimental item in a similar pattern to that of Experiment 3 and Experiment 2c, I fitted a series of linear regressions to see whether the magnitude of the by-item priming effects in each predicted the magnitude of the priming effect in Experiment 4.

The magnitudes of the by-item priming effects in Experiment 4 (*Ad-Hoc QI*) and Experiment 2c (*MI*) were significant predictors of the magnitude of the by-item priming effect in

	Estimate	SE	z	p
By-item magnitude of priming effect (Exp.4)	0.54	0.18	3.05	0.004
By-item magnitude of priming effect (Exp.3)	0.16	0.17	0.96	0.34
By-item magnitude of priming effect (Exp. 2a)	0.58	0.17	0.82	0.001

Table 6: results of linear regressions modelling whether the by-item magnitude of the priming effect in Experiment 2c & 3 & 4 predicts that in Experiment 5.

Experiment 5, while that of Experiment 3 (*I-Implicature*) was not. These findings suggest some similarity between the priming effects in Experiment 5 and those of Experiment 4 and 2c but not Experiment 3.

I conducted a second linear regression to examine the relationship between the by-item baseline rates of MI derivation and the by-item priming effect magnitude in Experiment 5; the by-item baseline rate of MI derivation was a significant predictor of the magnitude of the priming effect by experimental item ($\beta = 0.39, p < .003$). Given the relatively small size of the priming effect, I did not run any item-specific analyses.

6.3.8. Discussion Experiments 4 & 5

6.3.8.1. Overall priming effect: Experiment 5

I interpret the increase in the rate of MI derivation in the critical trials of Experiment 5 to indicate the priming of MI by the adjectival scalar primes. As indicated by the pairwise comparison (Table 6), there is no statistically significant difference between the likelihood of MI derivation in Experiment 5 and Experiments 2c, 3 and 4. This indicates that the priming effect may be of a similar nature in Experiment 5 and Experiments 2c, 3, and 4. Considering that the difference in the rate of MI derivation between the ‘primed’ and ‘unprimed’ trials is very small, I conclude that the priming effect in Experiment 5 is not localized to the ‘primed’ trials.

6.3.8.2. Locus of the priming effect

Experiments 4 and 5 suggest that the priming effect is not localized to the critical trials immediately preceded by two ‘correctly’ answered primes. In light of Experiment 3’s findings, I considered various explanations for a non-localized priming effect and ultimately concluded that a non-localized priming effect is likely the result of the priming of a higher-order, more generic mechanism. This conclusion is compatible with the non-localized priming effect observed in Experiments 4 and 5. Given that the priming effects in Experiments 3 – 5 are similarly non-localized, this may suggest that the primed mechanism is of a similar nature, e.g., higher-order and not specific to MI.

6.3.8.3. Shared and distinct mechanisms

According to the Gricean-inspired account, MI and QI are both types of EVA, as I outlined in The Similarities and Differences Between QI and MI in the Gricean-Inspired Perspective.2.3.6, Bott and Chemla (2016) reportedly primes the level of implicature sub-computation that is specific to QI but not shared with other types of purported EVA. Although Experiments 4 and 5 both involve EVA primes, there is no difference between the likelihood of MI derivation in the critical trials of Experiments 4 and 5 compared to Experiment 3. There are two possible explanations for this. Either the type of mechanism primed in Experiments 4 and 5 is not a mechanism specific to the computation of EVAs, e.g., is neither

responsible for the generation of the alternative nor its use, or the mechanism primed is EVA-specific, which is not compatible with the Gricean-inspired perspective of I-Implicature.

Similarly, there is no significant difference between the degree to which the ad-hoc QI trials and the adjectival scalar trials increase the likelihood of MI derivation. As discussed, the adjectival scalars are more similar in nature to MI than the ad-hoc QI trials are, given that they rely more heavily on context and are likely to entail a greater degree of calculation, according to the Gricean-inspired account. Considering this, MI and adjectival scalars are likely to share more levels of sub-computation than MI and ad-hoc QI, given that the latter is considered to be partially lexicalized and generated by default, by some perspectives (e.g., Horn, 1989; Levinson, 2000). However, there is no evidence to suggest that the adjectival scalars confer a larger priming effect than the ad-hoc QIs. Depending on the nature of the mechanism activated by the priming effect, this observation could have different interpretations.

For instance, if the priming effect is the result of the residual activation of a shared but generic, higher-order implicature-deriving mechanism, then the differences between the sub-computational mechanisms responsible for adjectival scalars and ad-hoc QI may be irrelevant. Nevertheless, this would suggest that the ad-hoc QIs involve some level of non-default, on-line computation— although, this is easily accounted for within existing theories. On the flip side, the priming effect may involve an EVA-specific mechanism.

The *combination* and *salience* models offer different perspectives on this mechanism. In the former, the *usage mechanism* is a separate sub-computational level that is responsible for the use of an alternative, i.e., its negation and subsequent combination with the articulated utterance. In contrast, the salience model posits implicature derivation to occur automatically, providing an alternative is sufficiently salient. According to this model, priming can only occur at the level of the alternative (Rees and Bott, 2018). If priming occurs at the level of the usage mechanism, then it follows that MI should be primed equally by different EVAs, regardless of the nature of their alternative. In contrast, if MI is primed at the level of the alternative, EVAs that involve similar types of alternatives, e.g., those that are contextually defined, may confer larger priming effects based on the assumption that shared mechanisms process them.

As described, the alternative in ad-hoc QI and adjectival scalar implicature differs in the degree to which it is context-dependent or lexicalized. While this may suggest that priming

occurs via the usage mechanism, it is unclear whether the alternative in ad-hoc QI and adjectival scalar generation is different enough to warrant a different search mechanism for each – after all, the alternative is at least partially context-dependent in both. That said, if priming occurs via the alternative, this explanation must accommodate the unexpected priming effect observed in Experiment 3. The nature of the primed mechanism and its ramifications for models of implicature derivation will be discussed in greater detail in 7.2.1

6.3.8.4. *Within- and between-category priming effects*

The priming effects observed in Experiments 4 and 5 appear to be fundamentally different to those of Experiment 2 but are similar to those in Experiment 3 – they are not localized and confer a far smaller degree of priming when answered ‘correctly’. This supports the idea that the difference in priming effects between Experiment 2 and Experiments 3 – 5 can be attributed to a distinction between within-category and between-category priming effects rather than to a more nuanced difference, given that the MI critical trials seem to be primed by non-MI prime trials in a similar manner, i.e., to around 21%. In turn, this may suggest that the between-category priming shows a ceiling effect, indicating that priming not involving MI-specific mechanisms may only have a small effect.

6.3.8.5. *Contextual adaptation, priming-via-the-alternative & contextual modulation*

In previous chapters, I explored alternative explanations for the priming effect observed across the studies beyond an activation-based account. One such explanation is Marty et al.’s (2022) theory of context adaptation. As previously noted, this perspective offers a framework for understanding the priming effect as the result of the prime trials updating the listener expectations of the conversational context, leading to the increased expectation of MI in the succeeding trials, which increases the rate of MI derivation, in turn. As discussed, the degree of context adaptation conferred by a stimulus is related to the degree to which it is unexpected – in short, unexpected constructions lead to listeners adapting their expectations.

Assuming that ad-hoc QI is at least partially lexicalized and that adjectival scalars are context-dependent, it follows that the latter is more unexpected, owing to their comparative infrequency. Considering this, an account of context adaptation likely predicts that the

adjectival scalars should confer a larger priming effect than ad-hoc QIs on MI derivation. Moreover, the adjectival scalar primes would likely contribute adaptations to the expectations of the conversational context that are more relevant to MI derivation, e.g., that the speaker seeks to elicit particularised, context-dependent implicature. However, there is no evidence to suggest that either condition confers a larger priming effect. Consequently, I do not consider the results of Experiment 4 and 5 to support an account of context adaptation.

As outlined in 5.1.9.6, the nature of the I-Implicature primes means that they may increase the salience of MI's alternative and contribute to the contrast between the prime and critical trials that renders MIs more obvious. However, there is no reason why the adjectival scalar primes and the ad hoc QI primes should increase the salience of a MI-specific alternative too – if the I-Implicature primes had, hypothetically. Moreover, Experiment 4 and 5's primes are not diametric to the MI critical trials in the same way as the I-Implicature primes and, therefore, do not generate the same degree of contrast. Given that there is no evidence that the primes in Experiment 4 and 5 confer a smaller priming effect than those in Experiment 3, I conclude that the mechanism of priming is unlikely to involve the salience of the alternative or a modulation of the conversational context.

6.3.8.6. *Nature of QI & MI*

Different perspectives on the nature of QI predict varying degrees of similarity with MI in terms of the mechanisms responsible for implicature derivation. Resultingly, different between-category priming effects are predicted between different types of QI and MI. Further, these predictions vary depending on the level at which the derivational mechanism is primed, as some levels of sub-computation may be shared but not others, as outlined above.

6.3.8.7. *The Gricean-inspired perspective*

Broadly, the Gricean-inspired perspective considers MI and QI to be governed by the exact derivational mechanism. The existence of a between-category priming effect between QI and MI in Experiments 4 and 5 suggests that there is a shared, or partially shared, mechanism responsible for the derivation of QI and MI and that this mechanism may be EVA-specific, as discussed. In turn, the evidence from Experiments 4 and 5 supports the Gricean-inspired perspective that MI is a type of EVA. Rett (2022) focuses on types of MI that arise from the

Maxim of Manner's third clause, 'Be brief or orderly' and concludes that these types of MI have more in common with QI than originally accounted for by Grice (1975) and the existence of a priming effect between the two lends support to this account.

Assuming that the priming effect in Experiments 4 and 5 targets an EVA-specific mechanism, the lack of evidence to suggest that the magnitude of the priming effects is different between the ad-hoc QI primes and adjectival scalar primes is most compatible with Grice's (1975) account of implicature, alongside other Pragmatic Theories of QI derivation (e.g., Russel, 2006; Guerts, 2009). According to Grice, all three are particularised conversational implicature and therefore require on-line, non-default calculation. As such, ad-hoc QI and adjectival scalar QI are likely derived by the same set of sub-computations. Considering this, their derivation is likely to be equally similar to that of MI and, therefore, likely to elicit similar priming effects – this is supported by the similar priming effects across Experiment 4 and 5.

In contrast, for both Horn (1972; 1989) and Levinson (2000), there are notable differences between ad-hoc QI and adjectival scalars. For Horn, the former is the product of a generalised Horn Scale, and therefore, the alternative is not context-dependent to the same degree as that of adjectival scalars. However, both do require on-line calculation, although this is considered to occur by default. From Horn's perspective, while the same overarching mechanism calculates both implicatures, some levels of sub-computation may differ and, therefore, may be shared with MI derivation or not, e.g., the mechanism responsible for the retrieval of context-dependent alternatives. Given this, Horn's account likely predicts a larger priming effect in Experiment 5 than in Experiment 4, which is not supported by the results of either experiment.

On the other hand, Levinson (2000) considers high-frequency QI to be automatically derived. Considering this, the Horn Scale $\langle 'x', 'x' \text{ and } 'y' \rangle$ likely incorporated into the lexical representation of the determiner 'a' and the corresponding implicature '+> only an 'x'' the default interpretation of an ad-hoc scalar. Although ad-hoc scalars do require some level of contextual input, this likely applies after the automatic enrichment of the scalar term. Levinson considers non-high-frequency adjectival scalars and MI to be particularised implicatures. Consequently, their derivation requires on-line calculation. If ad-hoc QIs are derived automatically or partially so, then the derivational mechanism is likely less activated or activated at some levels of sub-computation but not others. As a result, the ad-hoc QI

primes should elicit little or no residual activation compared to the adjectival scalar primes. However, this does not predict a similar priming effect to that observed in Experiments 4 and 5.

Of the accounts comprising the Gricean-inspired perspective, Grice's (1975) original account is most compatible with the results of Experiments 4 and 5. This may suggest that the mechanisms that ad-hoc QIs and adjectival scalars share are the same mechanisms they share with MI derivation. If the shared mechanism(s) are of a more fundamental nature, this may suggest that ad-hoc QI is indeed calculated in a similar fashion to adjectival scalar implicature and MI.

6.3.8.8. *The Grammatical Account*

The Grammatical Account (e.g., Fox, 2007; Chierchia, 2006) considers QI derivation to be the product of a grammatical operator. Considering this, the Grammatical Account does not predict a between-category priming effect between QI and MI, given that the two are derived by very different mechanisms. Nevertheless, Experiment 4 and 5's findings are not explicitly incompatible with the Grammatical Account as the perspective considers pragmatic information to play a role in the interpretation of QI, e.g., for suppressing the enrichment in irrelevant contexts, say. Considering this, the priming effect could be explained as the result of a mechanism of QI interpretation rather than one of QI derivation. Nevertheless, the results of Experiments 4 – 5 seem to suggest that the mechanism responsible for MI and QI derivation is at least partially shared and that the derivation, or interpretation of ad-hoc QI and adjectival scalars, does have a significant pragmatic element. This finding is more compatible with the Gricean-inspired perspective of QI derivation than that of the Grammatical Account.

6.3.8.9. *Relevance Theory*

A crucial distinction between the Gricean-inspired perspective and that of Relevance theory is that the former proposes a modular account of conversational implicature²²(Fodor, 1983), while the latter proposes a unified account (Sperber and Wilson, 1995). Although different

²² The extent to which this account is modular is discussed in 7.3

mechanisms are responsible for the derivation of implicature in the Gricean-inspired perspective, Relevance Theory considers all conversational implicature to be derived by a unified, generic inferential mechanism, although this mechanism does perform different functions (e.g., implicature derivation, explicature derivation, disambiguation, a.o.) (Carston, 2006; 2002). Considering this, Relevance Theory likely predicts no difference in priming effect between Experiment 4 and 5, considering both ad-hoc QI and adjectival scalars are derived using the same mechanism – this is supported by findings of Experiments 4 and 5.

6.3.8.10. Conclusion

Overall, I interpret the results of Experiment 4 and 5 to indicate that ad-hoc QI primes and adjectival scalar primes can prime the derivation of MI in the succeeding critical trials. The priming effect appears to be non-localised, which may indicate that the priming effect likely targets a higher-order mechanism than that primed in Experiment 2.

The priming effects appear to be of a fundamentally different nature to that of Experiment 2, but there is no evidence to suggest that they are substantially different from one another and that of Experiment 3. I interpret this to suggest that the mechanisms primed in Experiments 4 and 5 are of a similar nature and may be a mechanism of sub-computation involved in the derivation of ad-hoc QI, adjectival scalars and MI – this mechanism may be EVA-specific. I interpret the results of Experiments 3 – 5 to suggest that this mechanism may be shared with I-Implicature, too. However, this necessitates some clarification concerning the nature of I-Implicature, as outlined in Section 7.4.1.3.

Assuming that the priming effect is activation-based, I have detailed possible levels of sub-computation by which the priming effect may occur – these have different ramifications for the nature of QI, QI-subtypes and MI. Ultimately, I conclude that the priming effect is likely of a higher order than that of the MI-to-MI priming effect. However, the nature of this mechanism is discussed further in 7.3 and may take different forms.

I consider the results of Experiments 4 and 5 to be most compatible with the Gricean-inspired perspective's account of conversational implicature derivation, although they are not incompatible with accounts like the Grammatical Account and that of Relevance Theory.

6.4. Experiment 6: Scalar Ambiguity

6.4.1. Motivations

There is a possibility that the priming effect observed in Experiments 2–5 is not the result of activation-based, mechanistic priming but instead an artefact of the experimental paradigm. For example, because a higher proportion of trials in Experiments 3–5 do not involve MI compared to Experiment 2 and Experiment 1a, prolix expressions may be relatively more marked in Experiments 3–5. In turn, this could increase the rate of MI derivation in the critical trials, as participants may be more attuned to these marked expressions due to the structure of the experiment rather than genuine pragmatic. Other contextual factors, e.g., QUDs, perceptions of the speaker’s communicative strategies, and interpretative cues from the task itself, may also influence the participant’s responses in the succeeding critical trials.

To address these concerns, Experiment 6 uses ‘scope ambiguity’ primes that do not involve any pragmatic enrichment. This provides a baseline to assess the impact of the prime trials in Experiments 3 – 5 involving different types of conversational implicature. This approach aligns with the methodologies of other studies that pursue similar lines of investigation (e.g., Chemla and Spector, 2011; Breheny et al., 2006; Bott and Noveck, 2004), where null or control conditions are used to ensure that the observed effects are attributable to the manipulated effects, rather than the experimental paradigm more broadly.

6.4.1.1. Theoretical Background – Scope Ambiguity

Take the following sentence:

‘Every participant completed a block of trials’.

Example 6-2

Here, there are two possible interpretations. On one hand, the utterance may describe a situation (Situation 1) in which one, single trial block is completed by all participants. On the other hand, the utterance may describe a situation in which each participant completed a

different block of trials. This phenomenon is known as *scope ambiguity* and arises when there is conflict between the scope of two different operators.

In Example 6-2, if scope is assigned to the quantifiers in the order in which they appear, ‘every’ may be assigned scope over ‘a’ – this is known as *surface scope* (Situation 1). Conversely, *inverse scope* (Situation 2) allows for the ‘a’ to be assigned scope over the proposition. Although context can privilege one interpretation over another, scope ambiguity is not considered a pragmatic phenomenon or a type of implicature (Ruys and Winter, 2010; Musolino et al., 2000; Ioup, 1975).

6.4.2. Participants

Participants (N = 159) were recruited via Prolific (www.prolific.com) [2024]. All participants were adults (≥ 18 years old) and self-identified English speakers who had not previously participated in any of the study’s experiments. Incomplete data was rejected from N = 11 of the participants; therefore, data was collected from N = 148 participants.

6.4.3. Design

The design of Experiment 6 was the same as that of previous experiments. Like Experiment 4, Experiment 6 involved 42 different critical trials and 28 different prime trials. While the 42 critical trials were the same as those used in previous experiments, the 28 prime trials involved scope ambiguity primes. The prime trials were modelled after those used in Guerts and Pouscoulous (2008; 2009) and featured cards with shapes connected to one another in different configurations by lines – as described below. The experimental design was between-participants and each participant responded to 14 primed critical trials, 28 prime trials and 56 filler trials, therefore 98 trials in total.

Participants were randomly assigned to one of three different trial clocks – each contained 14 unique critical trials, the same 28 prime trials and the same 56 filler trials. The trials were presented in the following order: *filler* → *filler* → *filler* → *prime #1* → *prime #2* → *critical trial*, and the specific experimental items in each configuration was completely randomized as too was the left-right orientation in which the images appeared.

6.4.4. Materials

6.4.4.1. Critical Trials

The critical trials were identical to those of previous experiments and consisted of the following components: 1) a caption reading “**Dr Smith: ‘Can you see a [item label]?’**” which contained one of the 42 item labels presented e.g., ‘t-shirt’, 2) a second caption, reading ‘Alan: **‘Yes, I [/can] see [...]**’ and contained a prolix description 3) the marked and unmarked images representative of the mentioned object and 4) static images of both ‘Alan’ (a man in a black shirt wearing a VR headset) and ‘Dr Smith’ (a man dressed in a lab coat). See the critical trial for the item ‘tuxedo’ below:

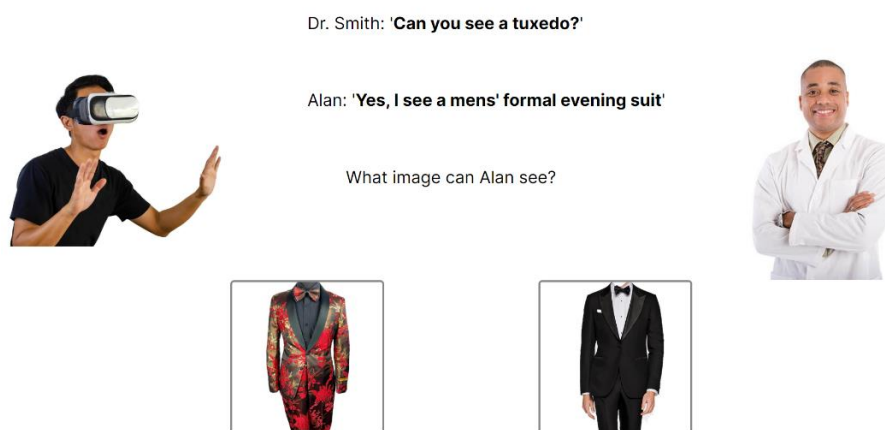


Figure 6-9: MI critical trial for the item 'tuxedo'

6.4.4.2. Prime Trials

The prime trials were similar to those used in previous studies and involved the following components: 1) a caption reading “**Dr Smith: ‘What image can you see?’**” 2) a second caption, reading ‘Alan: **‘I [/can] see an image in which every [/each] [object] is connected to a [object]**’ 3) two images, one compatible with interpretation of the caption as having surface scope and one compatible with a semantic interpretation of the caption as having inverse scope 4) static images of both ‘Alan’ (a man in a black shirt wearing a VR headset) and ‘Dr Smith’ (a man dressed in a lab coat).

The objects mentioned in the captions were fruits or shapes and were sometimes modified with an adjective, e.g., ‘blue circle’ or ‘apple’ (a full list is presented in 9.9 and 9.10). The images that represented the interpretation of the caption with surface scope involved 6 objects, 3 of each type mentioned in the caption. Each of the 3 shapes was individually connected to one of the 3 other shapes by a black vertical line (see below). On the flip side, the images representing the interpretation of the caption with inverse scope involved 3 instances of the first-mentioned object and 1 instance of the second-mentioned object – all 3 instances of the former were connected with by a black line to the single second-mentioned object.

As mentioned, the images represented different interpretations of the scope of the operators in the captions and therefore elicited scope ambiguity. There were no contextual clues that dictated which interpretation was preferred and therefore participants could choose freely between the two images. Given this, I did not pre-test the prime images to ascertain which scope interpretation was preferred, as this is largely irrelevant – both surface scope and inverse scope are thought to be interpreted using non-pragmatic means. That said, the natural ambiguity of the prime trials may warrant a consideration of conversational context more generally, which allows for the role of such considerations in the priming effect in Experiments 2 – 5 to be assessed. Note, any pragmatic resolution of scope ambiguity is not a function of conversational implicature.

Given that the choice of scope interpretation is not thought to be important, the primes were unguided.

The adjectival scalar prime for the objects ‘blue circle’ and ‘purple circle’ is depicted in Figure 6-10 below:

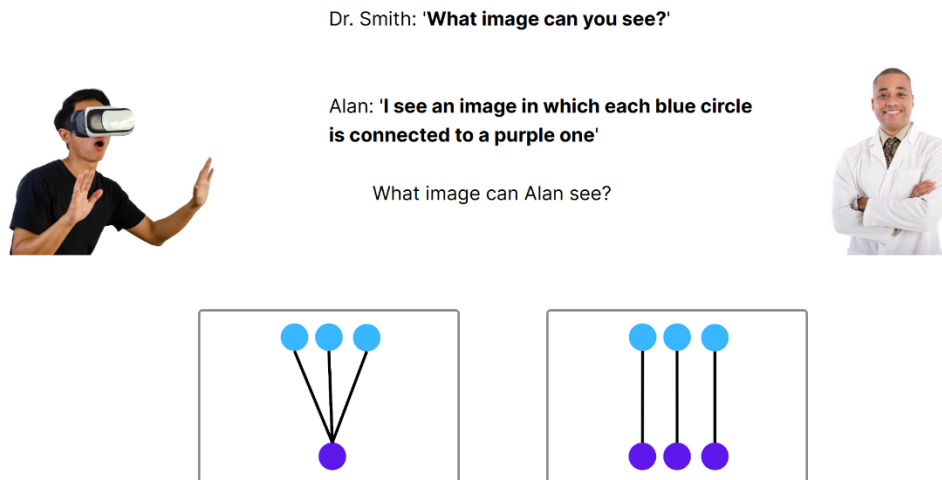


Figure 6-10: scope ambiguity prime trial for the item 'blue circle' in Experiment 6

6.4.4.3. Filler Trials

Experiment 3 used the same fillers as Experiment 2a, the 'Yes', No-Marked-Unmarked and No-False-False filler trials.

6.4.5. Procedure

The experiment was conducted using the Gorilla Experiment Builder (www.gorilla.sc) and involved participants recruited through Prolific (www.prolific.com) [2024]. Following the collection of consent and demographic information, participants were presented with instruction screens identical to those used in Experiments 3 – 5.

Instruction Screen 1: *“Welcome to the task! In this task you will be given sets of two images and a caption. You will be tasked with matching the caption to one of the images - you will do this by using by clicking on the image”*

Instruction Screen 2: *“In this experiment, you will be presented with a fictional scenario involving Alan and a researcher. Alan is participating in a visual recognition experiment and has been tasked with telling the researcher what image he can see in the VR headset. The researcher will show Alan two images in the VR headset, one will then disappear. Alan is*

tasked with communicating which image stays, i.e., which image he can still see. Try not to think too hard about your answer!”

6.4.6. Predictions

I make the following predictions for Experiment 6:

- The scope ambiguity primes will not lead to an increase in the rate of MI derivation in the succeeding critical trials.

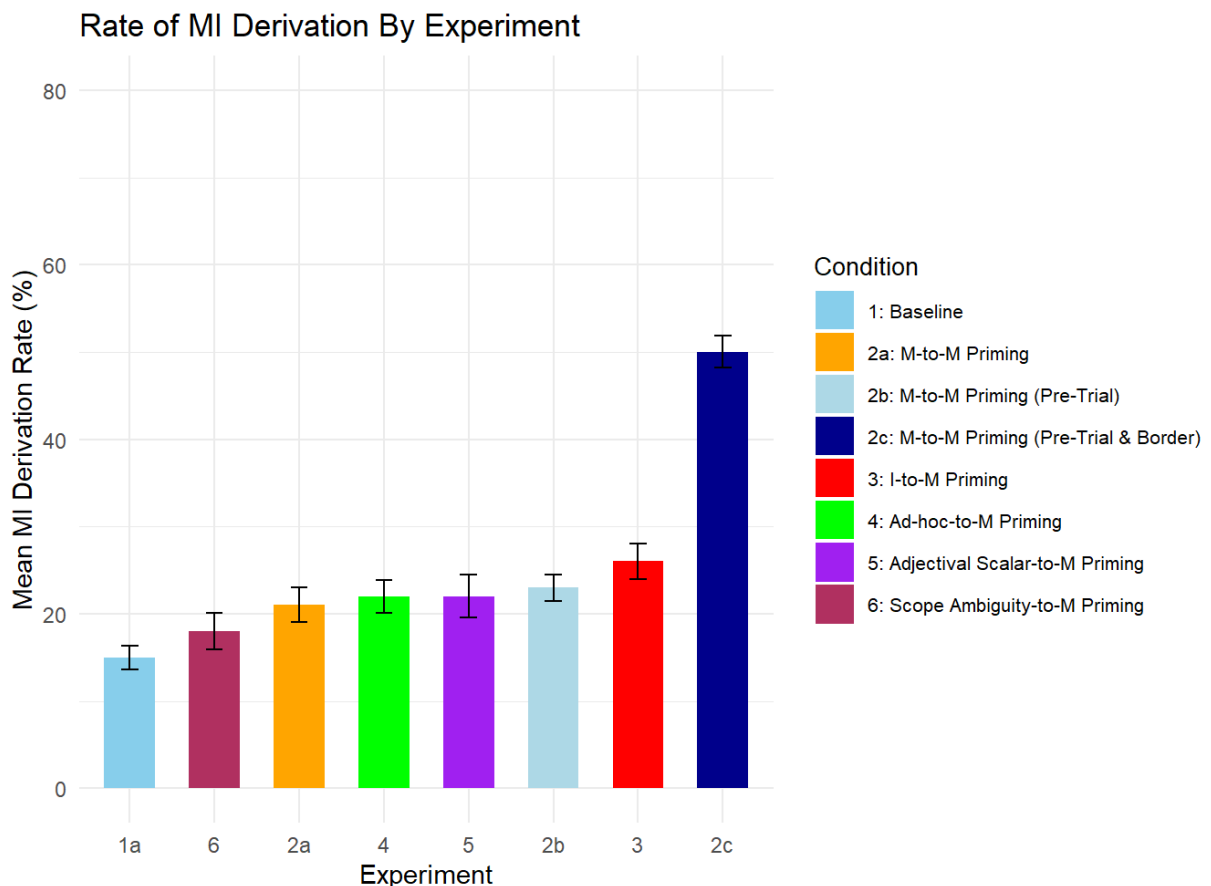
6.4.7. Coding

The selection of the marked image was interpreted to indicate MI derivation in Experiment 6's critical trials and coded as 'correct'. The filler trials were coded in the same way as previous experiments, and the selection of the image compatible with an interpretation of inverse scope was coded as 'correct' – this choice was random.

6.4.8. Results & Analysis

The rate of correct image selection was > 95% for each participant across the 'Yes' filler trials. Therefore, no participants were excluded from the analysis. The mean rate of marked image selection in the critical trials of Experiment 6 was 18%. In the primes, the rate of inverse scope selection was 35%. The rate of correct image selection was 95% for the 'Yes' trials, 62% for No-Marked-Unmarked fillers, and 49% for the No-False-False fillers. The rate of marked image selection increased from the baseline (15% → 18%).

The mean rate of MI derivation in the critical trials varied between experimental items ($r=0.63$, $sd=0.13$). The item ‘glasses’ had the lowest rate of MI derivation at 0% (baseline = 14%), while the item ‘cowboy boot’ had the highest at 63% (baseline = 48%). The difference between the baseline and primed rates of MI varied by item ($r=0.38$, $sd=0.08$), with the smallest ‘priming’ effect (or largest ‘negative priming’ effect) observed in the item ‘dress’ at -16% and the largest in ‘flipflops’ at 23%. Below, Example 6-3:, depicts the ordered mean rates of MI across Experiments 1a – 6:



Example 6-3: The mean rate of MI derivation across Experiments 2c, 2a, 3, 4,5 and 6, ordered by magnitude, with error bars standard error of the by-item mean.

6.4.8.1. The effect of the primed condition

A generalized linear mixed model using a binomial distribution and logit link function (logistic regression) was fitted to the full dataset, with condition (Primed: Scope Ambiguity vs. Baseline) as a fixed effect. The model included a random intercept and slope for the experimental item and a random intercept for the participant. The overall model was

significant ($\beta = -2.41$ $p < .001$), while the main effect of the condition was not a significant predictor of the likelihood of MI derivation ($\beta = 0.35$, $p = 0.12$).

In line with recent studies (e.g., Rees, Carter and Bott, 2023) and in accordance with recommendations concerning null hypothesis testing (Dienes, 2014; Lee and Wagenmakers, 2014; Wetzels & Wagenmakers, 2012), I used Bayes Factors to interpret the non-significant findings of Experiment 5. The Bayesian logistic regression model was fitted with condition (Primed: Scope Ambiguity vs. Baseline) as a fixed effect and used with a Cauchy prior ($r =$), a weakly informative prior, as per recommendations for default priors in logistic regression models (Gelman et al., 2008) and was calculated using JASP (JASP Team, 2024).

The analysis provided weak evidence against the null hypothesis (Primed: Scope Ambiguity vs. Baseline) ($BF_{01}=0.41$), which indicated that the data was 2.44 times more likely under the alternative model than the null model. The posterior mean estimate for the effect of the condition was $\beta=0.11$ (95% credible interval: [0.000, 0.251]), which corresponded to an odds ratio of 1.11 (95% CI: [1.00, 1.28]). However, the model explained only 0.1% of the variance ($R^2= .001$). Although the model provides weak evidence to suggest a small increase in the likelihood of a correct response in the Scope Ambiguity condition, the practical significance of this effect size small and uncertain.

6.4.8.2. Comparison to previous experiments

I conducted a pairwise comparison using the logistical regression models used to analyse the main effects of each experiment to assess the difference in the likelihood of MI derivation in Experiment 6 compared to Experiments 4, 3, and 2c. The results indicate that there were no significant correlations between the likelihood of MI derivation in Experiment 6 and the priming effects between Experiment 5 and Experiments 4, 3, and 2c.

<i>Contrast</i>	Estimate	SE	z	p
Scope Ambiguity vs. Lexical Scalar (Exp. 5)	0.22	0.16	1.9	0.91
Scope Ambiguity vs. Ad-Hoc QI (Exp. 4)	0.22	0.22	1.01	0.91
Scope Ambiguity vs. I-Implicature (Exp.3)	0.42	0.22	1.92	0.4
Scope Ambiguity vs. MI (Exp. 2a)	-0.22	0.21	-1.01	0.91

Table 7: results of pairwise comparison between the logistical regressions used to model the main effect of condition in Experiments 6 and Experiments 5, 4, 3, and 2a.

6.4.8.3. Item Analysis

To investigate if the variation in the size of the priming effect in Experiment 4 varied between experimental items in a similar pattern to that of Experiment 3 and Experiment 2c, I fitted a series of linear regressions to see whether the magnitude of the by-item priming effects in each predicted the magnitude of the priming effect in Experiment 4.

	Estimate	SE	z	p
By-item magnitude of priming effect (Exp. 5)	0.03	0.12	2.95	0.005
By-item magnitude of priming effect (Exp. 4)	0.21	0.16	1.33	0.19
By-item magnitude of priming effect (Exp.3)	0.35	0.13	2.81	0.008
By-item magnitude of priming effect (Exp. 2a)	0.4	0.14	2.81	0.007

Table 8: results of linear regressions modelling whether the by-item magnitude of the priming effect in Experiment 2c, 3, 4 and 5 predicts that in Experiment 4

The magnitudes of the by-item priming effects in Experiment 4 (*Ad-Hoc QI*) and Experiment 2c (*MI*) were significant predictors of the magnitude of the by-item priming effect in Experiment 5, while that of Experiment 3 (*I-Implicature*) was not. These findings suggest some similarity between the ‘priming’ effects in Experiment 5 and those of Experiment 4 and 2c, but not Experiment 3.

I conducted a second linear regression to examine the relationship between the by-item baseline rates of MI derivation and the by-item priming effect magnitude in Experiment 6; the by-item baseline rate of MI derivation was not a significant predictor of the magnitude of the priming effect by experimental item ($\beta = 0.16, p < .33$).

6.4.9. Discussion

Although there is a slight numerical increase in the rate of marked image selection in Experiment 6, there is insufficient evidence to suggest that this increase is statistically significant. Even if it were, the practical significance of such an increase is minimal, as indicated by the statistical analyses. Nonetheless, given that the priming effects discussed in this thesis are generally small, the modest increase observed in Experiment 6 warrants acknowledgment.

While the scope ambiguity prime trials do not involve implicature derivation, they likely engage higher-order pragmatic mechanisms to resolve the ambiguity pragmatically – although this cannot be done. As a result, it is hypothetically possible that these trials contribute contextual factors that could slightly increase the likelihood of marked implicature

derivation. However, any such effect is clearly distinct from the priming effects observed in Experiments 2–5, which are larger and statistically significant. Furthermore, there remains insufficient evidence to confirm that the increase observed in Experiment 6 is statistically significant, and even if it were, it would lack practical significance.

Although the lack of a significant difference between the effects of Experiment 6 and Experiments 2 – 5 is unexpected if there is no significant increase in the likelihood of MI in Experiment 6 but there is said to be in Experiments 2 – 5, it should be noted that the nature of the pairwise comparison does not model differences in the likelihood of a binary outcome (e.g., MI or no MI derivation), but rather measures differences in effect size, and is furthermore insensitive to item-level variation. Consequently, it is likely that such a comparison lacks the appropriate sensitivity with which to discern the difference between the effects of Experiment 6 and Experiments 2 – 5, especially considering the size of the increase in MI derivation is small across all of the studies, which might preclude the observation of a difference in the likelihood of MI derivation between them.

7. Conclusions

7.1. Summary of key findings

The overarching purpose of this thesis was to investigate the nature of MI and the interrelation between MI and another type of conversational implicature. Experiment 1 provided two experimental baselines, one for MI and I-Implicature, using a novel implicature-eliciting paradigm inspired by Bott & Chemla (2016), a.o. Although the difference in the selection of the unmarked image in the MI baseline and I-Implicature baseline experiments (Experiment 1a: 15% vs. Experiment 1b: 13%) was not large, the difference was statistically significant and indicated that the experimental paradigm can elicit both MI and I-Implicature²³. Moreover, the results suggest that MI is a heterogeneous phenomenon, given the inter-item variation in the rates of MI derivation.

Experiment 2 investigated MI-to-MI within the category by having participants complete prime trials that elicited MI before the MI critical trials. Experiment 2a involved un-guided

²³ Recall that comparing Experiment 1a to Experiment 1b is imperfect as the same types of inferences are not available in each critical trial – see 4.4 for discussion. While I have compared them here, this is only to illustrate that the MI critical trials do lead to an increase in the selection of the marked image.

primes, which led to the majority of the prime trials being ‘incorrectly’ answered (as predicted by Experiment 1a). Nevertheless, there was an increase in the rate of MI in the critical trials of Experiment 2a compared to the baseline (15% → 21%). This effect was localised to the trials that immediately succeeded the ‘correctly’ answered primes.

Experiment 2b involved a pre-trial training block and updated instructions to increase the salience of possible MI elicitation in the critical trials; however, this only led to a slight numerical increase in the rate of MI derivation (Experiment 2a: 21% → Experiment 2b: 23%). While the increase between the baseline and Experiment 2b was statistically significant, the increase between Experiment 2c and 2b was not.

Experiment 2c took a more direct approach to increase the ‘correct’ prime answers rate and involved primes in which a green border indicated the ‘correct’, i.e., pragmatic, answer. This condition increased the rate of correct primes and significantly increased the rate of MI in the succeeding trials beyond that of the previous conditions (→ Experiment 2c: 50%).

Experiments 3, 4 and 5 investigated between-category priming using MI critical trials and prime trials with different types of conversational implicature to investigate the extent to which they shared a derivational mechanism with MI. Experiment 3 investigated I-Implicature-to-MI priming and found a significant increase in the rate of MI derivation (15% → Experiment 3: 26%). Experiments 4 and 5 investigated QI-to-MI priming using two types of QI with different degrees of context dependency. Experiment 4 involved ad-hoc QI and found a significant priming effect (15% → Experiment 3: 22%). Experiment 5 involved adjectival scalars and found a significant priming effect (15% → 22%).

To determine whether the increase in MI derivation in the critical trials of Experiments 2 – 5 was due to the priming of an implicature-specific mechanism or an artefact of the experimental paradigm, Experiment 6 used non-implicature (i.e., scope ambiguity) primes. While there was a small numerical increase in the rate of marked image selection (15% → Experiment 6: 18%), this effect was not of practical significance, and its statistical significance was unclear but unlikely.

The difference between the baseline rates and primed rates of MI per and their statistical significance per experiment are summarised below:

Table 9: Overview Of The Results Of Experiments 2a – 6.

Type of Prime	Experiment 1a: Baseline	Rate of MI derivation	Significance of the difference from baseline ²⁴
Exp. 2a: <i>MI</i>	15% →	21%	$p=0.012^*$
Exp. 2b: <i>MI</i>		23%	$p=<.001^{***}$
Exp. 2c: <i>MI</i>		50%	$p=<.001^{***}$
Exp. 3: <i>I-Implicature</i>		26%	$p=<.001^{***}$
Exp. 4: <i>ad-hoc QI</i>		22%	$p = 0.031^*$
Exp. 5: <i>adjectival scalars</i>		22%	$p = 0.005^{**}$
Exp. 6: <i>scope ambiguity</i>		18%	$p=0.12$

7.1.1. Key conclusions & Unanswered Questions

In the discussion of the individual results, I have made several key conclusions.

Overarchingly, I conclude that the experimental paradigm can elicit MI and I-Implicature – alongside other types of inference. As per the results of Experiments 2 – 5, I concluded that MI can be primed, both by other MI and by other types of implicature, namely I-Implicature and QI. Having weighed up several explanations for the priming effect, I conclude that the

²⁴ Significance code: 0 ‘***’, 0.001 ‘**’, 0.01 ‘*’, 0.5 ‘.’

priming effect is the result of the activation of a shared mechanism involved in the derivation of MI and QI and is not an artefact of the experimental paradigm (Experiment 6).

The paradigm elicits two separate priming effects. Firstly, there appears to be a within-category priming effect that operates locally and confers a larger priming effect than the between-category primes (Experiment 2). Secondly, there appears to be a between-category priming effect that is not localised and confers a far smaller priming effect than the within-category primes. The second type of priming effect may be more generic and may not involve levels of implicature sub-computation specific to MI. There is no evidence to suggest that the strength of the between-category priming effect varies significantly between implicature types (Experiments 3 – 5).

However, these conclusions leave many questions unanswered. For instance, the exact nature of the mechanism is unclear, which has ramifications for our understanding of the nature of both MI and other types of implicature.

Sections 7.2 - 7.4 seek to discuss these questions in more detail, addressing four main points:

- What is the nature of the primed mechanism?
- What does the priming effect tell us about the nature of MI?
- What does the priming effect tell us about the nature of I-Implicature and QI?
- What are the ramifications for theories of implicature derivation?

7.2. Implications for the nature of the priming effect(s)

As described, the nature of the priming effect is uncertain. In 3.4.1, I outlined various perspectives on priming. The perspective of priming assumed by this thesis is modelled on the view that priming is the product of residual activation. Typically, the structural priming literature conceives priming as the product of the residual activation of a linguistic structure or representation – see Branigan & Pickering (2017) for an overview. For instance, Raffray and Pickering (2010) and Bott and Chemla (2015) present studies that seek to prime the logical form representations of scopal interpretations (i.e., surface scope and inverse scope).

In contrast, Bott and Chemla (2016), Rees and Bott (2018) and subsequent studies (e.g., Meyer & Feiman, 2021; Rees, Carter & Bott, 2023) present investigations of implicature

derivation that assume a *mechanistic* perspective, whereby priming is considered the result of the residual activation of mechanism responsible for implicature derivation – or a sub-computation of it – rather than of a linguistic representation associated with the implicature²⁵. Furthermore, the idea that specific mechanisms that govern implicature derivation are shared or partially shared across different types of implicature is a core assumption of the current thesis. Despite this, such ‘mechanisms’ do not enjoy a great deal of conceptual clarity in the referenced literature and their qualities remain largely unexplained and unexplored.

The current thesis has assumed that a ‘mechanism’ is a domain-specific process of language comprehension (or implicature comprehension, especially) involved in processing related phenomena. The line of investigation also assumed that multiple mechanisms are involved in implicature comprehension and that they can be shared to different degrees across different phenomena, e.g., as concluded by Meyer and Feiman (2021). Moreover, I have assumed that a ‘mechanism’ is not an ad-hoc, specific function but rather an overarching cognitive process that is used for ad-hoc, specific functions, i.e., I have not considered the processing of *litotes*, say, to constitute a mechanism, but have considered this to be an ad-hoc function of a more generic mechanism. Furthermore, I have assumed that mechanisms may be more or less generic and, therefore, may exist in hierarchical relation to one another – this perspective is discussed below.

Evidently, the nature of any priming effect is dependent on the subject of its priming. Considering this, due attention should be given to the assumptions that underpin the concept of a derivational ‘mechanism’ in studies that seek to prime them. Although theoretical accounts of implicature derivation have presented various ‘mechanisms’ in great detail (e.g., Grice, 1975; Levinson, 2000; Horn, 1984, a.o.) and these mechanisms are discussed in referenced work (e.g., Rees & Bott, 2016; Meyer & Feiman, 2021), they should not be taken as models of cognition, *per se*²⁶. Considering this, the application of a psychologically defined phenomenon like ‘residual activation’ (or priming, generally) to linguistic

²⁵ That said, Bott & Chemla (2016) do present a representation account of how EVA-specific priming may occur.

²⁶ Recall, Levinson (2000) explicitly considers the process of implicature derivation to involve the coarticulation of various other mechanisms and the ‘steps’ of the ‘core account’ of implicature derivation are not necessarily discreet or chronological. Similarly, neither Grice (1975) nor Horn (1989; 1984) claim that these steps are.

mechanisms should warrant their redefinition in terms of cognition, such that there is a clearer image of what *cognitive* mechanisms exist in implicature derivation.

This thesis does not begin with a fixed definition of what a shared ‘mechanism’ is, owing to the fact that the nature of this mechanism is investigated in Experiments 2 – 6. Nevertheless, future research that involves mechanistic priming could benefit from examining the assumptions that underpin the concept of a derivational ‘mechanism’ and the extent to which it is considered cognitive, linguistic or both. Section 7.2 attempts to shed some light on how this mechanism may be structured.

7.2.1. Locus of the priming effect

As discussed, there appears to be a difference between the priming effect of Experiment 2 and the priming effects in Experiments 3 – 5. Considering this, the locus of the priming effect is unclear. In the discussion of the previous experiments, I outline possible loci of the priming effects in each experiment. To recap, Rees and Bott (2018) and Meyer and Feiman (2021) conclude that priming can occur at some levels of sub-computation without occurring in others, and it is concluded that a lack of priming effect may indicate that a mechanism cannot be accessed (e.g., Rees & Bott, 2018) or may not be shared between phenomena (e.g., Meyer & Feiman, 2018; Bott & Chemla, 2016). Bott and Chemla (2016) and the current study observed that the strength of a priming effect can vary depending on whether the priming effect is within or between category. Moreover, these effects can be localized or not localized to directly primed trials.

Considering this, the results of the current study pose the question: What mechanism(s) have been primed? To answer this question, I propose the following explanation. As discussed in Meyer and Feiman (2021), different types of enrichment may share different levels of sub-computation. Typically, the *search for the alternative* and the *usage mechanism* are thought to be two mechanisms of implicature computation (Meyer & Feiman, 2021). As described, Bott and Chemla (2016) consider priming via either as a possible explanation for the observed effect, while Rees and Bott (2018) conclude that priming may only occur via the alternative. Whether or not the direct priming of the usage mechanism is possible depends on whether a salience or combination model of implicature derivation is assumed.

Experiments 2 – 5 show two unique priming effects, one functions within category, is localized and is stronger, the other functions between category and is not localized, is weaker

and is relatively uniform between phenomena. This suggests that the priming effects are fundamentally different. Given that activation-based accounts appear to assume that ‘activation’ is homogenous, i.e., that mechanisms or representations do not have multiple co-existing levels or types of activation, it seems unlikely that the same mechanism of implicature derivation can be activated in *different* ways by different stimuli²⁷. As such, the difference between the priming effect in Experiment 2 and Experiments 3 – 5 is best explained by the perspective that different mechanisms have been primed in each.

Considering this, the findings of Experiments 2 – 5 are a better fit with accounts of implicature derivation that posit multiple levels of sub-computation that exist as independent mechanisms that can be individually primed, e.g., the combination model. In the case of MI derivation, Experiment 2’s primes may activate a mechanism specific to MI derivation. As per a combination model, this may be the mechanism responsible for the search for the alternative. In contrast to Rees & Bott’s (2018) findings, salient alternatives did not prime MI to the same degree as MI primes – compare Experiment 2 to Experiment 3. This may indicate that priming-via-the-alternative is mechanistic rather than representational, i.e., the mechanism itself must be activated rather than the representation. Alternately, the representation for ‘I-Implicature’ used in I-Implicature and in MI may be different – this contrasts with the Gricean-inspired perspective.

In sum, this perspective posits two loci for the priming effect observed in the current study. The first is at the level of the search for the alternative. The second is at the level of the usage mechanism. This perspective accommodates the divergent priming effects of Experiment 2 – 5 within existing theories of implicature derivation. Moreover, this perspective also accounts for the numerical increase in the rate of MI derivation in the ‘semi-primed’ trials of Experiment 2c (15% → 30%)²⁸, as both the search for the alternative and the usage mechanism are likely primed in tandem by the MI prime trials therefore delivering a localised and non-localised effect.

However, it remains unclear why priming the search mechanism would lead to a localised priming effect, while priming the usage mechanism may not. To answer this question, I draw

²⁷ While it is possible that different stimuli lead to more or less activation of a mechanism, it appears unlikely that a mechanism can be activated by different stimuli in a way that leads to fundamental differences in the priming effect (e.g., localised vs. non-localised) without this being indicative of either a different type of activation or priming via a different route.

²⁸ Statistical analysis has not ascertained if this is significant, however.

attention to the fact that the ‘incorrect’ MI prime trials and the ‘correct’ prime trials in Experiments 3 – 5 are of a fundamentally different nature. If implicature derivation is suppressed at the point of the search for the alternative, then the ‘incorrect’ MI primes involve the suppression of MI, which may lead to the suppression of subsequent MI – following the assumptions of an activation-based account. In contrast, the ‘incorrect’ I-Implicature and QI primes do not involve the suppression of MI, so may not lead to the same suppression of succeeding MI. From this perspective, the localised priming effect in Experiment 2 may be the product of the dynamic activation and suppression of the search mechanism. In contrast, if the usage mechanism is activated independently, it may not be suppressed by the ‘incorrect’ trials in the same way. Although this is a speculative explanation, it does account for the differences between the observed effects. An alternative perspective is that the mechanism primed between categories is not accounted for by either the salience or the combination models of implicature derivation and may operate at an implicature-generic level.

7.3. Implications of the structure of the ‘mechanism’

As discussed, the nature of the ‘mechanism’ involved in the priming effects of Experiment 2 – 5 is unclear. That said, the nature of the priming effects does give us some insight into how the mechanism is structured.

Rees & Bott (2018) point out the general tendency of researchers to describe the derivation of implicature as a series of independent steps, e.g., Meyer and Feiman (2021), Tomlinson, Bailey and Bott, (2013), Breheny, Ferguson, and Katsos (2013) and Katsos and Bishop (2011), a.o. Undoubtedly, such perspectives are inspired by the step-based approach to implicature derivation posited by the Gricean-inspired perspective. As described above, I argue that such accounts conflate the individual mechanisms involved in theoretical implicature derivation with the individual mechanisms involved in the cognition of implicature to some extent – or at least do not give the boundaries between the two due consideration. Because activation-based priming targets cognitive mechanisms, the findings of Experiments 2 – 5 can provide insight into how the mechanisms governing implicature derivation are structured and whether these map onto theoretical mechanisms.

Generally, priming is a useful tool for investigating the domain-generality of different processes (Zeng et al., 2021). If the priming effects in Experiment 2 and Experiments 3 – 5 function via different mechanisms, then this suggests some modularity to the process of

implicature derivation (Fodor, 1975) as different mechanisms have been accessed independently of one another. As mentioned, this is the perspective assumed by many researchers who seek to investigate the derivation of implicature. Nevertheless, this assumption is not necessarily predicted by the Gricean-inspired account, whereby the theoretical steps that constitute implicature derivation are not considered to be representative of explicit cognitive steps. However, the results from Experiment 2 – 5 suggest that there are separate mechanisms involved in implicature derivation that can be independently primed.

While Bott and Chemla (2016), Meyer and Feiman (2021) and Marty et al. (2022) provide evidence to suggest that between-category priming effects are weaker than within-category priming effects, the priming effects exist within broad-scope implicature categories (i.e., within QI). Considering this, it is not possible to ascertain whether the within-category priming advantage is the result of the sameness of the primes and critical trials or the result of more activation across more levels of sub-computation. Experiments 2 – 5 provide evidence for two distinct priming effects that fall along the lines of broad-scope implicature categories (i.e., MI vs. non-MI), which indicates that there may be both implicature-specific and implicature-generic mechanisms responsible for implicature derivation.

The potential existence of implicature-specific and implicature-generic mechanisms provides evidence to support the perspectives of previous studies that have assumed this type of relationship (e.g., Bott & Chemla, 2016; Rees & Bott, 2018; Meyer & Feiman, 2021). Moreover, this suggests that the modular nature of the theoretical models of implicature derivation posited by Grice (1975), Levinson (2000) and Horn & Ward (2006) may reflect a cognitive reality. On the flip side, evidence for the coarticulation of independent cognitive mechanisms may suggest that unified accounts like that of Relevance Theory (Carston, 2006; Sperber & Wilson, 1995) may not reflect cognitive nuances associated with implicature derivation²⁹.

The existence of different types of shared and non-shared mechanisms may suggest some hierarchical ordering in the way implicature is derived. While hierarchical processing is a widely assumed component of various models of cognition and language (e.g., Friederici, 2012), there is little experimental evidence to suggest that the mechanisms responsible for implicature processing are ordered hierarchically. Evidence of highly reactive domain-

²⁹ Although Relevance Theory considers the mind to be modular (Sperber & Wilson, 2002) the cognitive process that results in pragmatic enrichment is not modular.

specific implicature mechanisms (i.e., a MI-specific search mechanism) and less reactive, domain-generic implicature mechanisms (i.e., the usage mechanism) may provide anecdotal evidence that implicature derivation is hierarchical, whereby more generic mechanisms may be shared across various implicature subtypes (e.g., Experiments 3 – 5), but not with other types of enrichment (e.g., Bott & Chemla, 2016), and more specific mechanisms may be shared between the same, or closely-related, implicature subtypes (Experiment 2).

The results of Experiments 3 – 5 do not suggest that the strength of the priming effect varies depending on the degrees of similarity between implicature subtypes – although a counter perspective is discussed in 7.4.3.1. Assuming a hierarchical model of implicature processing, this could be indicative of a type of informational encapsulation whereby there is limited lateral interaction between processes lower in a hierarchy³⁰, such that priming may not be able to occur between implicature subtypes at lower levels in the processing hierarchy, even if the mechanisms are similar to one another. In this sense, there may not be an individual mechanism responsible for the search for MI-specific alternatives; instead, priming may not occur between this level of the cognitive mechanism.

Although the current study does not provide explicit evidence in support of this hypothesis, evidence for a hierarchical ordering of the mechanisms associated with implicature derivation suggests that the Gricean-inspired perspective's focus on layered conversational heuristics may reflect the cognition of implicature.

7.3.1. Conclusions about the nature of mechanism

Although much of the above discussion is speculative, I do make some tentative conclusions regarding the nature of the mechanism primed in Experiments 2 – 5. Firstly, I conclude that two different mechanisms are primed in the experiments, one specific to MI and another more generic. This conclusion is based on the fact that the effects in Experiment 2 and Experiments 3 – 5 differ in their magnitude, the degree to which they are localized and the type of prime used. This conclusion entails a degree of modularity to the cognitive mechanism responsible for implicature derivation, such that different (sub-) mechanisms can be accessed independently of one another. I tentatively conclude that the mechanisms may be ordered hierarchically and that there may be limited interaction between lateral mechanisms, such that the degrees of shared computation between more specific types of implicature are hard to see.

³⁰ This idea is loosely inspired by Fodor's model of cognition (1983) and by frameworks of predictive coding, e.g., Friston, 2005.

Overall, this account is most compatible with a combination model of implicature derivation – however, an alternative explanation is discussed in the section below.

7.4. Ramifications theories of conversational implicature

I will provide an outline of the broader implications of the findings of Experiments 1 – 5 for existing theories of conversational implicature derivation.

7.4.1. The Gricean-inspired perspective

The Gricean-inspired account predicts MI and QI to be governed by the same mechanism, although there are considered to be fundamental differences between the phenomena. These are outlined below:

Table 10: Qualities of QI and MI.

<u>Qualities of Implicature</u>	QI	MI
Defeasible?	✓	✓
Non-Detachable?	✓	✗
Non-Conventional?	✓	✓
Context-dependent?	✓	✓
Generalised?	✓ *	✗
Lower-bounded?	✓	✗
Locally enriched?	✓ *	✗
Semantic equivalence of the alternative?	✗	✓

One major source of difference between QI and MI from a Gricean-inspired perspective is the nature of the alternative. For MI, the alternative is semantically equivalent to the articulated utterance, and the alternative is lower-bounded, i.e., it is a minimal interpretation. In contrast, for QI, the alternative is not semantically equivalent to the utterance and is upper-bounded, i.e., is stronger or more informative.

Consequently, the conclusion that QI can prime MI derivation but not via the exact same mechanisms as MI-to-MI priming occurs is compatible with the Gricean-inspired perspective. Considering the nature of the alternative in QI and MI is different, the idea that a different search mechanism may be responsible for each is not controversial and provides a clear explanation for the loci of the different priming effects in Experiment 2 and Experiments 3 – 5. However, the observation that QI-to-MI priming may occur by an overarching mechanism but that MI-to-MI priming may involve a domain-specific mechanism is best accommodated by a combination model of implicature derivation, which is not immediately compatible with the Gricean-inspired perspective which typically considers the activation of the usage mechanism to occur automatically (Rees & Bott, 2018), which is incompatible with the idea that it possesses its own level of activation and can be independently primed.

However, although Rees & Bott (2018) suggest that the usage mechanism may not be able to be primed independently, it should be noted that this conclusion is made with regards to within-QI priming. Here, the priming effect is understood to occur primarily via the alternative search mechanism based on the observation that alternative primes and usage primes confer the same priming effect. While this may be true for phenomena that share a level of sub-computation responsible for the alternative, if the phenomena do not share this level of sub-computation (e.g., MI and QI), then the between-category priming of this mechanism may not be possible, and a priming effect may only occur via the residual activation of the usage mechanism, even if the usage mechanism is activated automatically by a mechanism of sub-computation that is not shared by the phenomena.

Although the findings of Experiment 2 – 5 fit more comfortably with a combination model of implicature derivation, they may not be incompatible with the type of salience models posited by the Gricean-inspired account, which contrasts with Rees & Bott’s (2018) conclusions.

7.4.1.1. *Ad-hoc vs. Adjectival scalars*

The Gricean-inspired perspective is divided in the degree to which it considers ad-hoc QI and adjectival scalars to be processed in the same way, outlined below:

Table 11: Qualities Of Different Types Of Implicature As Per The Neo-Gricean Perspective.

Qualities:	<i>Default?</i>		<i>Calculated?</i>		<i>Particularized?</i>	
	Ad-hoc QI	Adjectival Scalar	Ad-hoc QI	Adjectival Scalar	Ad-hoc QI	Adjectival Scalar
Grice (1975)	x	x	✓	✓	✓	✓

Horn (1972, 1989)	✓	✓	✓	✓	✗	✓
Levinson (2000)	✓	✓	✗	✓	✗	✓

As discussed, the lack of a significant difference in the strength of the priming effect between Experiments 4 and 5 is most compatible with a Gricean account of QI derivation. However, in light of the conclusions made in light of Experiments 2 – 3 this perspective can be expanded upon.

If the between-category priming effect observed in Experiments 3 – 5 occurs via the priming of the usage mechanism, then the difference in the context-dependency of the alternative involved in ad-hoc QI and adjectival scalar derivation may be irrelevant. Consequently, providing the usage mechanism is sufficiently activated, both types of QI prime may prime subsequent MI equally. As discussed, non-calculated, default implicature generation is unlikely to involve the same usage mechanism as that of MI – or at least not activate it to the same degree as other implicature. Considering this, the findings of Experiments 2 – 5 may be more compatible with Grice's (1975) and Horn's (1972, 1989) perspectives than that of Levinson (2000). Nevertheless, it is unclear whether Levinson (2000) would consider the potentially context-dependent ad-hoc QIs elicited in Experiment 4 to involve no on-line processing given the ambiguity they entail. Indeed, the QIs used in Experiment 4 and 5 may be more context-dependent than initially thought.

Considering this, I cannot say with any certainty whether the results of Experiments 4 – 5 are better accounted for by Grice (1975), Horn (1972, 1989) or Levinson (2000). That said, Grice's (1975) original account provides the simplest explanation for the lack of difference between the two.

7.4.1.2. *Manner Implicature*

Overarchingly, the findings of Experiments 1 – 5 suggest that MI does comprise a unique implicature subtype, as accounted for by the Gricean-inspired perspective. Experiments 1a and 2 suggest that there is a psychological reality to the marked-to-marked pairing that MI elicits. Moreover, the difference between the priming effects in Experiment 2 and Experiments 3 – 5 (e.g., localisation and magnitude) suggest that the heterogeneous inferences elicited by the 42 critical trials used in Experiments 1a – 6 do comprise a specific type of implicature that is governed by mechanisms that are not shared with other types of implicature, i.e., QI and I-Implicature. Given that the Gricean-inspired perspective considers the nature of the alternative to differ in QI and MI derivation but for the same overarching

implicature mechanism to derive both phenomena, the results of the current study are compatible with the Gricean-inspired perspective that considers MI a type of EVA that shares some, but not all, of its sub-computations with QI.

The findings of Experiments 2 – 5 do suggest that there is a more apparent distinction between MI and QI than that posited by Horn (1984). While Grice (1975) and Levinson (2000) delineate form-based and information-based implicature more explicitly, Horn's account allows for MI to be derived via the Q-Principle in some cases, like QI. At the same time, the evidence from the current study suggests that there is a distinction between the mechanisms responsible for governing each; only one type of MI is investigated. Considering this, the findings are not incompatible with Horn's perspective, as it is impossible to ascertain whether the primed mechanisms pertain to all MI.

As outlined, the type of MI used in the critical trials of Experiments 1a – 6 is highly dependent on context and is not associated with a particular marked form as other types of MI may be, e.g., *periphrastic causatives*. Given that the Gricean-inspired perspective considers implicature to be derived by a set of domain-specific mechanisms, the between-category priming of prolixity-based MI and QI and I-Implicature may be more unexpected than between-category priming involving more generalised MI. Considering the priming effect in Experiment 1a – 6 involves the most 'extreme' case of MI, i.e., the subtype that is least similar to QI and I-Implicature, the existence of a priming effect provides convincing evidence to suggest that MI shares an implicature-level derivational mechanism with QI and I-Implicature, beyond only sharing lower levels of computation, e.g., say, a mechanism responsible for generating alternative syntactical representations in the case of periphrastic vs. non-periphrastic causatives.

In turn, this suggests that MI is a phenomenon at the level of 'implicature' and comprises a distinct category thereof. Further investigation would benefit from using a variety of different MI with differing degrees of similarity to QI and I-Implicature to ascertain whether there are different mechanisms involved in the derivation of more generalised MI that may be involved in that of QI and I-Implicature, which the current study does not do. This may provide evidence for Rett's (2020) perspective that MI derived via the Maxim of Manner's third submaxim, '*Be brief or orderly*', e.g., litotes, evaluatives and periphrasis, a.o., are more similar to QI than the Gricean-inspired perspective considers them to be.

7.4.1.3. *I-Implicature*

As outlined in 2.3.5, there is no unified account of I-Implicature in the Gricean-inspired perspective. Although Grice (1975) and Horn (1984) do not give much attention to the nature of I-Implicature, Levinson's account involves the extensive documentation of phenomena reported to be instances of I-Implicature (Levinson, 2000). The existence of an equal priming effect between I-Implicature and QI on MI derivation is not predicted by either Horn or Grice's accounts, given that it is not considered a type of conversational implicature – explicitly, at least.

Moreover, whether or not Levinson's account predicts priming is unclear. On the one hand, Levinson considers I-Implicature to be derived automatically and not to involve the same level of calculation as other types of implicature – this does not predict a priming effect between I-Implicature and MI. On the other hand, Levinson explicitly rejects the interpretation that I-Implicature is a type of implicature that is derived via a specific mechanism and concedes that the cognition of different types of I-Implicature can occur using different algorithms, which makes it more difficult to make predictions concerning the mechanisms it may share with MI and other implicature.

Given that the current study assumes an activation-based account of priming and assumes the domain-specificity of 'mechanisms' involved in implicature derivation, then the finding that I-Implicature can prime MI is compatible with Levinson's (2000) account of I-Implicature, providing Levinson allows for I-Implicature to share certain levels of sub-computation with other types of inference (i.e., MI) and allows for the non-default computation of some types of I-Implicature. While it is unclear whether the account posited by Levinson (2000) allows for this, this possibility is not explicitly rejected by Levinson and may be an explanation for the results in Experiment 3.

However, assuming this perspective has ramifications for the way that I-Implicature – or at least *implicature-to-stereotype* – is conceived. Given that the 'core account' (Meyer & Feiman, 2021) assumes the usage mechanism to be EVA-specific (i.e., it negates the alternative and combines it with the utterance's surface interpretation), then I-Implicature must be a type of EVA if it primes MI via this mechanism. This contrasts with the perspective of the Gricean-inspired account, which does not account for I-Implicature derivation to occur in the same way as MI and QI derivation.

7.4.2. Other accounts of implicature derivation

7.4.2.1. Relevance Theory

Broadly, evidence that MI derivation can be primed by MI, QI, and I-Implicature is compatible with Relevance Theory's perspective on implicature derivation (Carston, 2005; Sperber & Wilson, 1995). Relevance Theory posits a unified cognitive mechanism that is broadly responsible for all types of pragmatic enrichment. Consequently, MI, QI, and I-Implicature are all governed by the same cognitive mechanism, which appears to predict the between-category priming effect observed in Experiments 3 – 5 and the within-category priming effect observed in Experiment 2.

Nevertheless, Relevance Theory's inferential mechanism is not considered to comprise of domain-specific sub-computations. Consequently, while Relevance Theory may predict the priming effect observed in Experiments 2 – 5, it does not predict there to be a difference between the within-category and between-category priming effects observed when comparing Experiment 2 to Experiments 3 – 5. Moreover, as discussed in light of Experiment 3, it is unclear whether a unified, generic cognitive mechanism could be primed the type of dynamic way observed across the experiments. Given that Relevance Theory's inferential mechanism experiences perpetual activation during conversation, it appears unintuitive to suggest that the derivation of implicature could substantially increase its activation.

Consequently, while the results of Experiments 2 – 6 are broadly compatible with Relevance Theory, there are points of contention when specific patterns of priming are considered.

7.4.2.2. The Grammatical Account

The Grammatical Account (e.g., Fox, 2007; Chierchia, 2006) is largely concerned with scalar implicature, so cannot be used to make predictions concerning within-category MI priming. That said, the findings of Experiment 2 are compatible with the Grammatical Account for this very reason. However, accounts of I-Implicature and QI are of more concern to the Grammatical Account and scalar implicature and I-Implicature-like phenomena have been accounted for as grammaticalized, or partially-grammaticalized (Chierchia, 2004; Chierchia, Fox & Spector, 2012). Nevertheless, MI is typically considered to be an entirely pragmatic phenomenon, which means that the between-category priming effect observed in Experiment 3 – 5 is not predicted by the Grammatical Account. While the Grammatical Account does consider pragmatics to play a role in cancelling syntactically-based scalar implicature, which

may explain the results of Experiments 3 – 5, it is presently unclear why priming should occur via such a mechanism.

7.4.2.3. *Probabilistic Accounts*

The findings of Experiments 2 – 5 are compatible with probabilistic accounts for implicature derivation too, e.g., game-theory approaches (e.g., Lewis, 196; Parikh, 2001; Franke, 2009; Blutner 2000; Bergen, et al., 2016, a.o.) Bayesian approaches (e.g., Lassiter and Goodman, 2014; Potts et al., 2015) and constraint-based approaches to pragmatic reasoning (e.g., Frank & Goodman, 2012; Goodman & Stuhlmüller, 2013; Degen & Tanenhaus, 2014, 2019).

Considering the likelihood of pragmatic enrichment is modelled probabilistically and is sensitive to context factors, e.g., world knowledge, QUDs, etc., it follows that the perceived probability of implicature elicitation should increase in instances in which prior implicature derivation occurs and should be stronger within-category than between-category.

Nevertheless, the findings of Experiment 2c suggest that conversational context is not taken into account, either consciously or subconsciously, when answering the critical trials, given that training the participants to spot MI does not increase the rate of MI derivation in the critical trials. Considering that explicit instruction should increase the perceived probability of subsequent implicature derivation, it is unclear whether the results of Experiment 2c are compatible with the predictions of probabilistic accounts of implicature derivation.

7.4.3. *Adjusting the Gricean-inspired perspective?*

While the findings of Experiment 2 – 5 are broadly compatible with the Gricean-inspired perspective, I propose some tentative clarifications to the way that MI and I-Implicature may be conceived that streamline this account.

Based on the findings of the current study and the intuition that certain types of I-Implicature are computed in a fashion similar to that of MI and QI, I propose that I-Implicature may be derived using an alternative. If I-Implicature is an EVA then the derivation of the types of I-Implicature elicited in the trials used in Experiments 1b and 3 may look as follows:

Table 12: *Step-by-Step derivation of I-Implicature.*

Steps	Utterance
1) Speaker said ‘x’	John said ‘Hello’ to the secretary and then he smiled

2) Speaker chose to say 'x' instead of a more <i>informative</i> alternative (y)	e.g., John said 'Hello' to the <i>secretary, who happened to be a robot</i> , and then he smiled
3) If Speaker meant 'y', they would have said it.	e.g., John said 'Hello' to the <i>secretary, who happened to be a robot</i> , and then he smiled
4) Assuming the Speaker is rational, the fact that they did not say the alternative means - '¬ y'	
5) Speaker means 'x and ¬ y'.	[John said 'Hello' to the secretary and then he smiled] and ¬ [anything more]

Although I have used QI as a possible avenue for I-Implicature derivation, I take no firm stance on which *Gricean* maxim is responsible for I-Implicature. While the Maxim of Quantity may be responsible for the generation of some types of I-Implicature, Relevance may play a role in others. At first glance, it is unclear whether I-Implicature implicates that there is not a more *informative* alternative or not a more *relevant* alternative. Nevertheless, the process delivers results in line with Levinson's (2000) I-Heuristic's Speaker's Maxim 'Do not say more than is required' and either the Q-Heuristic's Recipients Corollary 'What isn't said isn't' or the I-Heuristics' 'What is simply described is stereotypically exemplified', or both.

Interestingly, an EVA account of I-Implicature may explain why some I-Implicature can be laid bare by the inclusion of the word 'only' – as other types of EVAs are (Levinson, 2000):

A: 'Are they dating?'

B: 'They only went for a coffee!'

I +> They did not go for a coffee in more relevant or more informative way than suggested.

I ++> [...], i.e., the coffee was not romantic'

Example 7-1

Accounting for I-Implicature in this way may resolve I-Implicature's apparent incompatibility with the qualities of context-dependency, non-conventionality and defeasibility – issues

discussed in 2.3.5.1. Moreover, there is no need for the processing of a potentially context-dependent stereotypical concept, e.g., that of ‘going for a coffee’, which presumably reduces the cognitive cost associated with I-Implicatures, which are undoubtedly frequent in conversation and therefore would benefit from parsimonious derivation. Here, the secondary inference does most of the ‘leg work’ and is likely relevance-based and contextually-dependent and variable in its specificity, e.g., it may enrich the utterance to a large or small degree depending on conversational factors or may not be derived at all.

However, the theoretical reconstruction of I-Implicature is not an aim of this thesis and discussion concerning the nature of I-Implicature will be left open ended. Nevertheless, the between-category priming effect observed in Experiment 3 suggests that I-Implicature and QI and MI have more in common than predicted by the Gricean-inspired perspective.

7.4.3.1. *Primary and Secondary Inferences*

As discussed 2.3 and identified by Levinson (2000) conversational implicature subtypes appear to have a basic structure (e.g., for MI ‘+> in a marked way’) alongside various ad-hoc, context-dependent enrichments (e.g., ‘dance’ may I-Implicate ‘+> the waltz’). The Gricean-inspired programme conceives of MI as distinct from other types of conversational implicature given its purportedly ad-hoc nature, yet QI and certain types of MI may be more similar than the Gricean-inspired account acknowledges, as identified by Rett (2020). Although MI are undoubtedly more context-dependent, considering they are elicited by marked forms, the similarities between MI, QI and I-Implicature – and indeed, other types of implicature – may be more easily observed if ‘primary’ and ‘secondary’ inferences are considered separately. For instance, the implicatures ‘+> not in a more informative way’ (QI), ‘+> not in a more informative/relevant way’ (I-Implicature) and ‘+> not in a more stereotypical way’ (MI) appear more uniform than when ad-hoc instances of each are considered.

Although I do not make any claims about the nature of primary and secondary implicature, nor the mechanisms by which they are derived, it appears important that the distinction between the two is considered when the nature of implicature is investigated, both theoretically and experimentally. If not, then boundaries drawn between different phenomena (e.g., I-Implicature and MI) may inadvertently reflect the difference between primary and secondary inferences, rather than the typological differences they are intended to. Assuming

this perspective, the results of Experiments 2 – 5 may not be as surprising as initially considered. Indeed, if EVAs are generated by an EVA-specific mechanism and if this mechanism generates primary rather than secondary implicatures, then a cross-priming effect should be observed between phenomena, regardless of their degree of context-specificity.

As discussed in 2.3.3.4, there is uncertainty concerning where context-dependent information first enters the mechanism of implicature derivation. In general, the Gricean-inspired perspective requires context-dependent alternatives to be generated before the entire implicature is generated – for instance, to derive MI, I-Implicature must first be generated and negated. In the case of MI derivation, deriving the ‘right’ MI is reliant on the selection of the appropriate I-Implicature and anticipation that its negation selects an appropriately marked referent. Presumably, such a process is cognitively burdensome. Instead, a more streamlined mechanism may first derive primary inferences which are then enriched via another mechanism – this idea is alluded to by Levinson (2000).

Assuming this perspective opens up another interesting interpretation of the priming effect in Experiments 2 – 6. Here, the different priming effects may not act independently on an alternative-specific search mechanism and a usage mechanism – as described above – but rather on a generic EVA usage mechanism (i.e., the between category effect) and on whatever mechanism may be responsible for the derivation of ad-hoc secondary inferences (i.e., the within-category priming effect) – after all, clothing-specific MI was almost uniform between the MI trials. Although addressing this question falls outside the scope of the present thesis, it is an interesting avenue for theoretical and experimental investigation.

7.5. Research Outlook

7.5.1. Limitations

Naturally, Experiments 1 – 6 are not without their limitations. One overarching limitation of the studies is the non-randomness of the assignment of the participants to either the baseline or experimental conditions. The experiments were conducted in chronological order such that baseline (Experiment 1) was established prior to the running of Experiments 2 – 6.

Establishing the experimental baseline was an iterative task and multiple baseline experiments were conducted, though they have not been reported in this thesis. The baseline studies varied slightly (e.g., the number and ‘type’ of fillers, the size and positioning of the images, etc.) and were conducted with different pools of participants and were subject to asynchronous re-runs. The baselines showed a remarkably stable rate of MI of approximately 15%. Consequently, given the stable baseline rate of MI derivation in the paradigm and

restraints on the number of participants I was able to recruit in Experiments 2 – 6, Experiment 1a was conducted independently to Experiments 2 – 6. As such, participants recruited to in Experiments 2 – 6 could not be assigned to the baseline condition. Although this runs counter to usual experimental recommendations, the baseline condition's stability meant that I was confident that the baseline rate of MI derivation does not vary across sampling periods. That said, future research may benefit from rectifying this.

An open question is whether the binary judgement measure of MI derivation partially obscures the observation of a larger – or different - priming effect in Experiments 2 – 5. The decision to use a measure of binary judgement was made in line with Bott & Chemla's (2016) study, with an effort to ensure the MI derivation was as unambiguous as possible – as discussed, participants are unlikely to select the marked image unless deriving MI. That said, other measures like reaction time or mouse- and eye-tracking metrics may be useful, too. For instance, reaction time metrics could be used to assess whether the correctly answered MI critical trials take longer to complete than incorrectly answered MI trials, I-Implicature trials or the guided prime trials, which is predicted by the Gricean-inspired account's perspective that MIs require additional processing. As covered extensively, the numerical increase in the rate of MI between the prime trials and the critical trials is comparatively small. Given this, other measures may be more sensitive in ascertaining the effect of the primed trials and may provide evidence of a larger priming effect than the measures used in the current study.

Finally, the use of the unscreened data (i.e., the 'primed' and 'unprimed') as the basis of the comparison between experiments may limit the sensitivity of the data analysis to the priming effect. The use of the unscreened data was motivated by the fact that screening the data of Experiment 2a would lead to a loss of statistical power, i.e., the dataset would be too small. Although this was not the case in Experiments 2b – 6, using the screened data from these experiments and comparing them alongside Experiment 2c would be problematic, given the latter contains majority 'incorrect' responses. While Experiment 2b and 2c could have been used as the basis of a comparison between the within- and between-category priming effects, they did not represent a 'neutral' pseudo-baseline as they included elements (i.e., the guided primes and the pre-trial acclimatisation task) that were not present in Experiments 3 – 6. Moreover, the use of screened data, e.g., as in Bott & Chemla's study (2016), presupposes a localised priming effect, contingent on the correct answering of two primes trials immediately preceding the relevant critical trial, specifically. Naturally, this is an arbitrary definition of priming and runs counter to the observation of non-localised priming in the study. Consequently, when comparing the between- and within-category priming effects, the use of the unscreened data from Experiment 2c with the unscreened data from Experiments 3 – 6 appeared to be the least contentious option.

However, future research would benefit from harmonising the paradigm(s) used in Experiments 2 – 6 such that a more straight-forward comparison between the within- and between-category priming effects can be made, either by including the same manipulations across all experiments (e.g., guided primes, pre-trials) or by adapting the paradigm to ensure higher rates of MI-to-MI priming, such that there exists a viable screened dataset.

7.5.2. Contributions

The thesis contributes to research in experimental and theoretical pragmatics in multiple ways but also raises questions for future research. The critical contributions of the thesis and its research outlook are outlined below.

Manner Implicature

The thesis presents a novel methodology for eliciting MI in experimental settings and priming MI with other types of conversational inference. Unlike previous studies that have investigated MI experimentally, e.g., Wilson (2017) and Katsos and Wilson (2016), the current study does so over a wide range of different ad-hoc QIs. The range of the rate of MI derivation in Experiment 1a provides empirical evidence that MIs are ad-hoc, context-dependent phenomena, as predicted by theories of conversational implicature.

However, while Experiment 1a captures a heterogeneity not observed in previous studies, the extent to which MI is elicited in the 42 items used in Experiments 1a – 6 is unclear.

Moreover, the factors that modulate the availability of MI are uncertain, although the perceived markedness of the referents may play a role. Future experimental research into the nature of MI may benefit from investigating the role the objective proximity (e.g., word length) and the subjective markedness of an utterance has in eliciting MI as these factors were not measured by either of the item screening tasks.

Because of the construction of the current study's critical trials, the role of the alternative in MI derivation and MI priming cannot be properly investigated. To better understand the role of the alternative specifically, the paradigm could be adapted to modulate the extent to which the alternative is represented in both the critical trials (both linguistically and visually) and prime trials, allowing the paradigm to contribute to discussion concerning the role of the alternative in implicature derivation and implicature priming (e.g., Rees & Bott, 2018; Waldon & Degen, 2020).

Finally, Experiment 1a and Experiment 2 deal with just one type of MI – specifically, *proximity-based MI*. Consequently, the MIs present in the experiment (i.e., the 42 experimental items) are ad-hoc, particularised and therefore non-uniform, to a degree. While I expect the findings of Experiment 2 to hold true for other types of MI, it appears intuitive that the observed priming effects would be more robust in more uniform types of MI, for instance, *periphrastic causatives*. Future research could provide insight into whether priming occurs

within different types of MI (e.g., periphrastic-causative-to-periphrastic-causative) and between different types of MI (e.g., prolixity-based-to-periphrastic-causative, or vice-versa).

I-Implicature

Although not a primary concern of this thesis, the theoretical and experimental exploration of I-Implicature makes important contributions to the understanding of I-Implicature – which, like MI, is an underexplored type of inference. The findings of Experiment 1b provide novel evidence – as far as I know – of the psychological reality of *inferences to stereotype* across a large caddy of stimuli. Moreover, Experiment 3 provides evidence suggesting that I-Implicature can be used to prime different types of conversational implicature and may be computed similarly.

These findings have important theoretical ramifications, as discussed. Consequently, I recommend that future theoretical and experimental investigations consider certain types of I-Implicature to be *conversational implicatures* to the same extent as QI and MI. However, the extent to which all types of I-Implicature can be accounted for as EVAs remains to be determined. Future research should ascertain the processing of different types of I-Implicature and seek to establish the boundaries of the mechanism(s) that derive them, addressing the ambiguity Levinson (2000) leaves regarding their processing.

Mechanisms of Implicature Derivation

A large proportion of this thesis is dedicated to discussion concerning the nature of the mechanism that governs implicature derivation. The thesis contributes to a growing body of research that uses priming methods to investigate the extent to which mechanisms are shared, partially shared, or separate in the derivation of different types of implicature, e.g., Bott & Chemla (2016), Meyer and Feiman (2021), Mary et al. (2024). The findings in the thesis suggest that the paradigm accesses different mechanisms and that they are not shared between MI and other implicature subtypes.

However, it is unclear whether the distinct mechanisms represent a difference between between-category and within-category priming effects (i.e., MI vs other implicature), a difference between domain-specific priming effects (e.g., context-dependent vs. lexicalised, or partially-lexicalised implicature) or a difference based on the ‘trigger’ of the enrichment (i.e., form vs. information). Consequently, future research may investigate whether there are similar priming effects between ‘lexicalised’ MI – e.g., periphrastic causatives and ad-hoc

MIs, or whether less context-dependent MI are more readily primed by ad-hoc QI, adjectival scalars and I-Implicature than prolixity-based MI is. In turn, such experimental investigation contributes to the discussion concerning the degree to which different types of implicature share levels of sub-computation.

8. References

- Andrews, E. (2012). *Markedness*. In R. I. Binnick (Ed.), *The Oxford Handbook of Tense and Aspect*. Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780195381979.013.0007>
- Antoniou K, Grohmann KK, Kambanaros M, Katsos N. The effect of childhood bilectalism and multilingualism on executive control. *Cognition*. 2016 Apr;149:18-30. doi: 10.1016/j.cognition.2015.12.002. Epub 2016 Jan 8. PMID: 26774217.
- Antoniou, K., & Katsos, N. (2017). The effect of childhood multilingualism and bilectalism on implicature understanding. *Applied Psycholinguistics*, 38(4), 787–833. <https://doi.org/10.1017/S014271641600045X>.
- Antoniou, K., & Katsos, N. (2017).** The effect of childhood multilingualism and bilectalism on implicature understanding. *Applied Psycholinguistics*, 38(4), 787–833. <https://doi.org/10.1017/S014271641600045X>
- Antoniou, K., Cummins, C., & Katsos, N. (2016). Why only some adults reject under-informative utterances. *Journal of Pragmatics*, 99, 78–95. <https://doi.org/10.1016/j.pragma.2016.05.001>
- Atlas, J. D., & Levinson, S. C. (1981). It-clefts, informativeness, and logical form: Radical pragmatics (revised standard version). In P. Cole (Ed.), *Radical Pragmatics* (pp. 1–61). Academic Press.
- Baker, R., Doran, R., McNabb, Y., Larson, M., & Ward, G. (2009). On the non-unified nature of scalar implicature: An empirical investigation. *International Review of Pragmatics*, 1(2), 211–248. <https://doi.org/10.1163/187730909x12538045489854>
- Barbet, C., & Thierry, G. (2018). When *some* triggers a scalar inference out of the blue: An electrophysiological study of a Stroop-like conflict elicited by single words. *Cognition*, 177, 58–68. <https://doi.org/10.1016/j.cognition.2018.03.013>
- Bargh, J. A., & Chartrand, T. L. (2000). Studying the mind in the middle: A practical guide to priming and automaticity research. In H. Reis & C. Judd (Eds.), *Handbook of research methods in social psychology* (pp. 1–39). Cambridge University Press.
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3), 10.1016/j.jml.2012.11.001. <https://doi.org/10.1016/j.jml.2012.11.001>
- Barr, D., & Keysar, B. (2002). Anchoring and comprehension in linguistic precedents. *Journal of Memory and Language*, 46, 391–418.
- Bassi, I., Del Pinal, G., & Sauerland, U. (2021). Presuppositional exhaustification. *Semantics and Pragmatics*, 14(11), 1–48. <https://doi.org/10.3765/sp.14.11>

- Baxter, S. M., Ilicic, J., Kulczynski, A., & Lowrey, T. M. (2017). Using sublexical priming to enhance brand name phonetic symbolism effects in young children. *Marketing Letters*, 28(4), 565–577. <https://doi.org/10.1007/s11002-017-9430-9>
- Benz, A., & Jasinskaja, K. (2017). Questions under discussion: From sentence to discourse. *Discourse Processes*, 54(3), 177–186. <https://doi.org/10.1080/0163853X.2017.1316038>
- Bergen, L., Levy, R., & Goodman, N. (2016). Pragmatic reasoning through semantic inference. *Semantics and Pragmatics*, 9, 1–83. <https://doi.org/10.3765/sp.9.20>
- Blass, R. (1990). *Relevance relations in discourse*. Cambridge University Press.
- Blutner, R. (2000). Some aspects of optimality in natural language interpretation. *Journal of Semantics*, 17(3), 189–216. <https://doi.org/10.1093/jos/17.3.189>
- Bock, J. K. (1986). Syntactic persistence in language production. *Cognitive Psychology*, 18(3), 355–387. [https://doi.org/10.1016/0010-0285\(86\)90004-6](https://doi.org/10.1016/0010-0285(86)90004-6)
- Bock, K., & Griffin, Z. M. (2000). The persistence of structural priming: Transient activation or implicit learning? *Journal of Experimental Psychology: General*, 129(2), 177–192. <https://doi.org/10.1037/0096-3445.129.2.177>
- Bond, N. (2018). *Aretha Franklin's long history of legendary feuds*. *New Zealand Herald*. Retrieved from <https://www.nzherald.co.nz/entertainment/aretha-franklins-long-history-of-legendary-feuds/JCRZHASYBN4RBX3TZ3NQOI2UMM/>
- Bott, L., & Chemla, E. (2013). Pragmatic priming and the search for alternatives. *Cognitive Science*, 35(35). <https://escholarship.org/content/qt9wr510rh/qt9wr510rh.pdf?t=os5zqh>
- Bott, L., & Chemla, E. (2016). Shared and distinct mechanisms in deriving linguistic enrichment. *Journal of Memory and Language*, 91, 117–140. <https://doi.org/10.1016/j.jml.2016.04.004>
- Bott, L., & Frisson, S. (2022). Salient alternatives facilitate implicatures. *PLOS ONE*, 17(3), e0265781. <https://doi.org/10.1371/journal.pone.0265781>
- Bott, L., & Noveck, I. A. (2004). Some utterances are underinformative: The onset and time course of scalar inferences. *Journal of Memory and Language*, 51(3), 437–457. <https://doi.org/10.1016/j.jml.2004.05.006>
- Boudewyn, M. A., Zirnstein, M., Swaab, T. Y., & Traxler, M. J. (2014). Priming prepositional phrase attachment: Evidence from eye-tracking and event-related potentials. *Quarterly Journal of Experimental Psychology*, 67(3), 424–454. <https://doi.org/10.1080/17470218.2013.815237>
- Branigan, H. P., & Pickering, M. J. (2017). An experimental approach to linguistic representation. *Behavioral and Brain Sciences*, 40, e282. <https://doi.org/10.1017/S0140525X16002028>

- Breheny, R. (2008). A new look at the semantics and pragmatics of numerically quantified noun phrases. *Journal of Semantics*, 25(2), 93–139. <https://doi.org/10.1093/jos/ffm016>
- Breheny, R. (2019). “Language processing, relevance and questions,” in *Relevance, Pragmatics and Interpretation*, eds B. Clark, K. Scott, and R. Carston (Cambridge: Cambridge University Press), 42–52.
- Breheny, R., Ferguson, H. J., & Katsos, N. (2013). Taking the epistemic step: Toward a model of on-line access to conversational implicatures. *Cognition*, 126(3), 423–440. <https://doi.org/10.1016/j.cognition.2012.11.012>
- Breheny, R., Katsos, N., & Williams, D. (2006). Are generalised scalar implicatures generated by default? An on-line investigation into the role of context in generating pragmatic inferences. *Cognition*, 100(3), 434–463. <https://doi.org/10.1016/j.cognition.2005.07.003>
- Breheny, R., Katsos, N., & Williams, J. (2006). Are generalized scalar implicatures generated by default? An online investigation into the role of context in generating pragmatic inferences. *Cognition*, 100(3), 434–463. <https://doi.org/10.1016/j.cognition.2005.07.003>
- Brown, P., & Levinson, S. C. (1987). *Politeness: Some universals in language usage*. Cambridge University Press.
- Brown-Schmidt, S. (2009). Partner-specific interpretation of maintained referential precedents during interactive dialog. *Journal of Memory and Language*, 61(2), 171–190. <https://doi.org/10.1016/j.jml.2009.04.003>
- Bumford, D., & Rett, J. (2021). Rationalizing manner-driven evaluativity inferences. In *Proceedings of Sinn und Bedeutung* (Vol. 25)
- Cai, Z. G., & Zhao, N. (2024). Structural priming: An experimental paradigm for mapping linguistic representations. *Language and Linguistics Compass*, 18(2). <https://doi.org/10.1111/lnc3.12507>
- Carston, R. (1998). Informativeness, relevance and scalar implicature. *Relevance Theory*, 179–200. <https://doi.org/10.1075/pbns.37.11car>
- Carston, R. (2002). *Thoughts and Utterances: The Pragmatics of Explicit Communication*. Oxford: Blackwell.
- Carston, R. (2005). A note on pragmatic principles of least effort. *UCL Working Papers in Linguistics*, 271–278.
- Carston, R. (2006). Relevance theory and the saying/implicating distinction. In L. R. Horn & G. Ward (Eds.), *The handbook of pragmatics* (pp. 633–656). Blackwell Publishing. <https://doi.org/10.1002/9780470756959.ch30>
- Chahboun, S., Vulchanov, V., Saldaña, D., Eshuis, H., & Vulchanova, M. (2017). Can you tell it by the prime? A study of metaphorical priming in high-functioning autism in comparison

with matched controls. *International Journal of Language and Communication Disorders*, 52(6), 766–785. <https://doi.org/10.1111/1460-6984.12314>

Chang, F., Janciauskas, M., & Fitz, H. (2012). Language adaptation and learning: Getting explicit about implicit learning. *Language and Linguistics Compass*, 6(5), 259–278. <https://doi.org/10.1002/lnc3.337>

Chemla, E., & Singh, R. (2014). Remarks on the experimental turn in the study of scalar implicature, Part II. *Language and Linguistics Compass*, 8, 387–399. <https://doi.org/10.1111/lnc3.12080>

Chemla, E., & Spector, B. (2011). Experimental evidence for embedded scalar implicatures. *Journal of Semantics*, 28(3), 359–400. <https://doi.org/10.1093/jos/ffq023>

Chen, X., & Hartsuiker, R. J. (2021). Do particle verbs share a representation with their root verbs? Evidence from structural priming. *Journal of Memory and Language*, 120, 104255. <https://doi.org/10.1016/j.jml.2021.104255>

Chen, X., Wang, S., & Hartsuiker, R. J. (2023). Do structure predictions persevere to multilinguals' other languages? Evidence from cross-linguistic structural priming in comprehension. *Bilingualism*, 26(4), 653–669. <https://doi.org/10.1017/s1366728922000888>

Chevallier, C., Wilson, D., Happé, F., & Noveck, I. (2010). Scalar inferences in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 40(9), 1104–1117. <https://doi.org/10.1007/s10803-010-0960-8>

Chia, K., & Kaschak, M. P. (2022). It's not you, it's me: Some speakers elicit structural priming more reliably than others. *Collabra: Psychology*, 8(1). <https://doi.org/10.1525/collabra.36312>

Chierchia, G. (2004). Scalar implicatures, polarity phenomena, and the syntax/pragmatics interface. In A. Belletti (Ed.), *Structures and beyond: The cartography of syntactic structures* (Vol. 3, pp. 39–103). Oxford University Press.

Chierchia, G. (2006). Broaden Your Views: Implicatures of Domain Widening and the “Logicality” of Language. *Linguistic Inquiry*, 37(4), 535–590. <https://doi.org/10.1162/ling.2006.37.4.535>

Chierchia, G., Fox, D., & Spector, B. (2012). The grammatical view of scalar implicatures and the relationship between semantics and pragmatics. In C. Maienborn, K. von Stechow, & P. Portner (Eds.), *Semantics: An international handbook of natural language meaning* (Vol. 3, pp. 2297–2331). Mouton de Gruyter.

Clark, H. H. (1996). *Using language*. <https://doi.org/10.1017/cbo9780511620539>

Coderre, E. L., Chernenok, M., Gordon, B., & Ledoux, K. (2017). Linguistic and non-linguistic semantic processing in individuals with autism spectrum disorders: An ERP study. *Journal of Autism and Developmental Disorders*, 47(3), 795–812. <https://doi.org/10.1007/s10803-016-2985-0>

- Corley, M., & Scheepers, C. (2002). Syntactic priming in English sentence production: Categorical and latency evidence from an Internet-based study. *Psychonomic Bulletin & Review*, 9(1), 126–131. <https://doi.org/10.3758/bf03196267>
- Crepaldi, D., Morone, E. A., Arduino, L. S., & Luzzatti, C. (2014). Morphological processing of printed nouns and verbs: Cross-class priming effects. *Journal of Cognitive Psychology*, 26(4), 433–460. <https://doi.org/10.1080/20445911.2014.895007>
- Cummins, C., & Rohde, H. (2015). Evoking context with contrastive stress: Effects on pragmatic enrichment. *Frontiers in Psychology*. <https://doi.org/10.3389/fpsyg.2015.01779>
- de Carvalho, A., Reboul, A. C., Van der Henst, J. B., Cheylus, A., & Nazir, T. (2016). Scalar implicatures: The psychological reality of scales. *Frontiers in Psychology*, 7.
- Degen, J. (2023). The rational speech act framework. *Annual Review of Linguistics*, 9(1), 519–540. <https://doi.org/10.1146/annurev-linguistics-031220-010811>
- Degen, J., & Goodman, N. D. (2014). Lost your marbles? The puzzle of dependent measures in experimental pragmatics. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 36, 397–402.
- Degen, J., & Tanenhaus, M. K. (2011). Making inferences: The case of scalar implicature processing. In *Proceedings of the 33rd Annual Meeting of the Cognitive Science Society* (pp. 3299–3304).
- Degen, J., & Tanenhaus, M. K. (2015). Processing scalar implicature: A constraint-based approach. *Cognitive Science*, 667–710. <https://doi.org/10.1111/cogs.12171> (2015)
- Degen, J., & Tanenhaus, M. K. (2019). Constraint-based pragmatic processing. In C. Cummins & N. Katsos (Eds.), *The Oxford Handbook of Experimental Semantics and Pragmatics* (pp. 21–38).
- Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in Psychology*, 5, 781. <https://doi.org/10.3389/fpsyg.2014.00781>
- Diesendruck, G., & Markson, L. (2011). Children’s assumption of the conventionality of culture. *Child Development Perspectives*, 5(3), 189–195. <https://doi.org/10.1111/j.1750-8606.2010.00156.x>
- Dufour, S., & Frauenfelder, U. H. (2016). Inhibitory phonetic priming: Where does the effect come from? *Quarterly Journal of Experimental Psychology*, 69(1), 180–196. <https://doi.org/10.1080/17470218.2015.1045911>
- Dufour, S., & Meynadier, Y. (2019). Temporary ambiguity in whispered word recognition: A semantic priming study. *Journal of Cognitive Psychology*, 31(2), 157–174. <https://doi.org/10.1080/20445911.2019.1573243>
- Dyson, L., Morgan, J., & Herbert, R. (2021). Semantic processing in aphasia: Evidence from semantic priming and semantic interference. *Language, Cognition and Neuroscience*, 36(4), 491–516. <https://doi.org/10.1080/23273798.2020.1849749>

- Fairchild, S., & Papafragou, A. (2021). The role of executive function and theory of mind in pragmatic computations. *Cognitive Science*, *45*, e12938. <https://doi.org/10.1111/cogs.12938>
- Ferré, P., Guasch, M., García-Chico, T., & Sánchez-Casas, R. (2015). Are there qualitative differences in the representation of abstract and concrete words? Within-language and cross-language evidence from the semantic priming paradigm. *Quarterly Journal of Experimental Psychology*, *68*(12), 2402–2418. <https://doi.org/10.1080/17470218.2015.1016980>
- Ferreira, V. S., Bock, J. K., Wilson, M., & Cohen, N. J. (2005, November). Structural persistence in anterograde amnesia: Evidence for implicit learning. Paper presented at the 46th Annual Meeting of the Psychonomic Society, Toronto, Canada.
- Fine, A. B., Jaeger, T. F., Farmer, T. A., & Qian, T. (2013). Rapid expectation adaptation during syntactic comprehension. *PLoS ONE*, *8*(10), e77661. <https://doi.org/10.1371/journal.pone.0077661>
- Fleischer, Z., Pickering, M. J., & McLean, J. F. (2012). Shared information structure: Evidence from cross-linguistic priming. *Bilingualism*, *15*(3), 568–579. <https://doi.org/10.1017/s1366728911000551>
- Fodor, J. A. (1975). *The language of thought*. Harvard University Press.
- Foppolo, F., Guasti, M. T., & Chierchia, G. (2012). Scalar implicatures in child language: Give children a chance. *Language Learning and Development*, *8*(4), 365–394.
- Foppolo, F., Mazzaggio, G., Panzeri, F., & Surian, L. (2021). Scalar and ad-hoc pragmatic inferences in children: Guess which one is easier. *Journal of Child Language*, 1–23. <https://doi.org/10.1017/S030500092000032X>
- Fox, D. (2007) Free choice and the theory of scalar implicatures. *Presupposition and implicature in compositional semantics*, ed. by U. Sauerland and P. Stateva, 71–120. Basingstoke, UK: Palgrave Macmillan.
- Frank, M. C., & Goodman, N. D. (2012). Predicting pragmatic reasoning in language games. *Science*, *336*(6084), 998–998. <https://doi.org/10.1126/science.1218633>
- Franke, M. (2009). *Signal to act: Game theory in pragmatics* (Doctoral dissertation). University of Amsterdam.
- Franke, M., & Jäger, G. (2016). Probabilistic pragmatics, or why Bayes' rule is probably important for pragmatics. *Zeitschrift Für Sprachwissenschaft*, *35*(1), 3–44. <https://doi.org/10.1515/zfs-2016-0002>
- Franklin, M. S., Dien, J., Neely, J. H., Huber, E., & Waterson, L. D. (2007). Semantic priming modulates the N400, N300, and N400RP. *Clinical Neurophysiology*, *118*(5), 1053–1068. <https://doi.org/10.1016/j.clinph.2007.01.012>
- Friederici, A. D. (2012). The cortical language circuit: From auditory perception to sentence comprehension. *Trends in Cognitive Sciences*, *16*(5), 262–268. <https://doi.org/10.1016/j.tics.2012.04.001>

- Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 360(1456), 815–836. <https://doi.org/10.1098/rstb.2005.1622>
- Garrod, S., & Clark, A. (1993). The development of dialogue co-ordination skills in schoolchildren. *Language and Cognitive Processes*, 8(1), 101-126.
- Garrod, S., & Pickering, M.J. (2013). Dialogue: Interactive Alignment and Its Implications for Language Learning and Language Change.
- Garrod, S., Tosi, A., & Pickering, M. J. (2018). Alignment during interaction. In: Rueschemeyer, S. A., & Gaskell, M. G. (Eds.), *The Oxford Handbook of Psycholinguistics*, 2nd Edition (pp. 575–93). Oxford, UK: Oxford University Press.
- Gazdar, G. (1979). *Pragmatics: Implicature, presupposition and logical form* (2nd ed., reprint). Academic Press.
- Gelman, A., Jakulin, A., Pittau, M. G., & Su, Y.-S. (2008). A weakly informative default prior distribution for logistic and other regression models. *The Annals of Applied Statistics*, 2(4), 1360–1383. <https://doi.org/10.1214/08-AOAS191>
- Gerner, M., & Ling, Z. (2020). Zero morphemes in paradigms. *Studies in Language*, 44, 1-26.
- Geurts, B. (2006). Take “five”: The meaning and use of a number word. In S. Vogeleer & L. Tasmowski (Eds.), *Non-Definiteness and Plurality* (pp. 311–329).
- Geurts, B. (2009). Scalar Implicature and Local Pragmatics. *Mind & Language*, 24(1), 51–79. <https://doi.org/10.1111/j.1468-0017.2008.01353.x>
- Geurts, B. (2010). *Quantity Implicatures*. Cambridge University Press.
- Geurts, B., & Pouscoulous, N. (2008). No scalar inferences under embedding. In P. Égré & G. Magri (Eds.), *Presuppositions and implicatures* (pp. 111–132). MIT Working Papers in Linguistics. (2009)
- Geurts, B., & Pouscoulous, N. (2009). Embedded implicatures?! *Semantics and Pragmatics*, 2(4), 1–34. <https://doi.org/10.3765/sp.2.4>
- Goodman, N. D., & Lassiter, D. (2015). Probabilistic semantics and pragmatics: Uncertainty in language and thought. *Chapter draft*. Stanford University.
- Goodman, N. D., & Stuhlmüller, A. (2013). Knowledge and implicature: Modeling language understanding as social cognition. *Topics in Cognitive Science*, 5(1), 173–184. <https://doi.org/10.1111/tops.12007>
- Gotzner, N., Solt, S., & Benz, A. (2018a). Adjectival scales and three types of implicature. *Semantics and Linguistic Theory*, 28, 409–432.
- Gotzner, N., Solt, S., & Benz, A. (2018b). Scalar diversity, negative strengthening, and adjectival semantics. *Frontiers in Psychology*, 9. <https://doi.org/10.3389/fpsyg.2018.01659>

- Grice, H. P. (1975). Logic and Conversation. *Speech Acts*, 41–58. https://doi.org/10.1163/9789004368811_003
- Grice, H. P. (1989). *Studies in The Way of Words*. Cambridge, MA: Harvard University Press.
- Grodner, D. J., Klein, N. M., Carbary, K. M., & Tanenhaus, M. K. (2010). Some, and possibly all, scalar inferences are not delayed: Evidence for immediate pragmatic enrichment. *Cognition*, 116, 42–55. <https://doi.org/10.1016/j.cognition.2010.03.014>
- Grodner, D., Klein, N., Carbary, K., & Tanenhaus, M. K. (2010). “Some,” and possibly all, scalar inferences are not delayed: Evidence for immediate pragmatic enrichment. *Cognition*, 116(1), 42-55. <https://doi.org/10.1016/j.cognition.2010.03.014>
- Hartsuiker, R. J., & Bernolet, S. (2017). The development of shared syntax in second language learning. *Bilingualism: Language and Cognition*, 20(2), 219–234.
- Hartsuiker, R. J., & Kolk, H. H. J. (1998). Syntactic persistence in Dutch. *Language and Speech*, 41(2), 143–184. <https://doi.org/10.1177/002383099804100202>
- Hartsuiker, R. J., Bernolet, S., Schoonbaert, S., Speybroeck, S., & Vanderelst, D. (2008). Syntactic priming persists while the lexical boost decays: Evidence from written and spoken dialogue. *Journal of Memory and Language*, 58(2), 214–238. <https://doi.org/10.1016/j.jml.2007.07.003>
- Heim, I. (2007). Little. In M. Gibson & J. Howell (Eds.), *Proceedings of SALT 16*. Ithaca, NY: CLC Publications.
- Heyselaar E, Wheeldon L, Segaert K. Structural priming is supported by different components of nondeclarative memory: Evidence from priming across the lifespan. *J Exp Psychol Learn Mem Cogn*. 2021 May;47(5):820-837. doi: 10.1037/xlm0000955. Epub 2020 Nov 5. PMID: 33151717.
- Hirschberg, J. (1985). *A theory of scalar implicature* (Doctoral dissertation). University of Pennsylvania.
- Hirschberg, J. (1991). *A theory of scalar implicature*. Garland Publishing.
- Hochstein, L., Bale, A., & Barner, D. (2018). Scalar implicature in the absence of epistemic reasoning? The case of autism spectrum disorder. *Language Learning and Development*, 14(3), 224–240.
- Hoedemaker, R. S., & Gordon, P. C. (2014). It takes time to prime: Semantic priming in the ocular lexical decision task. *Journal of Experimental Psychology: Human Perception and Performance*, 40(6), 2179–2197. <https://doi.org/10.1037/a0037677>
- Horn, L. (1991). Duplex negatio affirmat...: The economy of double negation. Regional Meeting of the Chicago Linguistic Society. Part Two, 27, 80–106.

- Horn, L. (1984). Toward a new taxonomy for pragmatic inference: Q-based and R-based implicature. In D. Schriften (Ed.), *Meaning, form, and use in context: Linguistic applications* (pp. 11–42). Georgetown University Press.
- Horn, L. (1989). A natural history of negation. Horn, L. (2004). Implicature. In G. Ward & L. Horn (Eds.), *The handbook of pragmatics* (Vol. 26). John Wiley & Sons.
- Horn, L. (2000). Pick a theory (not just any theory): Indiscriminatives and the free-choice indefinite. In L. Horn & Y. Kato (Eds.), *Negation and polarity* (pp. 147–192). Oxford University Press.
- Horn, L. (2005). Current issues in Neo-Gricean pragmatics. *Intercultural Pragmatics*, 2(2), 191–204. <https://doi.org/10.1515/iprg.2005.2.2.191>
- Horn, L. (2007). *Neo-Gricean pragmatics: A Manichaeian manifesto*.
- Horn, L. (2009). WJ-40: Implicature, truth, and meaning. *International Review of Pragmatics*, 1(1), 3–34. <https://doi.org/10.1163/187731009x455847>
- Horn, L. (2012a). Implying and inferring. In K. Allan & K. M. Jaszczolt (Eds.), *The Cambridge Handbook of Pragmatics* (pp. 69–86). Cambridge University Press.
- Horn, L. R. (2012b). Implicature. In G. Russell & D. G. Fara (Eds.), *The Routledge Companion to Philosophy of Language* (pp. 53–66). Routledge.
- Horn, L. R., & Ward, G. (Eds.). (2006). *The handbook of pragmatics*. Blackwell Publishing. <https://doi.org/10.1002/9780470756959>
- Horn, Laurence. (1972). *On the Semantic Properties of Logical Operators in English*.
- Huang, Y. (2003). On neo-Gricean pragmatics. *International Journal of Pragmatics*, 13(1), 87–110
- Huang, Y. (2007). *Pragmatics*. Oxford University Press.
- Huang, Y. (2009). Neo-Gricean Pragmatics and the Lexicon. *International Review of Pragmatics*, 1(1), 118–153. <https://doi.org/10.1163/187731009x455866>
- Huang, Y. (2017). Neo-Gricean pragmatics. In Y. Huang (Ed.), *The Oxford handbook of pragmatics* (online ed.). Oxford Academic. <https://doi.org/10.1093/oxfordhb/9780199697960.013.12>
- Huang, Y. T., Spelke, E., & Snedeker, J. (2013). What Exactly do Numbers Mean? *Language Learning and Development*, 9(2), 105–129. <https://doi.org/10.1080/15475441.2012.658731>
- Huber, D. E., Clark, T. F., Curran, T., & Winkielman, P. (2008). Effects of repetition priming on recognition memory: Testing a perceptual fluency-disfluency model. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34(6), 1305–1324.
- Huttenlocher, J., Vasilyeva, M., & Shimpi, P. (2004). Syntactic priming in young children. *Journal of Memory and Language*, 50(2), 182–195. <https://doi.org/10.1016/j.jml.2003.09.003>

- Ioup, G. (1975). "Some Universals for Quantifier Scope". In *Syntax and Semantics*. Leiden, The Netherlands: Brill. https://doi.org/10.1163/9789004368828_003
- Ito, A., & Husband, E. M. (2022). Investigating shared and distinct mechanisms in semantic and syntactic enrichment: A priming study [Dataset]. *Figshare*. <https://doi.org/10.6084/m9.figshare.19161954>
- Jackson, C. N. (2017). Second language structural priming: A critical review and directions for future research. *Second Language Research*, 34(4), 539–552. <https://doi.org/10.1177/0267658317746207>
- Jaeger, T. F., & Snider, N. E. (2013). Alignment as a consequence of expectation adaptation: Syntactic priming is affected by the prime's prediction error given both prior and recent experience. *Cognition*, 127(1), 57–83. <https://doi.org/10.1016/j.cognition.2012.10.013>
- Janssens, L., & Schaeken, W. (2013). "But" how do we reason with it: An experimental investigation of the implicature stemming from "but." *Journal of Pragmatics*, 57, 90–107.
- JASP Team. (2024). *JASP (Version 0.19.1) [Computer software]*. <https://jasp-stats.org/>
- Jespersen, O. (1924). *The philosophy of grammar*. <https://api.taylorfrancis.com/content/books/mono/download?identifierName=doi&identifierValue=10.4324/9780203716045&type=googlepdf>
- Kadmon, N., & Landman, F (1993). *Any*. *Linguistics and Philosophy* 16(4):353–422.
- Kahneman, D. (2012). A proposal to deal with questions about priming effects. *Nature*. Retrieved from <https://www.nature.com>
- Kalish, C. W., & Sabbagh, M. A. (2007). Conventionality and cognitive development: Learning to think the right way. *New Directions for Child and Adolescent Development*, 2007(115), 1–9. <https://doi.org/10.1002/cd.178>
- Katsos, N., & Bishop, D. (2011). Pragmatic tolerance: Implications for the acquisition of informativeness and implicature. *Cognition*, 120(1), 67–81. <https://doi.org/10.1016/j.cognition.2011.02.015>
- Katsos, N., & Cummins, C. (2010). Pragmatics: From Theory to Experiment and Back Again. *Language and Linguistics Compass*, 4(5), 282–295. <https://doi.org/10.1111/j.1749-818x.2010.00203.x>
- Katsos, N., & Wilson, E. (2016). *Convergence and divergence between word learning and pragmatic inferencing*. Conference: Formal and Experimental Pragmatics Workshop, Tübingen, Germany. https://www.researchgate.net/publication/264942995_Convergence_and_divergence_between_word_learning_and_pragmatic_inferencing
- Katz, J. (1972). *Semantic Theory*. New York: Harper & Row.

- Keen, A. D., & Lee, J. (2022). Structural priming from production to comprehension in aphasia. *Aphasiology*, 38(1), 70–91. <https://doi.org/10.1080/02687038.2022.2159314>
- Khorsheed, A., & Gotzner, N. (2023). A closer look at the sources of variability in scalar implicature derivation: A review. *Frontiers in Communication*, 8. <https://doi.org/10.3389/fcomm.2023.1187970>
- Khorsheed, A., Md Rashid, S., Nimehchisalem, V., Geok Imm, L., Price, J., & Ronderos, C. R. (2022). What second-language speakers can tell us about pragmatic processing. *PLOS ONE*, 17, e0263724. <https://doi.org/10.1371/journal.pone.0263724>
- Khorsheed, A., Price, J., & van Tiel, B. (2022). Sources of cognitive cost in scalar implicature processing: A review. *Frontiers in Communication*, 7. <https://doi.org/10.3389/fcomm.2022.990044>
- Kianifard, F., Neter, J., Wasserman, W., & Kutner, M. H. (1990). Applied linear regression models. *Technometrics*, 32(3), 352. <https://doi.org/10.2307/1269127> (1983)
- Kline, M., Muentener, P., & Schulz, L. (2013). Transitive and periphrastic sentences affect memory for simple causal scenes. Proceedings of the 35th Annual Conference of the Cognitive Science Society.
- Kronmüller, E., & Barr, D. J. (2007). Perspective-free pragmatics: Broken precedents and the recovery-from-preemption hypothesis. *Journal of Memory and Language*, 56(3).
- Kursat, L., & Degen, J. (2020). Probability and processing speed of scalar inferences is context dependent. *CogSci*.
- Kutasi, T., Suffill, E., Gibb, C. L., Sorace, A., Pickering, M. J., & Branigan, H. P. (2018). Shared representation of passives across Scottish Gaelic and English: Evidence from structural priming. *Journal of Cultural Cognitive Science*, 2(1–2), 1–8. <https://doi.org/10.1007/s41809-018-0012-z>
- Ladusaw, W. (1979). *Polarity as inherent scope relations* (Doctoral dissertation). University of Texas, Austin.
- Landman, F. (1998). Plurals and Maximalization. In S. Rothstein (ed.), *Events and grammar*, 237–271. Kluwer, Dordrecht.
- Lassiter, D., & Goodman, N. D. (2015). How many kinds of reasoning? Inference, probability, and natural language semantics. *Cognition*, 136, 123–134. <https://doi.org/10.1016/j.cognition.2014.11.016>
- Lee, J., Man, G., Keen, A., & Castro, N. (2022). Priming sentence production in older adults: Evidence for preserved implicit learning. *Aphasiology*, 38(1), 1–21. <https://doi.org/10.1080/02687038.2022.2153326>
- Lee, M. D., & Wagenmakers, E.-J. (2014). *Bayesian cognitive modeling: A practical course*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139087759>

- Leffel, T., Cremers, A., Gotzner, N., & Romoli, J. (2019). Vagueness in implicature: The case of modified adjectives. *Journal of Semantics*, 36(2), 317–348. <https://doi.org/10.1093/jos/ffy020>
- Lehrer, A. (1985). Markedness and antonymy. *Journal of Linguistics*, 21(3), 397–429.
- Levinson, S. C. (1983). *Pragmatics*. Cambridge: Cambridge University Press.
- Levinson, S. C. (1990). Deixis. In W. Bright (Ed.), *Oxford International Encyclopedia of Linguistics* (pp. 343–344). Oxford University Press.
- Levinson, S. C. (2000). *Presumptive meanings: The theory of generalized conversational implicature*. MIT Press.
- Lewis, D. (1969). *Convention*. Harvard University Press.
- Lim, J. H., & Christianson, K. (2023). Cross-script L1–L2 and L2–L1 masked translation priming and phonological priming: Evidence from unbalanced Korean–English bilinguals. *International Journal of Bilingualism*, 27, 862–881.
- Loebell, H., & Bock, K. (2003). Structural priming across languages. *Linguistics*, 41(5), 791–824. <https://doi.org/10.1515/ling.2003.026>
- Mahowald, K., James, A., Futrell, R., & Gibson, E. (2016). A meta-analysis of syntactic priming in language production. *Journal of Memory and Language*, 91, 5–27. <https://doi.org/10.1016/j.jml.2016.03.009>
- Marina, S. (2006). After Grice: Neo- and Post-perspectives. *Journal of Pragmatics*, 38(12), 2223–2234. <https://doi.org/10.1016/j.pragma.2005.09.007>
- Marjanovic, K., & Crepaldi, D. (2018, November 2). Semantic and morphological cross-word priming during sentence reading. <https://doi.org/10.31219/osf.io/fkb7j>
- Marty, P., Cowan, J., Breheny, R., Romoli, J., & Sudo, Y. (2021, March). What primes what – an experimental framework to explore alternatives for SIs. Paper presentation at the 34th Annual CUNY Conference on Human Sentence Processing, New York, United States.
- Marty, P., Romoli, J., Sudo, Y., & Breheny, R. (2022, April 29). Implicature priming, inverse preference, and context adaptation. *OSF Preprints*. <https://doi.org/10.31219/osf.io/263xf>
- McCawley, J. (1978). Conversational implicature and the lexicon. In P. Cole (Ed.), *Syntax and Semantics* (Vol. 9, pp. 245–259). Academic Press.
- McCawley, James. 1978. Conversational implicature and the lexicon. In P. Cole (ed.), *Syntax and Semantics* 9, 245–259. New York: Academic Press.
- McNamara, T. P., & Holbrook, J. B. (2003). Semantic memory and priming. In A. F. Healy & R. W. Proctor (Eds.), *Handbook of psychology: Experimental psychology* (Vol. 4, pp. 447–474). Wiley. <https://doi.org/10.1002/0471264385.wei0416>

- Metzing, C. & Brennan, S. E. (2003). When conceptual pacts are broken: Partner-specific effects on the comprehension of referring expressions. *Journal of Memory and Language*, 49(2), 201–213
- Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology*, 90(2), 227–234. <https://doi.org/10.1037/h0031564>
- Meyer, M.-C., & Feiman, R. (2021). Priming reveals similarities and differences between three purported cases of implicature: Some, number and free choice disjunctions. *Journal of Memory and Language*, 120, 104206.
- Musolino, J., Crain, S., & Thornton, R. (2000). Navigating negative quantificational space. *Linguistics*, 38(1), 1–32. <https://doi.org/10.1515/ling.38.1.1>
- Nakagawa, E., Koike, T., Sumiya, M., Shimada, K., Makita, K., Yoshida, H., Yokokawa, H., & Sadato, N. (2022). The neural correlates of semantic and grammatical encoding during sentence production in a second language: Evidence from an fMRI study using structural priming. *Frontiers in Human Neuroscience*, 15. <https://doi.org/10.3389/fnhum.2021.753245>
- Neely, J. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner & G. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp. 264–336). Lawrence Erlbaum Associates. (2012)
- Noveck, I. (2018). *Experimental pragmatics: The making of a cognitive science*. Cambridge University Press.
- Noveck, I. A. (2001). When children are more logical than adults: Experimental investigations of scalar implicature. *Cognition*, 78(2), 165–188. [https://doi.org/10.1016/s0010-0277\(00\)00114-1](https://doi.org/10.1016/s0010-0277(00)00114-1)
- Noveck, I., Fogel, M., Van Voorhees, K., & Turco, G. (2022). When eleven does not equal 11: Investigating exactness at a number's upper bound. *PLOS ONE*, 17, e0266920. <https://doi.org/10.1371/journal.pone.0266920>
- Okanda, M., Asada, K., Moriguchi, Y., & Itakura, S. (2015). Understanding violations of Gricean maxims in preschoolers and adults. *Developmental Psychology*, 6, 901.
- Pankratz, E., & Van Tiel, B. (2021). The role of relevance for scalar diversity: A usage-based approach. *Language and Cognition*, 13(4), 562–594. <https://doi.org/10.1017/langcog.2021.13>
- Papafragou, A., & Musolino, J. (2003). Scalar implicatures: Experiments at the semantics–pragmatics interface. *Cognition*, 86(3), 253–282. [https://doi.org/10.1016/s0010-0277\(02\)00179-8](https://doi.org/10.1016/s0010-0277(02)00179-8)
- Papafragou, A., & Skordos, D. (2016). Scalar implicature. In J. Lidz, W. Snyder, & J. Pater (Eds.), *The Oxford Handbook of Developmental Linguistics*. <https://doi.org/10.1093/oxfordhb/9780199601264.013.26>

- Papafragou, A., & Tantalou, N. (2004). Children's Computation of Implicatures. *Language Acquisition*, 12(1), 71–82. https://doi.org/10.1207/s15327817la1201_3
- Parikh, P. (2001). *The use of language*. CSLI Publications.
- Pastor-Cerezuela, G., Tordera Yllescas, J. C., González-Sala, F., Montagut-Asunción, M., & Fernández-Andrés, M. I. (2018). Comprehension of generalized conversational implicatures by children with and without autism spectrum disorder. *Frontiers in Psychology*, 9. <https://doi.org/10.3389/fpsyg.2018.00272>
- Peloquin, B., & Frank, M. (2016). Determining the alternatives for scalar implicature. *Cognitive Science*.
- Pickering MJ, Branigan HP. Syntactic priming in language production. *Trends Cogn Sci*. 1999 Apr;3(4):136-141. doi: 10.1016/s1364-6613(99)01293-0. PMID: 10322467.
- Pickering, M. J., & Branigan, H. P. (1998). The representation of verbs: Evidence from syntactic priming in language production. *Journal of Memory and Language*, 39(4), 633–651. <https://doi.org/10.1006/jmla.1998.2592>
- Pickering, M. J., & Ferreira, V. S. (2008). Structural priming: A critical review. *Psychological Bulletin*, 134(3), 427–459. <https://doi.org/10.1037/0033-2909.134.3.427>
- Pickering, M. J., & Garrod, S. (2004). The interactive-alignment model: Developments and refinements. *Behavioral and Brain Sciences*, 27(2), 212–225. doi:10.1017/S0140525X04450055
- Politzer-Ahles, S. (2020). What can electrophysiology tell us about the cognitive processing of scalar implicatures? *Language and Linguistics Compass*, 14, e12401. <https://doi.org/10.1111/lnc3.12401>
- Politzer-Ahles, S., & Husband, E. M. (2018). Eye movement evidence for context-sensitive derivation of scalar inferences. *Collabra: Psychology*, 4, 1–13. <https://doi.org/10.1525/collabra.100>
- Potts, C. (2005). *The logic of conventional implicatures*. Oxford University Press. (2004)
- Potts, C. (2012). Conventional implicature and expressive content. In K. von Stechow, C. Maienborn, & P. Portner (Eds.), *Semantics: An international handbook of natural language meaning* (Vol. 3, pp. 2516–2536).
- Potts, C. (2013) Stanford Linguistics. Draft of Sep 30, 2013. For the Michigan Philosophy–Linguistics Workshop, Nov 22-24,. 2013.
- Potts, C. (2015a). Embedded implicatures as pragmatic inferences under compositional lexical uncertainty. *Manuscript*.
- Potts, C. (2015b). Negotiating lexical uncertainty and speaker expertise with disjunction. In *Proceedings of the Annual Meeting of the Berkeley Linguistics Society*, 41. <https://doi.org/10.20354/b4414110013>

- Raffray, C. N., & Pickering, M. J. (2010). How Do People Construct Logical Form During Language Comprehension? *Psychological Science*, 21(8), 1090–1097. <https://doi.org/10.1177/0956797610375446>
- Recanati, F. (2003). Embedded Implicatures. *Philosophical Perspectives*, 17(1), 299–332. <https://doi.org/10.1111/j.1520-8583.2003.00012.x>
- Recanati, F., (2004). *Literal meaning*. Cambridge University Press
- Rees, A., & Bott, L. (2017). Structural priming is a useful but imperfect technique for studying all linguistic representations, including those of pragmatics. *Behavioral and Brain Sciences*, 40, e307. <https://doi.org/10.1017/S0140525X17000553>
- Rees, A., & Bott, L. (2018). The role of alternative salience in the derivation of scalar implicatures. *Cognition*, 176, 1–14. <https://doi.org/10.1016/j.cognition.2018.02.024>
- Rees, A., Carter, E., & Bott, L.. (2023). Priming scalar and ad hoc enrichment in children. *Cognition*, 239. <https://doi.org/10.1016/j.cognition.2023.105572> (Rees, Carter, Bott)
- Rett, J. (2014). The Semantics of Evaluativity.
- Rett, J. (2020). Manner implicatures and how to spot them. *International Review of Pragmatics*, 12(1), 44–79. <https://doi.org/10.1163/18773109-01201105>
- Roediger, H. L., & Blaxton, T. A. (1987). Effects of varying modality, surface features, and retention interval on priming in word-fragment completion. *Memory & Cognition*, 15(5), 379–388. <https://doi.org/10.3758/bf03197728>
- Ronai, E., & Xiang, M. (2021). Pragmatic inferences are QUD-sensitive: An experimental study. *Journal of Linguistics*, 57(4), 841–870. <https://doi.org/10.1017/s0022226720000389>
- Russell, B. (2006). Against grammatical computation of scalar implicatures. *Journal of Semantics*, 23(4), 361–382. <https://doi.org/10.1093/jos/ffl008>
- Ruys, E., & Winter, Y. (2010). Quantifier scope in formal linguistics. In *Springer eBooks* (pp. 159–225). https://doi.org/10.1007/978-94-007-0479-4_3
- Sapir, E. (1944). Grading, A Study in Semantics. *Philosophy of Science*, 11(2), 93–116. <http://www.jstor.org/stable/184355>
- Sauerland, U. (2012). The computation of scalar implicatures: Pragmatic, lexical or grammatical? *Language and Linguistics Compass*, 6(1), 36–49. <https://doi.org/10.1002/lnc3.321>
- Sbisà, M. (2006). After Grice: Neo- and post-perspectives. *Journal of Pragmatics*, 38(12), 2223–2234. <https://doi.org/10.1016/j.pragma.2005.09.007>
- Schacter, D.L. & Buckner, R. L. (1998). Priming and the brain. *Neuron*, 20(2), 185–195. [https://doi.org/10.1016/S0896-6273\(00\)80448-1](https://doi.org/10.1016/S0896-6273(00)80448-1)

- Schoonbaert, S., Hartsuiker, R. J., & Pickering, M. J. (2007). The representation of lexical and syntactic information in bilinguals: Evidence from syntactic priming. *Journal of Memory and Language*, 56(2), 153–171. <https://doi.org/10.1016/j.jml.2006.10.002>
- Schwarz, F. (2017). Experimental pragmatics. In M. Aronoff (Ed.), *The Oxford research encyclopedia of linguistics*. Oxford University Press.
- Segaert, K., Menenti, L., Weber, K., & Hagoort, P. (2011). A paradox of syntactic priming: Why response tendencies show priming for passives, and response latencies show priming for actives. *PLoS ONE*, 6(10), e24209. <https://doi.org/10.1371/journal.pone.0024209>
- Segaert, K., Wheeldon, L., & Hagoort, P. (2016). Unifying structural priming effects on syntactic choices and timing of sentence generation. *Journal of Memory and Language*, 91, 59–80. <https://doi.org/10.1016/j.jml.2016.03.011>
- Sevold, C. A., Dell, G. S., & Cole, J. S. (1995). Syllable structure in speech production: Are syllables chunks or schemas? *Journal of Memory and Language*, 34(6), 807–820. <https://doi.org/10.1006/jmla.1995.1036>
- Skordos D, Papafragou A. (2016) *Children's derivation of scalar implicatures: Alternatives and relevance*. *Cognition*. 2016 Aug;153:6-18. doi: 10.1016/j.cognition.2016.04.006. Epub 2016 Apr 21. PMID: 27107741.
- Smith, M., & Wheeldon, L. (2001). Syntactic priming in spoken sentence production – an online study. *Cognition*, 78(2), 123–164. [https://doi.org/10.1016/s0010-0277\(00\)00110-4](https://doi.org/10.1016/s0010-0277(00)00110-4)
- Snoeren, N. D., Seguí, J., & Hallé, P. A. (2008). On the role of regular phonological variation in lexical access: Evidence from voice assimilation in French. *Cognition*, 108(2), 512–521. <https://doi.org/10.1016/j.cognition.2008.02.008>
- Son, M. (2020). Cross-linguistic syntactic priming in Korean learners of English. *Applied Psycholinguistics*, 41(5), 1223–1247. <https://doi.org/10.1017/s0142716420000545>
- Song, G., & Wolff, P. (2005). Linking perceptual properties to linguistic expressions of causation. In M. Achard & S. Kemmer (Eds.), *Language, culture and mind* (pp. 237–250). Stanford, CA: CSLI Publications.
- Spector, B. (2007). Aspects of the Pragmatics of Plural Morphology: On Higher-Order Implicatures.
- Sperber, D. and Wilson, D. (1995) *Relevance: Communication and Cognition*. 2nd Edition, Blackwell, Oxford.
- Sperber, D., & Wilson, D. (2002). Pragmatics, modularity, and mind-reading. *Mind & Language*, 17(1-2), 3–23. <https://doi.org/10.1111/1468-0017.00186>
- Spychalska, M., Kontinen, J., & Werning, M. (2016). Investigating scalar implicatures in a truth-value judgement task: Evidence from event-related brain potentials. *Language*,

Cognition and Neuroscience, 31(6), 817–840.

<https://doi.org/10.1080/23273798.2016.1161806>

Stiller, A. J., Goodman, N. D., & Frank, M. C. (2015). Ad-hoc implicature in preschool children. *Language Learning and Development*, 11(2), 176–190.

Taguchi, N., & Yamaguchi, S. (2019). Implicature comprehension in L2 pragmatics research. In N. Taguchi (Ed.), *The Routledge handbook of second language acquisition and pragmatics* (1st ed., pp. 16). Routledge. <https://doi.org/10.4324/9781351164085-3>

Tessler, M. H., & Franke, M. (2019). Not unreasonable: Why two negatives don't make a positive. *PsyArXiv Preprints*. <https://doi.org/10.31234/osf.io/tqjr2>

The Wall Street Journal. (2014, October 1). *Aretha Franklin on Adele, Taylor Swift, and divas* [Video]. The Wall Street Journal. <https://www.wsj.com/video/aretha-franklin-on-adele-taylor-swift-and-divas/DD876BBC-873B-45DA-A75A-7A824C622843>

Tieu, L., Romoli, J., Zhou, P., & Crain, S. (2015). Children's knowledge of free choice inferences and scalar implicatures. *Journal of Semantics*, 33(2), 269–298. <https://doi.org/10.1093/jos/ffv001>

Tomlinson, J. M., Bailey, T. M., & Bott, L. (2013). Possibly all of that and then some: Scalar implicatures are understood in two steps. *Journal of Memory and Language*, 69, 18–35. <https://doi.org/10.1016/j.jml.2013.02.003>

Tooley, K. M. (2023). Structural priming during comprehension: A pattern from many pieces. *Psychonomic Bulletin & Review*, 30(3), 882–896. <https://doi.org/10.3758/s13423-022-02209-7>

Tytus, A. E., & Rundblad, G. (2016). Cross-language priming as a means of investigating bilingual conceptual representations. *Linguistic Approaches to Bilingualism*, 6(4), 440–466. <https://doi.org/10.1075/lab.14020.tyt>

Van Boxtel, W., Cox, B., Keen, A., & Lee, J. (2023). Planning sentence production in aphasia: Evidence from structural priming and eye-tracking. *Frontiers in Language Sciences*, 2. <https://doi.org/10.3389/flang.2023.1175579>

Van Rooij, R., & Schulz, K. (2004). Exhaustive interpretation of complex sentences. *Journal of Logic Language and Information*, 13(4), 491–519. <https://doi.org/10.1007/s10849-004-2118-6>

van Tiel, B., & Pankratz, E. (2021). Adjectival polarity and the processing of scalar inferences. *Glossa: A Journal of General Linguistics*, 6(1), 32. <https://doi.org/10.5334/gjgl.1457>

van Tiel, B., Van Miltenburg, E., Zevakhina, N., & Geurts, B. (2014). Scalar diversity. *Journal of Semantics*, 31(2), 147–177. <https://doi.org/10.1093/jos/ffu017>

van Zwet, J. L., & Unsworth, S. (2024). Individual differences in structural priming in bilingual and monolingual children: the influence of perspective-taking. *Language and Cognition*, 16(4), 1026–1052. doi:10.1017/langcog.2023.71

Vicente, B. (2024). When is it legitimate to cancel a potential scalar implicature? The roles of the question under discussion and optimal relevance. *University of the Basque Country*. Retrieved from [SSRN](#)

von Stutterheim, C., & Klein, W. (1989). Referential movement in descriptive and narrative discourse. In R. Dietrich & C. F. Graumann (Eds.), *Language processing in social context* (pp. 39–76). North Holland.

Wada, R., Gillam, S.L., & Gillam, R.B. (2020). The Use of Structural Priming and Focused Recasts to Facilitate the Production of Subject- and Object-Focused Relative Clauses by School-Age Children With and Without Developmental Language Disorder. *American journal of speech-language pathology*, 1-13 .

Waldon, B., & Degen, J. (2020). Symmetric alternatives and semantic uncertainty modulate scalar inference. *Cognitive Science*.

<https://cogsci.mindmodeling.org/2020/papers/0029/index.html>

Ward, E. V. (2023). Age and processing effects on perceptual and conceptual priming.

Quarterly Journal of Experimental Psychology, 76(1), 1–14.

<https://doi.org/10.1177/1747021822109012>

Wei, R., Kim, S.-A., & Shin, J.-A. (2022). Structural priming and inverse preference effects in L2 grammaticality judgment and production of English relative clauses. *Frontiers in Psychology*, 13. <https://doi.org/10.3389/fpsyg.2022.845691>

Weingarten, E., Chen, Q., McAdams, M., Yi, J., Hepler, J., & Albarracín, D. (2016). From primed concepts to action: A meta-analysis of the behavioral effects of incidentally presented words. *Psychological Bulletin*, 142(5), 472–497. <https://doi.org/10.1037/bul0000030>

Wetzels, R., & Wagenmakers, E.-J. (2012). A default Bayesian hypothesis test for correlations and partial correlations. *Psychonomic Bulletin & Review*, 19(6), 1057–1064.

<https://doi.org/10.3758/s13423-012-0295-x>

Wilson E, Katsos N (2023). In a manner of speaking: progress and challenges in the study of the elusive manner implicature. PsyArXiv; 2023.

















Wilson, E. (2017). Children’s development of Quantity, Relevance and Manner implicature understanding and the role of the speaker’s epistemic state (*Doctoral dissertation*). University of Cambridge.

Wolff, P. (2003). Direct causation in the linguistic coding and individuation of causal events. *Cognition*, 88(1), 1–48. [https://doi.org/10.1016/S0010-0277\(03\)00004-0](https://doi.org/10.1016/S0010-0277(03)00004-0)





- Yang, X., Minai, U., & Fiorentino, R. (2018). Context-sensitivity and individual differences in the derivation of scalar implicature. *Frontiers in Psychology, 9*.
<https://doi.org/10.3389/fpsyg.2018.01720>
- Zeng, T., Mu, Y., & Zhu, T. (2021). Structural priming from simple arithmetic to Chinese ambiguous structures: Evidence from eye movement. *Cognitive Processing, 22*(2), 185–207.
<https://doi.org/10.1007/s10339-020-01003-4>
- Zhang, J., & Wu, Y. (2022). Epistemic reasoning in pragmatic inferencing by non-native speakers: The case of scalar implicatures. *Second Language Research*.
<https://doi.org/10.1177/02676583211069735>
- Zhang, Y. (2023). The role of phonological priming in backward priming effects of Chinese four-character words. *International Journal of Languages, Literature and Linguistics*.
- Zipf, G. K. (1949). *Human Behavior and the Principle of Least Effort: An Introduction to Human Ecology*. Cambridge, MA: Addison-Wesley.
- Zondervan, A., Meroni, L., & Gualmini, A. (2008). Experiments on the role of the question under discussion for ambiguity resolution and implicature computation in adults. *Proceedings of SALT, 18*, 765–777.

9. Appendices




















9.1. The item pairs used in the MI and I-Implicature critical and prime trials

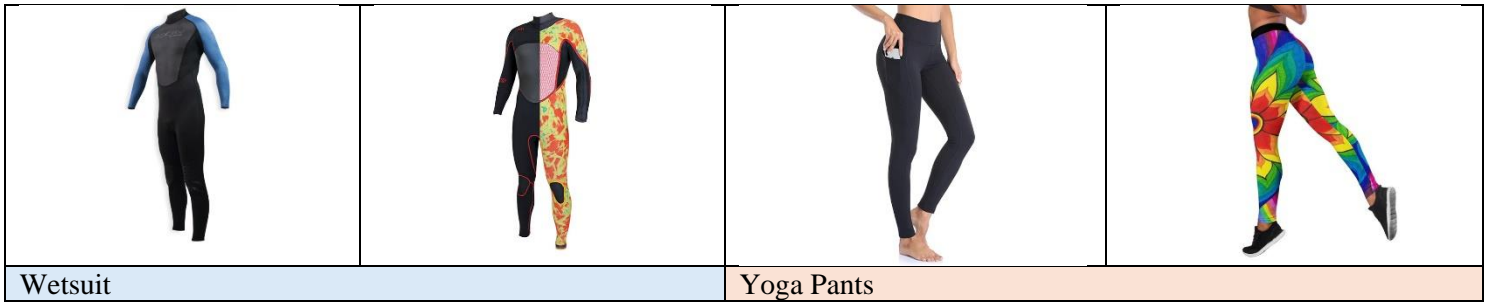
Unmarked	Marked	Unmarked	Marked
			
Apron		Bathing Suit	
			
Bathrobe		Beanie	
			
Belt		Bowtie	
			
Boxer Briefs		Bra	

			
Cap		Cowboy Boot	

Unmarked	Marked	Unmarked	Marked
			
Dress		Earring	
			
Flip Flops		Glasses	
			
Gloves		Goggles	
			
Hat		Handbag	
			
Headband		Headphones	

Unmarked	Marked	Unmarked	Marked
			
High Heels		Jeans	
			
Necklace		Polo Shirt	
			
Puffer Coat		Purse	
			
Ring		Scarf	
			
Shirt		Skirt	

Unmarked	Marked	Unmarked	Marked
			
Socks		Sunglasses	
			
Swimming Trunks		Tie	
			
Tights		T-shirt	
			
Tuxedo		Vest	
			
Wallet		Wedding Dress	



9.2. Trial Blocks (see Error! Bookmark not defined. for context)

Group 1	Group 2	Group 3
Bathrobe	Bathing Suit	Wetsuit
Cowboy Boot	Hat	Beanie
Purse	Goggles	Glasses
Wallet	Vest	Tuxedo
Headphones	Apron	Wedding Dress
Bowtie	Belt	Yoga Pants
Dress	Cap	Bra
Necklace	Headband	Handbag
Polo Shirt	Jeans	Flip Flops
Scarf	Ring	Puffer Coat
Tie	Skirt	Shirt
Boxer Briefs	Earring	T-shirt
Socks	High Heels	Gloves
Tights	Swimming Trunks	Sunglasses

9.3. MI Critical And Prime Trial Captions

Item	Caption
Apron	<i>I see a garment worn to protect clothing underneath</i>
Bathing Suit	<i>I see a garment used for swimming</i>
Bathrobe	<i>I see a gown made of towelling material</i>
Beanie	<i>I see a close-fitting hat</i>
Belt	<i>I see a strap of material worn around the waist to hold the trousers up</i>
Bowtie	<i>I see a necktie in the shape of a bow</i>
Boxer Briefs	<i>I see a garment worn underneath trousers</i>
Bra	<i>I see a garment worn over the breasts</i>
Cap	<i>I see a hat with a peak</i>
Cowboy Boot	<i>I see a pair of high-heeled boots with a pointed toe</i>
Dress	<i>I see a one-piece garment</i>
Earring	<i>I see a piece of jewellery worn on the earlobe</i>
Flip Flops	<i>I see light sandals with a thong between the toes</i>
Glasses	<i>I see a pair of lenses set in a frame that are used to correct vision</i>
Gloves	<i>I see some coverings for the hands</i>
Goggles	<i>I see some glasses used underwater</i>
Handbag	<i>I see a bag used for carrying personal items</i>
Hat	<i>I see a shaped covering for the head</i>
Headband	<i>I see a decorative band of fabric worn around the head</i>
Headphones	<i>I see a pair of earphones with a band worn over the head</i>
High Heels	<i>I see women's shoes that elevate the foot with a heel</i>
Jeans	<i>I see some casual trousers made from denim</i>
Necklace	<i>I see a piece of jewellery worn around the neck</i>
Polo Shirt	<i>I see a casual short-sleeved garment with buttons at the neck</i>
Puffer Coat	<i>I see an outer garment worn to keep warm</i>
Purse	<i>I see a small leather pouch for putting money in</i>
Ring	<i>I see a piece of jewellery worn on the finger</i>
Scarf	<i>I see some fabric worn around the neck</i>
Shirt	<i>I see a garment with a collar and sleeves and buttons down the front</i>
Skirt	<i>I see a piece of clothing that sits on the waist and hangs down the legs</i>
Socks	<i>I see an item of clothing worn on the feet</i>
Sunglasses	<i>I see a pair of tinted lenses in a frame</i>
Swimming Trunks	<i>I see trouser-like garments that are cropped at the leg</i>
Tie	<i>I see a strip of material worn around the collar</i>
Tights	<i>I see a close fitting, trouser-like, undergarment</i>
T-shirt	<i>I see a short-sleeved casual top</i>
Tuxedo	<i>I see a men's formal evening suit</i>
Vest	<i>I see an undergarment worn on the upper body</i>
Wallet	<i>I see a flat-folding case used to hold money</i>
Wedding Dress	<i>I see a dress that's worn by a bride</i>
Wetsuit	<i>I see a close-fitting suit used underwater</i>
Yoga Pants	<i>I see stretchy trousers used for exercise</i>

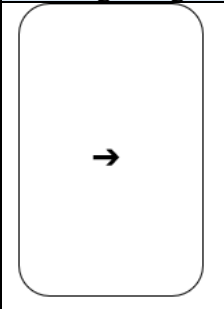
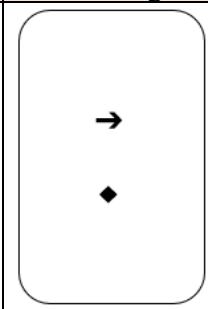
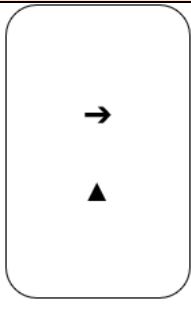
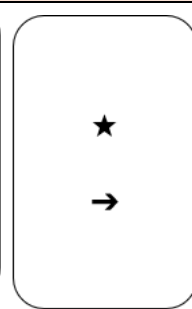
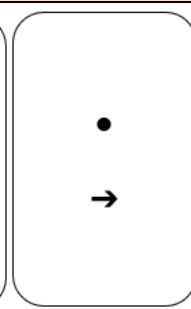
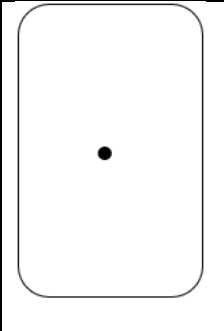
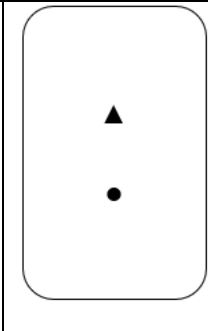
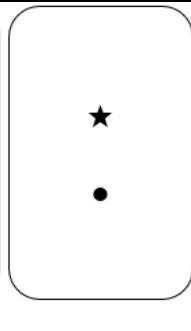
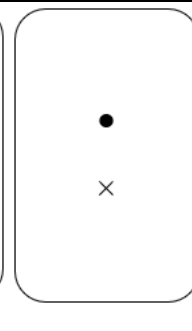
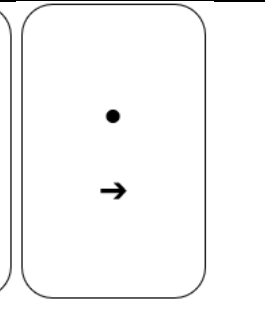
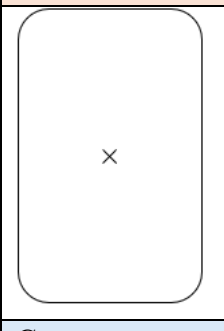
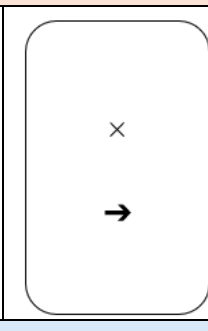
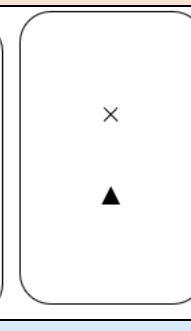

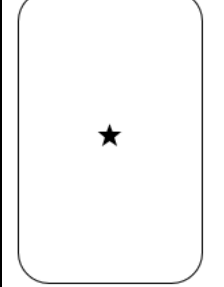
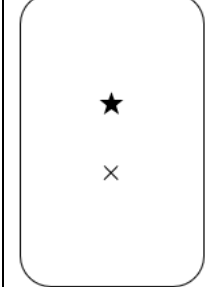
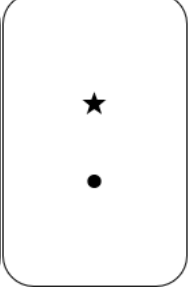
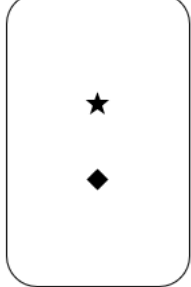
9.4. I-Implicature Critical And Prime Trial Captions

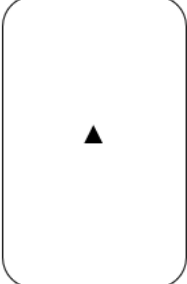
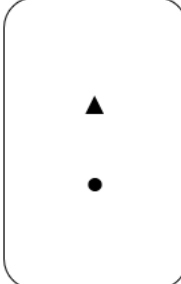
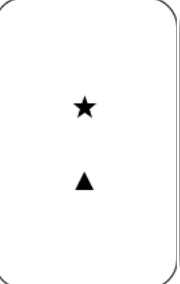
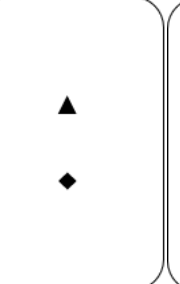
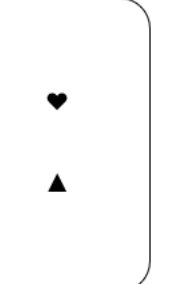
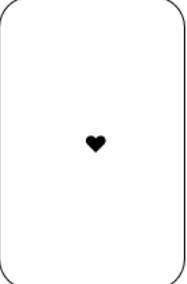

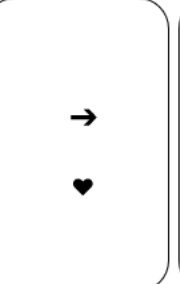
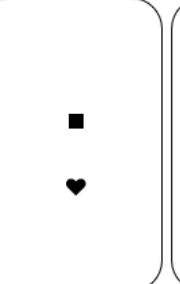

Item	Caption
Apron	<i>I can see an apron</i>
Bathing Suit	<i>I can see a bathing suit</i>
Bathrobe	<i>I can see a bathrobe</i>
Beanie	<i>I can see a beanie</i>
Belt	<i>I see a belt</i>
Bowtie	<i>I see a bowtie</i>
Boxer Briefs	<i>I see some boxer briefs</i>
Bra	<i>I see a bra</i>
Cap	<i>I see a cap</i>
Cowboy Boot	<i>I can see a pair of cowboy boots</i>
Dress	<i>I see a dress</i>
Earring	<i>I see an earring</i>
Flip Flops	<i>I see a pair of flip-flops</i>
Glasses	<i>I can see a pair of glasses</i>
Gloves	<i>I see some gloves</i>
Goggles	<i>I can see a pair of goggles</i>
Handbag	<i>I see a handbag</i>
Hat	<i>I can see a hat</i>
Headband	<i>I see a headband</i>
Headphones	<i>I can see some headphones</i>
High Heels	<i>I see some high-heels</i>
Jeans	<i>I see a pair of jeans</i>
Necklace	<i>I see a necklace</i>
Polo Shirt	<i>I see a polo-shirt</i>
Puffer Coat	<i>I see a puffer coat</i>
Purse	<i>I can see a purse</i>
Ring	<i>I see a ring</i>
Scarf	<i>I see a scarf</i>
Shirt	<i>I see a shirt</i>
Skirt	<i>I see a skirt</i>
Socks	<i>I see some socks</i>
Sunglasses	<i>I see some sunglasses</i>
Swimming Trunks	<i>I see some swimming trunks</i>
Tie	<i>I see a tie</i>
Tights	<i>I see some tights</i>
T-shirt	<i>I see a t-shirt</i>
Tuxedo	<i>I can see a tuxedo</i>
Vest	<i>I can see a vest</i>
Wallet	<i>I can see a wallet</i>
Wedding Dress	<i>I can see a wedding dress</i>
Wetsuit	<i>I can see a wetsuit</i>
Yoga Pants	<i>I can see some yoga pants</i>

9.5. Experiment 4: Ad-hoc Captions

Item	Caption
Arrow	<i>Yes, I see a card with an arrow</i>
Circle	<i>Yes, I see a card with a circle</i>
Cross	<i>Yes, I see a card with a cross</i>
Star	<i>Yes, I see a card with a star</i>
Heart	<i>Yes, I see a card with a heart</i>
Triangle	<i>Yes, I see a card with a triangle</i>

9.6. Experiment 4: Ad-Hoc Items

Strong Image	Weak Images				
					
Arrow					
					
Circle					
					
Cross					
					
Star					



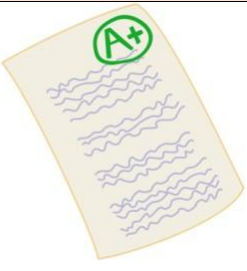
Strong Image	Weak Images				
					
Triangle					
					
Heart					
(N.B. Circle + Arrow, Triangle + Circle, Triangle + Triangle Diamond appeared twice each)					

Experiment 4: Ad-hoc Items


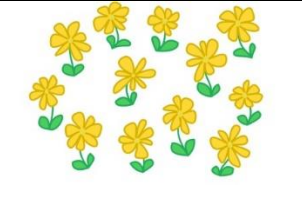
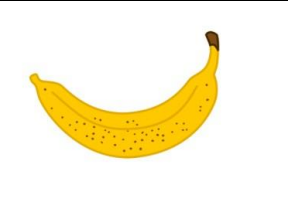
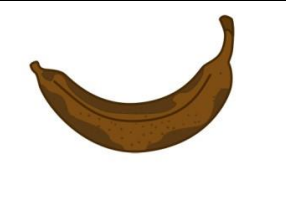


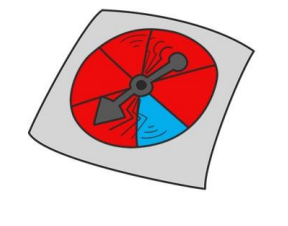
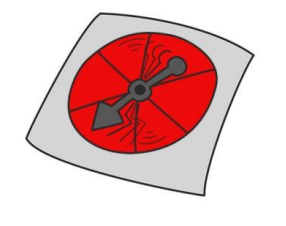
9.7. Experiment 5: Adjectival Scalar Captions

Item	Caption	Horn Scale
Fair	<i>I see a target with fair accuracy</i>	< fair, complete >
Chubby	<i>I can see a chubby man</i>	< chubby, obese >
Content	<i>I can see a content lady</i>	< content, happy >
Cool	<i>I can see a glass of cool water</i>	< cool, cold >
Ajar	<i>I can see an ajar door</i>	< ajar, open >
Drizzly	<i>I can see some drizzly weather</i>	< drizzly, rainy >
Breezy	<i>I can see some breezy weather</i>	< breezy, windy >
Low	<i>I see a bath in which the water is low</i>	< empty, low >
Mediocre	<i>I can see a test with a mediocre grade</i>	< mediocre, good >
Passable	<i>I can see a test with a passable score</i>	< passable, bad >
Ripe	<i>I can see a ripe banana</i>	< ripe, overripe >
Scarce	<i>I can see a field in which red flowers are scarce</i>	< scarce, absent >
Sleepy	<i>I can see a sleepy woman</i>	< sleepy, asleep >
Unlikely	<i>I can see a spinner that's unlikely to land on blue</i>	< unlikely, impossible >

9.8. Experiment 5: Adjectival Scalar Items

	Target 	Non-Target 	Target 	Non-Target 
Item:	Accuracy		Chubby	
Caption:	<i>I see a target with fair accuracy</i>		<i>I can see a chubby man</i>	
				
Item:	Content		Drizzly	
Caption:	<i>I can see a content lady</i>		<i>I can see some drizzly weather</i>	
				
Item:	Cool		Ajar	
Caption:	<i>I can see a glass of cool water</i>		<i>I can see an ajar door</i>	
				
Item:	Breezy		Low	
Caption:	<i>I can see some breezy weather</i>		<i>I see a bath in which the water is low</i>	
				
Item:	Mediocre		Passable	
Caption:	<i>I can see a test with a mediocre grade</i>		<i>I can see a test with a passable score</i>	

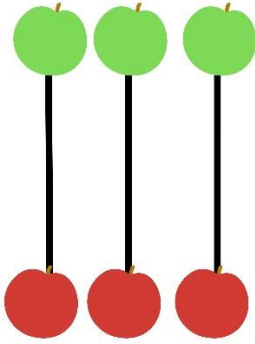
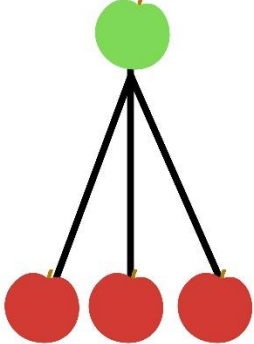
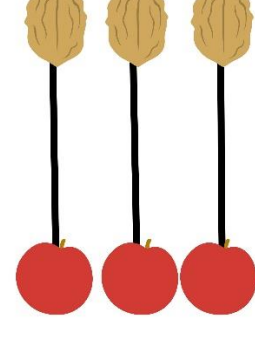
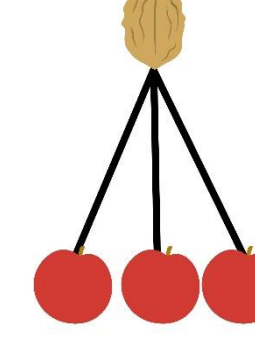
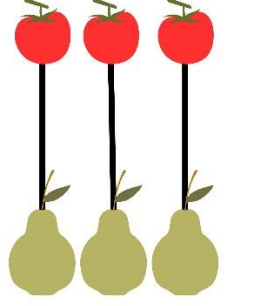
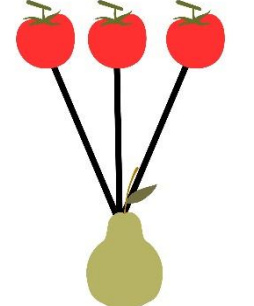
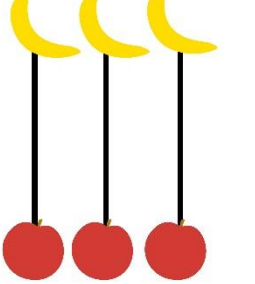
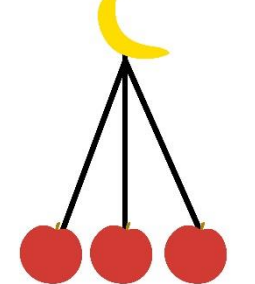
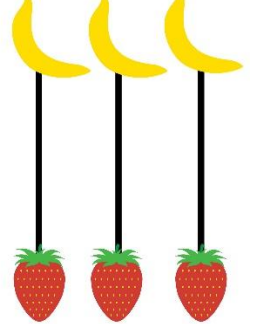
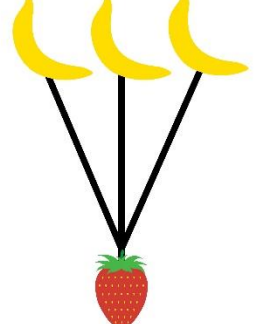
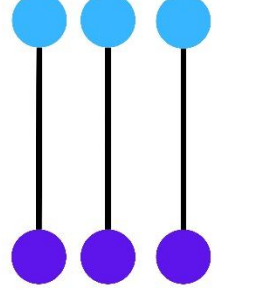
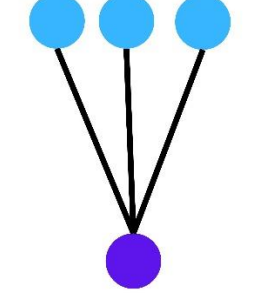
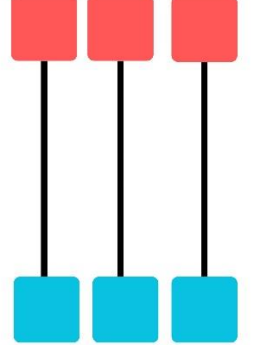
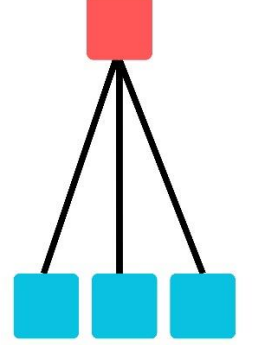
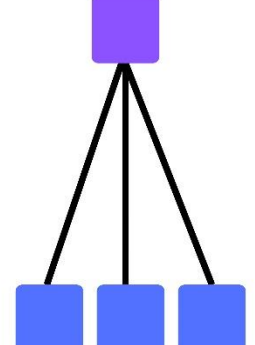
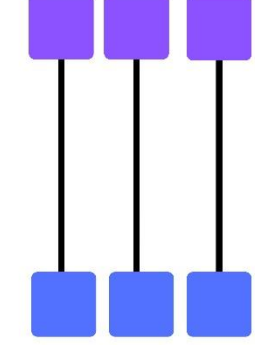
Experiment 5: Adjectival Scalar Items

	Target	Non-Target	Target	Non-Target
				
Item:	Scarce		Ripe	
Caption:	<i>I can see a field in which red flowers are scarce</i>		<i>I can see a ripe banana</i>	
				
Item:	Sleepy		Unlikely	
Caption:	<i>I can see a sleepy woman</i>		<i>I can see a spinner that's unlikely to land on blue</i>	

9.9. Experiment 6: Scope Ambiguity Captions

Item	Caption
Apple	<i>I see a picture in which each apple is connected to a nut</i>
Apple	<i>I see a picture in which each apple is connected to a pear</i>
Apple	<i>I see a picture in which each apple is connected to a banana</i>
Banana	<i>I see a picture in which each banana is connected to a strawberry</i>
Blue Circle	<i>I see an image in which each blue circle is connected to a purple one</i>
Blue Circle	<i>I see an image in which each blue circle is connected to a purple one</i>
Blue Square	<i>I see a picture in which each blue square is connected to another one</i>
Blue Square	<i>I see a picture in which each blue square is connected to a pink one</i>
Blue Square	<i>I see a picture in which each blue square is connected to a purple one</i>
Blue Square	<i>I see a picture in which each blue square is connected to a purple one</i>
Blue Square	<i>I see a picture in which each blue square is connected to a red one</i>
Blue Square	<i>I see a picture in which each blue square is connected to a red one</i>
Green Circle	<i>I see a picture in which each green circle is connected to a yellow one</i>
Green Square	<i>I see a picture in which each green square is connected to a red one</i>
Heart	<i>I see a picture in which each heart is connected to a triangle</i>
Heart	<i>I see a picture in which each heart is connected to a triangle</i>
Heart	<i>I see a picture in which each heart is connected to a triangle</i>
Heart	<i>I see a picture in which each heart is connected to a triangle</i>
Heart	<i>I see a picture in which each heart is connected to a triangle</i>
Heart	<i>I see a picture in which each heart is connected to a triangle</i>
Heart	<i>I see a picture in which each heart is connected to a triangle</i>
Lemon	<i>I see a picture in which each lemon is connected to a lime</i>
Lemon	<i>I see a picture in which each lemon is connected to a lime</i>
Orange	<i>I see a picture in which each orange is connected to a mango</i>
Purple Circle	<i>I see a picture in which each purple circle is connected to a blue one</i>
Purple Circle	<i>I see a picture in which each purple circle is connected to an orange one</i>
Red Apple	<i>I see a picture in which each red apple is connected to a green apple</i>
Red Circle	<i>I see a picture in which each red circle is connected to a grey one</i>
Red Circle	<i>I see a picture in which each red circle is connected to a green one</i>

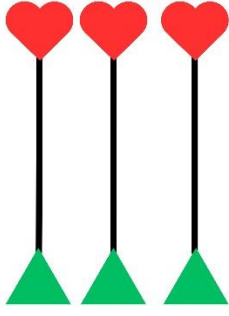
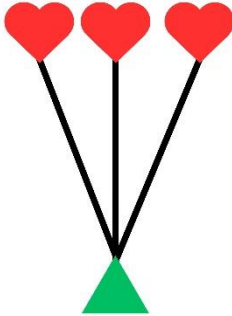
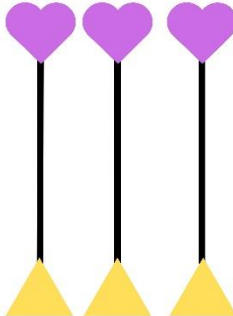
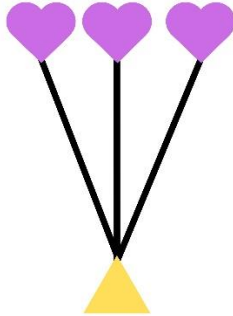
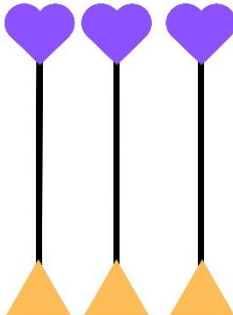
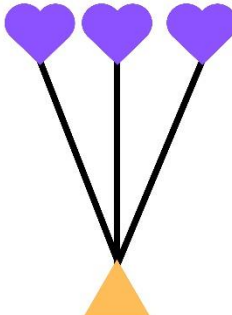
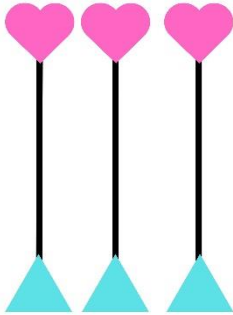
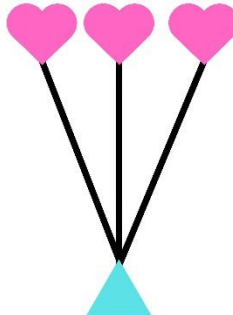
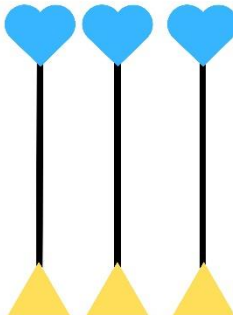
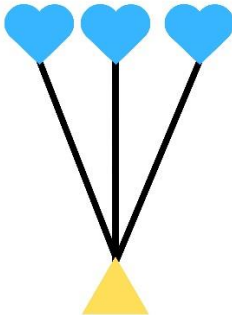
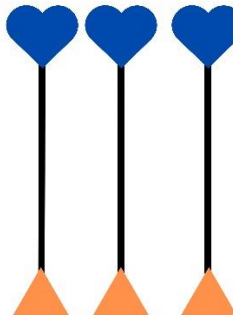
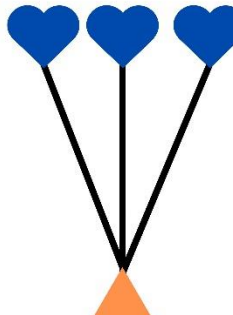
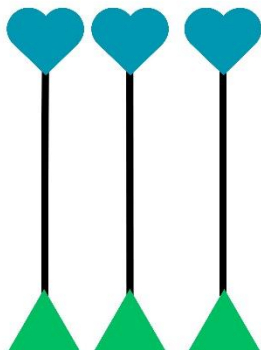
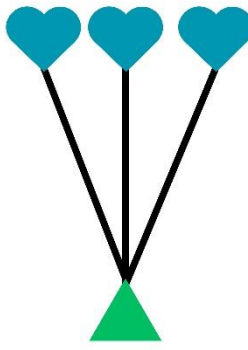
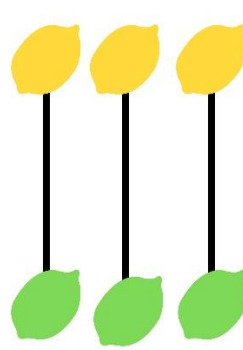
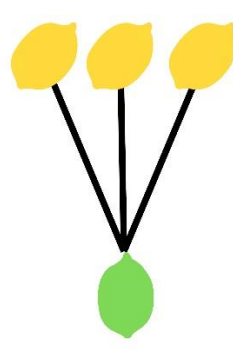
9.10. Experiment 6: Scope Ambiguity Items

			
Red Apple		Apple	
			
Apple		Apple	
			
Banana		Blue Circle	
			
Blue Square		Blue Square	

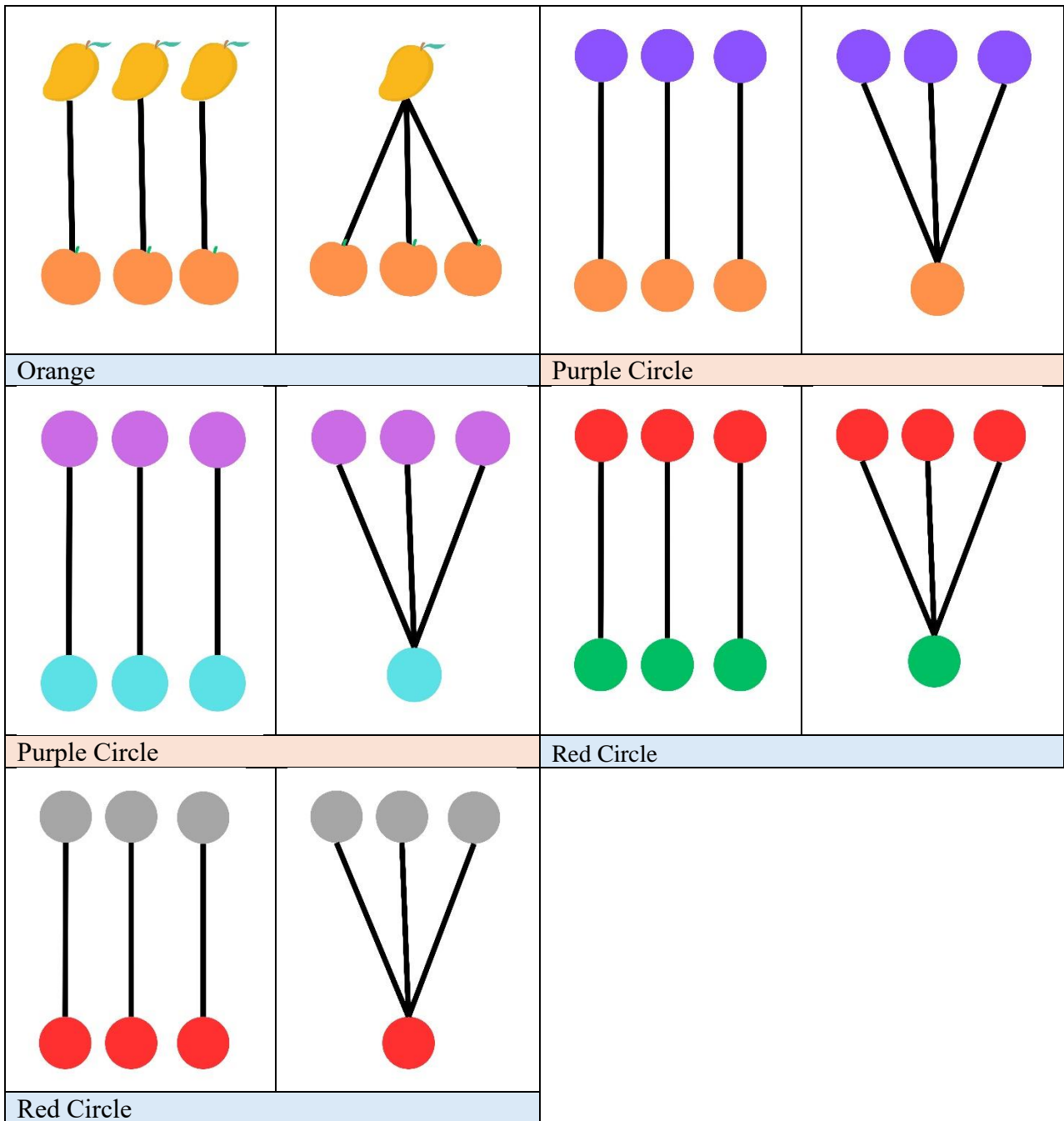
Experiment 6: Scope Ambiguity Items

Blue Square		Blue Square	
Blue Square		Blue Square	
Green Square		Green Circle	
Blue Square		Blue Square	













Experiment 6: Scope Ambiguity Items

			
Heart		Heart	
			
Heart		Heart	
			
Heart		Heart	
			
Heart		Lemon	

Experiment 6: Scope Ambiguity Items



9.11. Filler Trials

<p><i>Type</i></p> <p><i>Yes</i></p>	<p>Dr. Smith: 'Can you see a cat?'</p> <p>Alan: 'Yes, I can'</p> <p>What image can Alan see?</p>    
<p><i>No-Marked- Unmarked</i></p>	<p>Dr. Smith: 'Can you see a hammer?'</p> <p>Alan: 'No, I can't'</p> <p>What image can Alan see?</p>    
<p><i>No-False- False</i></p>	<p>Dr. Smith: 'Can you see a burger?'</p> <p>Alan: 'No, I can't'</p> <p>What image can Alan see?</p>    

9.12. Consent Form (All Experiments)

Consent

This experiment has been approved by the Ethics Committee.

No harm will come to you from taking part in this experiment. You have the right to stop at any time.

Thank you for agreeing to take part in this experiment! Before we continue, we need your consent to the following:

1. I consent to performing the task online
2. I understand and consent to my responses are being recorded and stored securely in a database
3. I understand and consent to my responses may be used anonymously for secondary research in the future

I consent to items 1-3 above

Next



9.13. Demographic Questionnaire

Demographics

Please answer the following questions about yourself.

Are you...

- Female
 Male
 Prefer Not To Say
 Other (please specify)

How old are you in years? (e.g., 25)

Next