



# Vagueness, presupposition and truth-value judgements

Jérémy Zehr

## ► To cite this version:

Jérémy Zehr. Vagueness, presupposition and truth-value judgements. Philosophy. Ecole normale supérieure - ENS PARIS, 2014. English. NNT : 2014ENSU0018 . tel-01303747

**HAL Id: tel-01303747**

**<https://theses.hal.science/tel-01303747>**

Submitted on 13 May 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Vagueness, Presupposition and Truth-Value Judgments

Jérémy Zehr

Institut Jean Nicod, UMR 8128  
École Normale Supérieure de Paris

## Committee Members

*Paul Égré\**, *Orin Percus\**  
*Reinhard Muskens\*\**, *Florian Schwarz\*\**  
*Márta Abrusán*, *Benjamin Spector*

A thesis submitted for the degree of  
*Doctor in Linguistics*

18 December 2014

\*Supervisors

\*\*Reviewers

# Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
1.1	Introducing Vagueness . . . . .	4
1.2	Introducing Presupposition . . . . .	8
1.3	The Project of Unifying Vagueness and Presupposition . . . . .	12
1.3.1	Previous M.A. Work . . . . .	12
1.3.2	Present Work . . . . .	15
<b>2</b>	<b>Trivalence, Vagueness and Presupposition</b>	<b>19</b>
2.1	Supervaluationism . . . . .	20
2.1.1	Presupposition: Van Fraassen's Proposal . . . . .	22
2.1.2	Vagueness: Fine's implementation . . . . .	27
2.2	Strong Kleene . . . . .	29
2.2.1	Presupposition: George's Implementation . . . . .	31
2.2.2	Vagueness: Tye's implementation . . . . .	36
2.3	Predictions and Comparisons . . . . .	38
<b>3</b>	<b>Experimental Literature on Vagueness and on Presupposition</b>	<b>47</b>
3.1	Vagueness . . . . .	48
3.2	Presupposition . . . . .	51
3.3	Summary . . . . .	57
<b>4</b>	<b>ST5: a 5-Valued System</b>	<b>59</b>
4.1	Truth-Value Judgments . . . . .	60
4.2	ST5 . . . . .	61
4.2.1	The Original 3-Valued ST System . . . . .	61
4.2.1.1	Two Notions of Satisfaction . . . . .	61
4.2.1.2	No Room for Presupposition . . . . .	63
4.2.2	Why Exclude Alternative Systems . . . . .	63
4.2.3	The ST5 System . . . . .	65

4.3	Hybrid Sentences . . . . .	67
4.3.1	Conjunctions, Disjunctions and Implications in ST5 . . . . .	67
4.3.1.1	An Example . . . . .	67
4.3.1.2	Left-Right Asymmetries . . . . .	68
4.3.2	Deriving Presuppositions From Words . . . . .	71
4.3.2.1	Presuppositional Expressions in Traditional Truth-Conditional Semantics . . . . .	72
4.3.2.2	Presuppositional Expressions in ST5 . . . . .	74
4.4	Conclusions . . . . .	82
<b>5</b>	<b>Three Experiments</b>	<b>83</b>
5.1	First Experiment: Testing ST5 . . . . .	84
5.1.1	Design . . . . .	85
5.1.1.1	Predictions . . . . .	88
5.1.2	Methods . . . . .	91
5.1.2.1	Materials . . . . .	91
5.1.2.2	Procedure and Participants . . . . .	91
5.1.3	Results . . . . .	92
5.1.4	Discussion . . . . .	92
5.1.4.1	Vagueness . . . . .	92
5.1.4.2	Presupposition . . . . .	94
5.2	Second Experiment: A 3-Response Paradigm . . . . .	99
5.2.1	Design . . . . .	99
5.2.1.1	Predictions . . . . .	100
5.2.2	Methods . . . . .	102
5.2.2.1	Materials . . . . .	102
5.2.2.2	Procedure and Participants . . . . .	102
5.2.3	Results . . . . .	103
5.2.4	Discussion . . . . .	105
5.3	Follow-Up Experiment: Investigating Participants' Motivations for Presupposition . . . . .	107
5.3.1	Design . . . . .	107
5.3.1.1	Predictions . . . . .	107
5.3.2	Methods . . . . .	108
5.3.2.1	Materials . . . . .	108
5.3.2.2	Procedure and Participants . . . . .	108

5.3.3	Results . . . . .	109
5.3.4	Discussion . . . . .	111
5.4	Conclusions . . . . .	112
5.4.1	Experimental Considerations . . . . .	112
5.4.2	Theoretical Considerations . . . . .	113
<b>6</b>	<b>4-Valued Systems for Vagueness and Presupposition</b>	<b>115</b>
6.1	A Bi-Lattice . . . . .	116
6.1.1	$\mathcal{P}$ as an Absorbing Truth-Value . . . . .	119
6.2	The Repair Strategy . . . . .	121
6.2.1	Weak Kleene and Strong Kleene . . . . .	121
6.2.2	Merging Weak and Strong Kleene . . . . .	122
6.2.3	Merging Strong and Middle Kleene . . . . .	124
6.3	Conclusions and Comparison with ST5 . . . . .	129
<b>7</b>	<b>Antonyms</b>	<b>132</b>
7.1	Borderline Contradictions . . . . .	132
7.2	Two Conceptions of Antonyms . . . . .	134
7.3	Pilot Experiment . . . . .	136
7.3.1	Design . . . . .	137
7.3.1.1	Predictions . . . . .	138
7.3.2	Methods . . . . .	141
7.3.2.1	Possible Biases in the Adjectives' and their Antonyms' Features . . . . .	141
7.3.2.2	Materials . . . . .	144
7.3.2.3	Procedure . . . . .	145
7.3.3	Results . . . . .	146
7.3.4	Discussion . . . . .	146
7.4	Main Experiment . . . . .	150
7.4.1	Design . . . . .	150
7.4.1.1	Predictions . . . . .	150
7.4.2	Methods . . . . .	151
7.4.2.1	Materials . . . . .	151
7.4.2.2	Procedure . . . . .	152
7.4.3	Results . . . . .	152
7.4.4	Discussion . . . . .	153

<b>8</b>	<b>Conclusions</b>	<b>157</b>
8.1	Theoretical Conclusions . . . . .	157
8.2	Theoretical Perspectives . . . . .	159
8.3	Empirical Conclusions . . . . .	160
8.4	Experimental Perspectives . . . . .	162
<b>A</b>	<b>Experiments on Vagueness and Presupposition</b>	<b>164</b>
A.1	Experiment 1 . . . . .	164
A.1.1	Instructions . . . . .	164
A.1.2	Materials . . . . .	165
A.1.3	Practice Item . . . . .	166
A.2	Experiment 2 . . . . .	166
A.2.1	Instructions . . . . .	166
A.2.2	Materials . . . . .	167
A.2.3	Trial Items . . . . .	167
<b>B</b>	<b>Experiment on Antonyms</b>	<b>169</b>
B.1	Instructions . . . . .	169
B.2	Materials . . . . .	170
	<b>Bibliography</b>	<b>171</b>

# Acknowledgments

I want to thank my supervisors, Paul Egré and Orin Percus, who spent so much time and effort supporting me prior to and during the last three years. They knew how to bring out the best in me. Their rigorous advice and their kind patience have been priceless.

I am very grateful to Reinhard Muskens and Florian Schwarz for accepting to be my thesis reviewers. I am also grateful to the other members of my committee, Márta Abrusán and Benjamin Spector. My thanks also go to the members of my mid-thesis committee, Denis Bonnay and Nicholas Asher, for their helpful feedback, and to Philippe Schlenker for valuable advice and discussion. My gratitude goes to Benjamin Spector again and to Emmanuel Chemla too, for their availability and all the enriching discussions that we had together.

This work has benefited from discussions with more people than I can name. Among these people are the participants of the ESSLLI 2012 Workshop on Trivalent Logic and Their Applications, the participants of the Paris-Munich Workshop, and the participants of the Oxford-Princeton-IJN Workshop. I am especially thankful to John Burgess for his very helpful reviews.

The Ph.D. students and the teachers that I met during my visiting quarter at UCLA contributed greatly to the early development of this thesis. I want to thank all of the members of the Linguistics Department for welcoming me and making my visit truly enjoyable. In particular, I owe a great deal to Heather Burnett for her personal and intellectual guidance, and to Dominique Sportiche for his help during these first months of my thesis.

I would never have pursued this dissertation project without the encouragement and training that I received from the amazing teachers that I had at the University of Nantes. In particular, I am indebted to Orin Percus, Hamida Demirdache, and Nicolas Guillot for introducing me to syntax and semantics. I would also like to thank my fellow graduate students; I am especially grateful to Nataša Knežević for

many stimulating discussions, and to Typhanie Prince for always providing moral support.

This thesis would not have been possible without the administrative and financial support of the Institut Jean Nicod – Institut d’Études Cognitives at the École Normale Supérieure. During my time at ENS, I also had the good fortune to meet many brilliant researchers from all over the world, including new friends whose presence and moral support have been essential to the completion of this thesis. My warmest thanks go to Bianca Cepollaro, Alexandre Cremers, Natasha Ivlieva, Jeremy Kuhn, Tristan Thommen, Lyn Tieu, and Erin Zaroukian for their precious feedback and advice, and for the invaluable moments of fun that we had together.

My final thanks go to two important friends, Jean Balavoine and Jean Demel, and to my family, for supporting me day after day.

---

This thesis has been granted by the Labex program “Frontiers in Cognition”. This work has been supported by ANR-10-LABX-0087 IEC and ANR-10-IDEX-0001-02 PSL; a part of this work was also supported by a ‘Euryi’ grant from the European Science Foundation (“Presupposition: A Formal Pragmatic Approach” to P. Schlenker). The ESF is not responsible for any claims made here.



# Chapter 1

## Introduction

The day I taught my first course of semantics, I presented a definition of *meaning* along the lines of (Heim & Kratzer 1998)’s, which I was presented with as an undergraduate student in linguistics: *to know what a sentence means is to know in what situations it is **true***. And very soon I showed that, as I also had come to realize five years before, this definition was unable to capture our intuitions about *presuppositional* sentences: these are sentences we perfectly understand, but that we are sometimes as reluctant to judge true as to judge false, even while possessing all potentially relevant information. But by the time I became an instructor, I had become well acquainted with another phenomenon that similarly threatens this truth-conditional definition of meaning: the phenomenon of *vagueness*. So I added the class of vague sentences to the discussion.

That both vague and presuppositional sentences threaten this fundamental definition shows the importance of their study for the domain of semantics. Under the supervision of Orin Percus, I therefore decided to approach the two phenomena jointly in my M.A. dissertation. By applying the tools developed for analyzing presupposition in truth-conditional semantics to the study of vagueness, I showed that it was possible to give a novel sensible account of the sorites paradox that has been puzzling philosophers since Eubulide first stated it more than 2000 years ago. This result illustrates how the *joint* study of two phenomena that were previously approached separately can bring new insights to long discussed problems.

This thesis aims at pursuing the joint investigation of the two phenomena, by focusing on the specific truth-value judgments that they trigger. In particular, theoretical literature of the last century rehabilitated the study of non-bivalent logical systems that were already prefigured during Antiquity and that have non-trivial con-

sequences for truth-conditional semantics.<sup>1</sup> In parallel, an experimental literature has been constantly growing since the beginning of the new century, collecting truth-value judgments of subjects on a variety of topics. The work presented here features both aspects: it investigates theoretical systems that jointly address issues raised by vagueness and presupposition, and it presents experimental methods that test the predictions of the systems in regard to truth-value judgments.

The next two sections of this chapter are devoted to the presentation of my objects of study, namely vagueness and presupposition; and the last section of this chapter exposes the motivations that underline my project of jointly approaching the two phenomena from a truth-functional perspective. Because the notions of truth-value judgments are at the core of the dissertation, I have to make clear what I mean by *bivalent* and *non-bivalent* truth-value judgments. When I say that a sentence triggers *bivalent truth-value judgments*, I mean that in any situation, a sufficiently informed and competent speaker would confidently judge the sentence either “True” or “False”. When I say that a sentence triggers *non-bivalent truth-value judgments*, I mean that there are situations where a competent speaker, even perfectly informed, would prefer to judge the sentence with a label different from “True” and “False”. In this chapter, I will remain agnostic as to what labels are actually preferred for each phenomenon, but the next chapters are mostly devoted to this question.

## 1.1 Introducing Vagueness

I said that vague sentences threaten the truth-conditional definition of meaning, for they trigger non-bivalent truth-value judgments. However this property is not a sufficient criterion to define the set of vague sentences, for presuppositional sentences trigger non-bivalent truth-value judgments too.

Each vague sentence in fact results from the use of a “vague *predicate*”, and the specificity of the truth-value judgments that obtain consequently results from the presence of this vague predicate. To facilitate the understanding of the reader, let me already give an instance of a vague predicate to keep in mind through the discussion: *young*. Every predicate can be seen as associated with a *positive extension* and a *negative extension*: entities either fall into its positive extension or they fall into its negative extension; and applying a predicate to an entity from its positive extension

---

<sup>1</sup>In particular, (Łukasiewicz 1922) appears to have had a great influence in bringing the topic of non-bivalence back in the discussion, by proposing a 3-valued system to analyze the problem of future contingents discussed by Aristotle.

results in a true proposition, whereas applying a predicate to an entity from its negative extension results in a false proposition. But one specificity of *vague* predicates is that we seem to lack a good criterion to precisely determine these extensions: there are some entities for which we just do not know whether they should belong to the positive or to the negative extension of a vague predicate, even when we know all relevant information about the entity. For instance, there are people, typically people whose age you will consider *average*, for which you do not know whether they should be described as *young* or not, even though you know their precise age and you are a competent speaker of English. These people constitute *borderline cases* for *tall*. (Sorensen 2013) begins his article “Vagueness” in the Stanford Encyclopedia of Philosophy with the following claim.

There is wide agreement that a term is vague to the extent that it has borderline cases. This makes the notion of a borderline case crucial in accounts of vagueness.

However, that an entity fails to clearly belong to the positive or to the negative extension of a predicate does not suffice to make this predicate vague and this entity a borderline case for the predicate. For instance, we can refuse to sort the color blue into either of the positive or negative extensions of *underage* on the grounds that the color blue is neither underage nor not underage, and argue that the question of determining whether the color blue is underage is just meaningless. Nonetheless, I want a notion of borderline cases that does not make the color blue a borderline case for *underage*: *underage* is *not* a vague predicate, it has *no* borderline case. As a matter of fact, the non-bivalent truth-value judgments that result from predicating *underage* of the color blue are qualitatively different from the non-bivalent truth-value judgments specific to vagueness. For this reason, I have to refine my definition of borderline cases.

I claim that the entities in the positive and negative extensions of a vague predicate can always be sorted along a scale. I believe that the proper definition of borderline cases should make reference to this scale: borderline cases are those entities that lie somewhere in the middle of the scale and that trigger non-bivalent truth-value judgments when described by the vague predicate. With this definition in mind, note that both the entities in the positive and negative extensions of *underage* and the entities in the positive and negative extensions of *young* can be sorted along the scale of age. However the color blue does not lie in the middle of this scale (indeed, it does not lie anywhere on the scale): for this reason, it cannot be said to be a borderline case

for *underage* (nor for *young* for the matter) and we will not conclude that *underage* is a vague predicate. On the contrary, if we use *young* to describe some persons that are located in the middle of the scale, we do get non-bivalent truth-value judgments: those persons are borderline cases for *young*, and their existence allows us to conclude that *young* is a vague predicate.

Recall that I said that we seem to lack a good criterion to precisely determine the positive and negative extensions of a vague predicate. In this respect, *underage* and *young* constitute a good minimal pair. Underage people can reasonably be said to be young and a certain proportion of non-underage persons can reasonably be said to be not young. However, whereas law provides us with a precise threshold to determine who is underage and who is not, there is no non-arbitrary way of drawing a clear cut-off point between young and not-young people. This property of vague predicates has a dramatic consequence: it yields sorites paradoxes. (1) exemplifies a sorites paradox built with the vague predicate *young*.

- (1) a. Any 10 year old person is to be considered young.
- b. If any 10 year old person is to be considered young, then any 10.5 year old person is to be considered young too.
- c. If any 10.5 year old person is to be considered young, then any 11 year old person is to be considered young too.
- d. ...
- e. If any 89.5 year old person is to be considered young,  
          then any 90 year old person is to be considered young too.
- f. Any 90 year old person is to be considered young.

The paradox works as following: (1-a) strikes us as true and (1-f) strikes us as false. However, (1-b) to (1-e) seem to be true: half a year is not felt a period long enough to make a young person not young. But accepting (1-a) along with (1-b) to (1-e) as true leads us to accept (1-f) as true. Yet we initially rejected (1-f) as false. Our intuitive judgments about the truth of each sentence in (1) and our reasoning on their basis yield contradictory judgments: there is a paradox.

From a purely descriptive perspective, what leads us to conclude the truth of (1-f) from the truth of (1-a) is the *iterated* application of the *inductive premises* (1-b) to (1-e). There is another version of the sorites paradox exemplified in (2), where the iterated inductive premises are replaced by a single *universal inductive premise*.

- (2) a. Any 10 year old person is to be considered young.

- b. Any person who is just slightly older than a person to be considered young is to be considered young too.
- c. Any 90 year old person is to be considered young.

The way in which people go through this version of the paradox seems very similar to the way in which people go through the version of the paradox with iterated premises in (1): in both cases arriving at the conclusion requires the consideration of a series of persons ranked by age, in which any two successive persons just slightly differ in age (this is called a *sorites series*). The set of inductive premises of the former version directly provides the sorites series, whereas it needs to be reconstructed from the universal premise in the latter. Even though the mental process responsible for the efficiency of the two versions may be similar, it is important to note that the forms of the premises are essentially different between the two versions. This has important consequences for authors who propose to “solve the sorites paradox”, as some of them treat each version differently.

Importantly, note that the paradox vanishes once we replace *young* by *underage* in both (1) and (2).

- (3)
  - a. Any 10 year old person is to be considered underage.
  - b. If any 10 year old person is to be considered underage, then any 10.5 year old person is to be considered underage too.
  - c. If any 10.5 year old person is to be considered underage, then any 11 year old person is to be considered underage too.
  - d. ...
  - e. If any 89.5 year old person is to be considered underage, then any 90 year old person is to be considered underage too.
  - f. Any 90 year old person is to be considered underage.
- (4)
  - a. Any 10 year old person is to be considered underage.
  - b. Any person who is just slightly older than a person to be considered underage is to be considered underage too.
  - c. Any 90 year old person is to be considered underage.

Both arguments are actually valid (*if* we accept the premises as true, we *must* accept the conclusion as true too), but they are *unsound*: there is a false premise in both cases. There is always a pair of persons in the sorites series that falsifies an inductive premise. For instance, if the law defines 21 as the threshold age in the state and for the legal problem under discussion, the iterated inductive premise *If any 20.5 year old*

*person is to be considered underage, then any 21 year old person is to be considered underage too* is false. Consequently, the falsity of this particular conditional makes the universal inductive premise false too.

In conclusion, we can say that what is specific to vague predicates is to have borderline cases, and that describing a borderline case with a vague predicate triggers non-bivalent truth-value judgments. This very vague nature of some predicates gives rise to sorites paradoxes. As paradoxes, they surely call for an explanation, but for the present purpose they are very welcome, in that they constitute a helpful diagnostic in determining my first object of study. In the end, vagueness is this property of words which associates them with borderline cases and which is responsible for the existence of sorites paradoxes; and the non-bivalent truth-value judgments associated with vagueness are those truth-value judgments that we observe when we describe a borderline case with a vague predicate.

## 1.2 Introducing Presupposition

I said that presuppositional sentences threaten the truth-conditional definition of meaning, for they trigger non-bivalent truth-value judgments. However this property is not a sufficient criterion to define the set of presuppositional sentences: as we just saw, vague sentences too trigger non-bivalent truth-value judgments.

The situations in which we observe non-bivalent truth-value judgments for presuppositional sentences are called situations of presupposition *failure*. I could therefore consider defining presupposition as a linguistic process than can *fail*, and presuppositional sentences as those sentences that trigger non-bivalent truth-value judgments resulting from a *failure*. And as a matter of fact, several types of failures have been analyzed in terms of presupposition: reference failure for proper names and definite description as in (5-a) and (5-b) (van Fraassen 1968), categorization failure for sortal predicates as in (5-c) (Thomason 1972) or more generally truth-value failure for sentences (Strawson 1950).

- (5)    a.    Pegasus has a white hind leg.
- b.    The king of France is bald.
- c.    The color of copper is forgetful.

The reasons why we are reluctant to give a “True” judgment but also to give a “False” judgment for (5-a) and for (5-b) are because Pegasus does not actually exist and because there is no king of France. The expressions *Pegasus* and *the king of*

*France* can be said to *fail* to refer to individuals. The reason why we are reluctant to give a “True” judgment but also to give a “False” judgment for (5-c) is because colors are not conscious entities. The predicate *forgetful* can be said to *fail* to apply to the color of copper. A definition of presuppositional sentences could therefore consist in drawing a list of different types of failures, associated with the linguistic material that can fail. (Beaver & Geurts 2013) outline such a list but also note the “ubiquity of presupposition”. This note reveals the vanity of the task.<sup>2</sup>

There is in fact a better way of determining whether a sentence is presuppositional or not. Rather than using the intuitive but informal notion of *failure* to sort out the set of presuppositional sentences, one can decide to label a sentence as *presuppositional* depending on whether particular inferences can be drawn when it interacts with a series of linguistic operators. In particular, one can look at the inferences that obtain when (5-a), (5-b) and (5-c) are questioned, as in (6-a), (6-b) and (6-c).

- (6) a. Does Pegasus have a white hind leg?
- b. Is the king of France bald?
- c. Is the color of copper forgetful?

- (7) Do you have a son?

Note first that no special inference can be drawn about the state of mind of a speaker who questions a non-presuppositional sentence. For instance, the only thing that we can reasonably infer upon hearing a speaker use (7) is that this speaker is ignorant about her interlocutor having sons,<sup>3</sup> but we cannot in particular infer that the speaker believes her interlocutor to have no more than one son.<sup>4</sup> Crucially, things are different with presuppositional sentences: a *positive* inference can be drawn about the state of mind of a speaker who questions a presuppositional sentence. For instance, upon hearing a speaker use (6-a), we can reasonably infer that she believes in the existence of an entity named Pegasus; upon hearing a speaker use (6-b), we can reasonably infer that she believes France to be a monarchy; and upon hearing a speaker use (6-c), we can reasonably infer that she believes that colors can be conscious. The observation of these inferences is a phenomenon called *presupposition projection*, and it happens with many other linguistic operators, such as negation.

---

<sup>2</sup>This is not to say that a *typology* of presuppositional expressions is vain though.

<sup>3</sup>And maybe also that the speaker believes that her interlocutor *could* have a son.

<sup>4</sup>Note that we could infer this from its affirmative counterpart (i) though.

- (i) You have a son.

- (8) a. Pegasus does not have a white hind leg.
- b. The king of France is not bald.
- c. The color of copper is not forgetful.

Once again, upon hearing a speaker use (8-a), we can reasonably infer that she believes in the existence of an entity named Pegasus; upon hearing a speaker use (8-b), we can reasonably infer that she believes France to be a monarchy; and upon hearing a speaker use (8-c), we can reasonably infer that she believes that colors can be conscious.<sup>5</sup>

I will henceforth call *presuppositional* any sentence that exhibits this behavior. In addition, I will call the proposition whose we infer the belief by the speaker the *presupposition* associated with the presuppositional sentence. Accordingly, I will say that (8-a) *presupposes* that there is an entity named Pegasus, that (8-b) *presupposes* that France is a monarchy and that (8-c) *presupposes* that colors can be conscious. And these presuppositions are said to *project* through interrogation and through negation in (6) and in (8). Finally, I will say that a presupposition is not *fulfilled* whenever the context is incompatible with the truth of the presupposition, and that there is a *presupposition failure* when this results in the infelicity of the sentence.

Note that in this discussion I said that the kind of entities that presuppose are *sentences*, and I accordingly associate presuppositions with *sentences*. But as it appeared when I introduced presupposition projection, the inferences that we test to determine whether a sentence is presuppositional concern *speakers' beliefs* in specific situations. For this reason, some authors such as (Stalnaker 1974) or more recently (Schlenker 2008) prefer to define *pragmatic* presuppositions as conditions on some aspects of the conversational context that have to be met for an utterance to be felicitous. Under this latter view, the kind of entities that presuppose are *speakers*, and accordingly presuppositions are associated with *speakers* too.

As Stalnaker notes, the two views are not incompatible, but they are not equivalent. To illustrate the idea that the two views are not incompatible, consider a sentence like (5-b), repeated in (9), which in (Strawson 1950)'s semantic approach to presupposition fails to get a truth-value, given that France is not a monarchy.

- (9) The king of France is bald.

---

<sup>5</sup>It appears that (8-a), (8-b) and (8-c) could also be used to respectively convey that Pegasus is not real, that France is not a Monarchy and that colors are unconscious entities. On this point, see the discussion of *local accommodation* in the next section.



Strawson’s approach of presupposition is semantic to the extent that a presupposition failure corresponds to a *sentence* lacking a *truth-value*. But Stalnaker shows that this can easily be incorporated within a pragmatic view of presupposition when one introduces a bridge principle saying that *speakers* presuppose that the sentences uttered in the conversation do not lack a truth-value. However, as an illustration of the non-equivalence of the semantic and the pragmatic positions, note that according to the latter speakers can presuppose a variety of things, thus making the utterance of a sentence possibly infelicitous *even if this sentence gets a truth-value*.

These last considerations call for some important technical distinctions. I just said that the utterance of a sentence might feel infelicitous even if the sentence has a truth-value. Now, there may be situations where the use of a *true* or a *false* sentence is infelicitous and where, for this reason, a speaker would nonetheless be reluctant to judge the sentence as true or as false. By this remark, I want to emphasize the distinction between the semantic *truth-value* of a sentence and the *truth-value judgments* that speakers give for a sentence. Importantly, there is no necessary correspondence between a technical truth-value and an observed truth-value judgment. In addition, note that the definition of “presupposition” that I adopted in the end is stated in terms of inferences (and more technically in terms of presupposition *projection*), but not in terms of truth-value judgments. For this reason, I leave open the possibility that speakers could give bivalent truth-value judgments in certain situations which nonetheless correspond to a presupposition failure. On these distinctions, I refer the reader to (von Fintel 2004).

In conclusion, we can say that what is specific to presuppositional sentences is precisely to be associated with presuppositions, and that using these sentences when the presuppositions are not fulfilled typically triggers non-bivalent truth-value judgments. I defined presuppositions as those *propositions* that project through negation and interrogation in particular. In the end, the non-bivalent truth-value judgments associated with presupposition are those truth-value judgments that we observe in most contexts which are incompatible with the truth of the presuppositions.

## 1.3 The Project of Unifying Vagueness and Presupposition

### 1.3.1 Previous M.A. Work

As noted earlier, I pursued a unified approach of vagueness and presupposition in my M.A. thesis. Given that vague predications of borderline cases neither yield a clear “True” judgment nor a clear “False” judgment, Orin Percus suggested to me that I might model vague predicates with partial functions, which would exclude borderline cases from their domain of definition. His suggestion thus echoed an approach to presupposition inspired by (Heim & Kratzer 1998), which models presuppositional sentences with partial functions that exclude from their domain of definition any context where the presupposition is not true.<sup>6</sup> This position allowed me to view vague sentences as sentences that presuppose they contain no predication of a borderline case. For instance, under this view the vague sentence (10-a) can be said to be associated with the presupposition (10-b): if (10-b) fails to be true (i.e. if I *am* a borderline case for *tall*), then I am not in the domain of definition of the semantic value suggested for *tall* and this supposedly prevents the observation of a clear “True” judgment as well as the observation of a clear “False” judgment.

- (10)    a.    I am tall.  
          b.    I am not a borderline case for *tall*.

I then considered several problems raised by vagueness that are much discussed in the literature. One of these problems was to give an account for the sorites paradox. As a solution, I proposed a new explanation of how the paradox emerges and how it can be solved, a solution in line with the parallel between presupposition and vagueness illustrated above. This solution exploits the notion of projection *accommodation*, and in particular of *intermediate* and *local* accommodation. As an example of intermediate accommodation, first consider (11-a), which is associated with the presupposition that I have a car, expressed in (11-b). To this extent, if I utter (11-a), I am committed in accepting (11-b).

- (11)    a.    I wash my car on Saturdays.

---

<sup>6</sup>This is in fact a speculation from their use of partial functions to model the definite article *the*, which is traditionally associated with the presupposition that there exists a unique entity corresponding to the complement noun phrase. In addition, I here use “context” in an informal way, but the intensional semantics that Heim & Kratzer eventually propose models sentences with functions from *possible-worlds* and *assignment functions* to truth-values.

- b. I have a car.

That (11-a) is associated with the presupposition expressed in (11-b) is manifestly due to the presence of a possessive pronoun (*my*) in (11-a). Now, notice that contrary to a possible expectation, I am not necessarily committed in accepting (12-b) if I utter (12-a), even though it contains a possessive pronoun (*her*). Rather, (12-c) seems to be a legitimate paraphrase of (12-a).

- (12) a. Every citizen washes her car on Saturdays.
- b. Every citizen has a car.
- c. Every citizen *who has a car* washes it on Saturdays.

Under the interpretation of (12-a) which makes (12-c) a good paraphrase, we say that the presupposition has been *intermediately accommodated*: while the possessive pronoun appears *in the verbal phrase*, the interpretation of the presupposition is processed *outside of the verbal phrase* and has the effect of restricting the domain of individuals considered in the quantification. I proposed that the same process can take place with the universal inductive premise of the sorites paradox, whose (13-a) is an instance. Recall that this approach views vague sentences as associated with the presupposition that they contain no predication of a borderline case. As a consequence, (13-b) paraphrases (13-a) under an interpretation resulting from *intermediate accommodation*.

- (13) a. Every man who is slightly shorter than a tall man is tall too.
- b. Every man who is slightly shorter than a tall man *and who is not a borderline case for tall* is tall too.

A very nice consequence of such a reading of the universal inductive premise is that it makes it true without making the sorites argument valid. Indeed, imagine that when going from the top to the bottom of the scale of heights, you first have clearly tall men, then *borderline cases* and finally clearly not tall men. Then it is true that every man who lies just below a tall man on this scale and *who is not a borderline case for tall* lies among the clearly tall men too. As a consequence, (13-a) is true under the reading expressed in (13-b). Nonetheless, it is manifest that accepting the universal inductive premise under this reading does not lead one to conclude that men at the bottom of the scale are tall. Hence, the reading that results from *intermediate accommodation* makes the universal inductive premise true, but it also makes the

sorites argument invalid: there is no contradiction in accepting its premises under this reading and rejecting its conclusion.

Intermediate accommodation allowed me to account for why we feel the inductive premise of the sorites paradox to be true, but it did not allow me to account for why we feel the argument to be valid. I proposed that this feeling is due to the existence of another reading of the universal inductive premise, which results from *local* accommodation. To illustrate local accommodation, first consider (14-a), which is associated with the presupposition that I have a gun, expressed in (14-b). To this extent, if I utter (14-a), I am committed in accepting (14-b).

- (14)    a.    I keep my gun in a drawer.  
           b.    I have a gun.

Once again, that (14-a) is associated with the presupposition expressed in (14-b) is manifestly due to the presence of a possessive pronoun (*my*) in (14-a). Now, notice that (15-b) is a natural reading of (15-a).

- (15)    a.    Every citizen keeps her gun in a drawer.  
           b.    Every citizen *has a gun and* keeps it in a drawer.

Under the reading of (15-a) paraphrased in (15-b), we say that the presupposition has been *locally accommodated*: as the possessive pronoun appears *in the verbal phrase*, the presupposition is processed as *coordinated with the verbal phrase*. Once again, we proposed that the same process can take place with the universal inductive premise of the sorites paradox. As a consequence, (16-b) paraphrases (16-a) under an interpretation resulting from a *local* accommodation.

- (16)    a.    Every man who is slightly shorter than a tall man is tall too.  
           b.    Every man who is slightly shorter than a tall man *is not a borderline case for tall and* is tall too.

Remember the scale of heights that we considered earlier: you first have clearly tall men, then *borderline cases* and finally clearly not tall men. This means that some borderline tall men lie just below some clearly tall men. As a consequence, (16-b) is false: some men are slightly shorter than a tall man but are borderline tall.<sup>7</sup> Note however that if we were to accept (16-b) as true, then we would have no choice but to

---

<sup>7</sup>Of course, this position rests on the arguable hypothesis that we can draw a sharp division between clear cases and borderline cases for *tall*. This is a view that I readily endorsed, given that it made it very natural to model vague predicates with partial functions, which can be thought of as defining a *third* sharp extension in addition to the positive and the negative ones.

consider every man on the scale of heights tall.<sup>8</sup> Hence, the reading that results from *local* accommodation makes the universal inductive premise false, but it also makes the sorites argument valid: if we were to accept the premises, we would consistently have to accept the conclusion too.

The final picture of the approach of the sorites paradox proposed in my M.A. thesis is the following:

- vague expressions are presuppositional expressions
- as such, they are subject to the same processes as any presuppositional expression
- the sorites paradox results from an underlying ambiguity of the universal inductive premise:
  - we accept the premise as *true* under the reading resulting from *intermediate* accommodation, which in fact makes the sorites argument invalid
  - but we do accept the sorites argument as *valid* under the reading of the premise resulting from *local* accommodation, which in fact makes the premise false

Note that if the universal premise is ambiguous between a *true* reading which makes the sorites paradox invalid and a *false* reading which makes the sorites paradox valid, we should also expect to observe “False” judgments for the premise and rejection of the sorites argument. And this is actually what happens. In fact, the judgments of speakers are not radical, and some of them do reject the premise as false while some do reject the validity of the sorites argument.

### 1.3.2 Present Work

This successful result bolstered me in the pursuit of a unified approach to vagueness and presupposition. As mentioned earlier, the starting point of my unifying project was the observation of a common departure from the clear “True” and the clear “False” truth-value judgments that we otherwise observe for non-vague, non-presuppositional sentences. However, in treating vague sentences as merely a variety of presuppositional sentence, the approach outlined above fails to account for the specificities of each phenomenon. In particular, it fails to account for the specific

---

<sup>8</sup>Another possibility would be to deny the first premise of the sorites argument and consider that every man is in fact not tall.

truth-value judgments triggered by vague sentences as opposed to the specific truth-value judgments triggered by presuppositional sentences. This observation straightforwardly calls for a system that would incorporate vagueness and presupposition as two different phenomena, but that would nonetheless give an explanatory account for the fact that both phenomenon involve non-bivalent truth-value judgments.

As highlighted above, truth-conditional approaches to meanings aim to establish how we determine whether sentences are *true* or *false*. From this perspective, one could propose to view the interpretation of a sentence as running an algorithm which outputs a truth-value corresponding to *true* or a truth-value corresponding to *false*. From there, a reasonable expectation is that we should always either observe a “True” judgment or a “False” judgment when a speaker interprets a sentence, depending on which truth-value her algorithm outputs. This observation is not borne out though, in particular when vague and presuppositional sentences are interpreted in specific contexts. I wanted to investigate the properties of these sentences which would be responsible for the interpretation process failing to eventually produce bivalent truth-value judgments. In particular, I wondered whether it was possible to posit a single algorithm to treat vague sentences and presuppositional sentences, or if one should prefer to define specific algorithms dedicated to each phenomenon. These considerations strike us as particularly relevant once we realize that vagueness and presupposition are parts of one global linguistic system, and that the two phenomena enter in interaction in many linguistic constructions like those listed in (17), where each sentence contains at least one vague and one presuppositional expression. I refer to these kinds of sentences as *hybrid sentences*.

- (17) a. The king of France<sub>presuppositional</sub> is bald<sub>vague</sub>.  
 b. You are rich<sub>vague</sub> but you don’t live in your<sub>presuppositional</sub> mansion.  
 c. I discovered<sub>presuppositional</sub> that you are old<sub>vague</sub>.

In this thesis, I try to answer these questions from a truth-functional perspective. This means that I will focus on what truth-values we should associate with vague and presuppositional sentences in different contexts, and how these truth-values can be used to derive the specific non-bivalent truth-value judgments that we observe.

Chapter 2 constitutes a synthetic review of the trivalent systems that have been proposed in the literature of vagueness and in the literature of presupposition. It shows that two main systems, supervaluationism and Strong Kleene, have been entertained in addressing vagueness as well as in addressing presupposition. Depending

on which phenomenon they propose to model, the different authors give different interpretations of the same system. Surprisingly enough, despite these common uses of one system across the phenomena, no author seems to have seriously considered a unified approach of vagueness and presupposition.

I give a succinct look at the extant experimental literature in Chapter 3. Even though this literature is already vast on both topics and is still growing, I focus on a few experiments that provide direct evidence for non-bivalent truth-value judgments. Not only do these experiments provide good evidence for speakers accessing non-bivalent truth-value judgments when evaluating vague descriptions of borderline cases and sentences with unfulfilled presuppositions, they also suggest that vague sentences allow for both *glutty* and *gappy* judgments, i.e. respectively truth-value judgments of the form *both true and false* and truth-value judgments of the form *neither true nor false*, whereas presuppositional sentences do give rise to the latter but never give rise to glutty judgments.

Chapter 4 can be seen as addressing the possibility of positing a single algorithm to treat vague sentences and presuppositional sentences. It introduces a totally ordered 5-valued logical system named *ST5*, that I developed at the beginning of my thesis on the basis of a system proposed in (Percus & Zehr 2012). *ST5* draws on a trivalent system developed for vagueness by (Cobreros, Egré, Ripley & van Rooij 2012) and it aims at unifying vagueness and presupposition by positing 5 *totally ordered* logical truth-values. This system naturally derives non-bivalent truth-value judgments for vague and presuppositional sentences, while associating each type with specific truth-values. Importantly, *ST5* makes direct predictions regarding the interaction between vagueness and presupposition in hybrid sentences like those in (17).

Chapter 5 presents two experiments that I subsequently conducted to test the predictions of *ST5* and more generally to empirically investigate the non-bivalent truth-value judgments associated with each phenomenon. The results do not conform to the predictions of *ST5*. In particular, in order to derive glutty judgments for vague descriptions of borderline cases, *ST5* appears to necessarily predict “True” judgments for negative sentences with unfulfilled presuppositions. However the truth-value judgments observed in the first experiment do not seem to go in this direction, and the truth-value judgments observed in the second experiment argue against this prediction. Chapter 5 also presents a follow-up experiment designed to elicit some problematic data for the presuppositional sentences tested in the second experiment. Unfortunately this follow-up did not produce conclusive results. I end this chapter with a practical discussion of the problems I encountered when experimentally probing

truth-value judgments for presupposition and with a theoretical discussion pointing in the direction of a partially ordered 4-valued system.

This is what I investigate in Chapter 6. This chapter can be seen as a counterpoint of Chapter 4, to the extent that I consider specific algorithms dedicated to each phenomenon. I first present a bi-lattice where vagueness and presupposition are conceived as entering in relation with plain truth and plain falsity on different dimensions. I then define a semantics for the logical operators, which relies on a total order that I derive from the bi-lattice. This makes direct predictions for the hybrid sentences that connect a vague and a presuppositional sentence. I first show that the truth-tables that obtain treat vagueness in a Strong Kleene way, whereas they treat presupposition in a Weak Kleene way. I then show that these truth-tables can in fact be derived from a joint implementation of an algorithm defining the Strong Kleene truth-tables and of an algorithm defining the Weak Kleene truth-tables. Finally, I discuss the possibility of addressing presupposition with a Middle Kleene algorithm in order to account for linearity effects, and how this option leads to make a stipulative choice when merging the Middle Kleene algorithm with the Strong Kleene algorithm.

Chapter 7 returns to vagueness by questioning the reality of gluttony judgments. Paul Egré and I investigate the acceptance of contradictory descriptions built with antonym adjectival phrases. We experimentally test two pragmatic theories of antonyms that derive similar *uses* for the adjectives, but which rests on different *underlying semantic* assumptions. The results that we obtain reveal that speakers accept in particular to describe borderline tall persons as “X is tall and not tall” but not as “X is tall and short”. We argue that these results are evidence for a view of lexical antonyms (*tall* vs. *short*) as semantic contraries rather than semantic contradictories.

Finally, I conclude this dissertation with Chapter 8, where I discuss the problems and possible solutions to a unified truth-functional approach of vagueness and presupposition that were put forward in this thesis.



## Chapter 2

# Trivalence, Vagueness and Presupposition

A reasonable reaction to the observation of non-bivalent truth-value judgments is to revise the bivalent logical system we use to model our truth-judgments. Maybe the most natural way to do it is to come up with a trivalent logical system, where the third value (noted #) is meant to be associated with non-bivalent truth-judgments. But this extension is not straightforward and many different implementations of the third value are theoretically possible. Among these possible implementations, two systems have received great attention since the middle of the twentieth century: Kleene's strong logic (henceforth Strong Kleene or SK) and supervaluationism. As will appear clearly in this presentation, these two systems hinge on very similar intuitions, to the point that both SK's and supervaluationism's connectives can be derived from the same bivalent considerations (see for instance (Spector 2012)). Note I use "Strong Kleene" and "SK" without committing myself to any particular view on how the system defines validity, in particular I remain agnostic on whether the third value # belongs to the set of designated values.

Both Strong Kleene and supervaluationism were used to model the phenomena of vagueness as well as the phenomena of presupposition. However, no author seems to have addressed the question of whether these systems could correctly model a language containing both vague and presuppositional expressions, even though natural languages are such languages. The two sections in this chapter are devoted to these two trivalent systems. Section 2.1 presents supervaluationism; Section 2.2 presents Strong Kleene. Each section discusses the various ways in which the trivalent systems were used and extended to account for the phenomena of vagueness and presupposition.

## 2.1 Supervaluationism

The system known as supervaluationism was first formalized by Baas van Fraassen in (van Fraassen 1966) to account for truth-value gaps in propositions involving non-referring terms such as “Pegasus”, and he later extended it to the analysis of presupposition (van Fraassen 1968). (Fine 1975) and (Kamp 1975) then adapted his system to vagueness ten years later.<sup>1</sup>

The truth-tables that classical logic provides for operators are bivalent. Therefore, if a proposition of value  $\#$  is part of a complex proposition, we cannot simply read the classical truth-tables to determine the truth-value of the complex proposition. The idea of supervaluationism is precisely to provide a general mechanism to determine this value on the basis of the classical truth-tables. This general mechanism makes use of the notion of *classical valuations* over a model. For the sake of the presentation, I will consider that the models over which we apply classical valuations must be models which associate atomic propositions with 1, 0 or  $\#$ , but other positions might be possible.<sup>2</sup> Regardless of one’s position towards the status of  $\#$ , the important point here is that supervaluationism builds on models which allow us to divide the set of atomic propositions in two categories: one category of atomic propositions that classical valuations map to a *systematically determined* value in  $\{0,1\}$ , and one category of atomic propositions that classical valuations can *freely* map to 0 or 1. Classical valuations are functions from propositions to  $\{0,1\}$ : they “repair” the model by mapping atomic propositions which initially receive the value  $\#$  to either 0 or 1. Besides, classical valuations necessarily map atomic propositions which initially receive a bivalent value to this bivalent value. In consequence, there is only one classical valuation over a model where each proposition already gets the value 1 or the value 0. However, there are two classical valuations over a model where exactly one atomic proposition initially receive the value  $\#$ . Logically, there are four

---

<sup>1</sup>Interestingly, (Cobreros, Egré, Ripley & van Rooij forthcominga) note that supervaluationism was already prefigured in (Mehlberg 1958)’s analysis of *vagueness*. As a matter of fact, supervaluationism has afterward been much more popular in the field of vagueness than in the field of presupposition, it seems quite reasonable to expect it to be initially developed for vagueness.

<sup>2</sup>For instance, in (van Fraassen 1966) the primary role of the model is to associate names to individuals and to associate predicates with a positive and a negative extension, but as the following excerpt shows (cf. *in the indicated manner*), he remains vague on whether the model associates formulas with truth-values independently of the classical valuation ( $v$ ):

if  $A$  is an atomic statement containing no nonreferring names, then  $v(A)$  is determined by the model, in the indicated manner

classical valuations over a model where exactly two atomic propositions initially receive the value #: one where they both receive 1, one where they both receive 0, and two where they get different values. Importantly, the truth-value of every complex proposition interpreted with a classical valuation will always be readable from the classical truth-tables, because every atomic proposition receive a bivalent truth-value when interpreted with a classical valuation.

The main idea of supervaluationism is to add another type of valuations, the *supervaluations*, which evaluate each proposition depending on the value it receives in the set of all possible classical valuations, as defined in Definition 2.1.1.

**Definition 2.1.1** (Supervaluation). *A supervaluation  $s$  over a model  $\mathcal{M}$  is a function from propositions to  $\{0, \#, 1\}$  such that, for any proposition  $\phi$ :*

- $s(\phi) = 1$  if and only if all the classical valuations over  $\mathcal{M}$  map  $\phi$  to 1 ( $\phi$  is supertrue in the model)
- $s(\phi) = 0$  if and only if all the classical valuations over  $\mathcal{M}$  map  $\phi$  to 0 ( $\phi$  is superfalse in the model)
- $s(\phi) = \#$  otherwise

Table 2.1 lists the set of the classical valuations over the given model, where  $\phi$  and  $\psi$  are assumed to be the only atomic propositions of value #.

	$\phi$	$\psi$	$\delta$	$\phi \wedge \psi$	$(\phi \wedge \psi) \rightarrow \delta$	$\phi \vee \neg\phi$	$\phi \wedge \neg\phi$
$\mathcal{M}$	#	#	1	?	?	?	?
$v_1$	0	0	1	0	1	1	0
$v_2$	0	1	1	0	1	1	0
$v_3$	1	0	1	0	1	1	0
$v_4$	1	1	1	1	1	1	0
$s$	#	#	1	#	1	1	0

Table 2.1: Example of a supervaluationist computation

As Table 2.1 makes clear, each single classical valuations assigns the same truth-value to every occurrence of the same atomic proposition in a complex sentence. I will say that supervaluationism *preserves the identity* of the (sub)propositions. A main direct consequence of this property is that supervaluationism preserves the laws of the excluded middle and of non-contradiction. To see this, look at the two last columns of Table 2.1, which show particular instances of a tautology and of a contradiction. They respectively get the truth-values 1 and 0 in every classical valuation and are

therefore respectively supertrue and superfalse. Had  $\phi$  and  $\psi$  received a bivalent truth-value in the initial model, this disjunction and this conjunction would still have been respectively supertrue and superfalse, because classical valuations respect the laws of classical logic. In contrast, note that even though  $(\phi \wedge \psi) \rightarrow \delta$  is supertrue in this model, it is not a tautology. Indeed, had the model under consideration assigned a different value to  $\delta$ , say 0, this complex proposition would not have been supertrue anymore.

Supervaluationism has the important property of *deriving* its connectives on purely *bivalent* considerations, whereas SK is usually seen as stipulating a set of trivalent truth-tables for its connectives. To this extent, one can consider that supervaluationism is a better motivated system than Strong Kleene. However, as discussed in Sect. 2.2.1, it turns out that SK too can be built on systematic bivalent considerations.

### 2.1.1 Presupposition: Van Fraassen's Proposal

As noted earlier, supervaluationism was first proposed by (van Fraassen 1966) to deal with arguments involving non-referring names, like (18) where the name “Mortimer” is assumed not to refer.

- (18)    a.    Mortimer is a man.
- b.    If Mortimer is a man, then Mortimer is mortal.
- c.    Mortimer is mortal.

Van Fraassen proposes a treatment of non-referring terms and a notion of validity that make this argument valid even if we endorse a view *à la* Strawson where atomic statements containing non-referring names are neither true nor false, and are therefore to be modeled with propositions of value  $\#$ . Basing again on Strawson's position toward failure of definite descriptions, (van Fraassen 1968) proposes the following characterization of presupposition:

- (19)     $A$  necessitates  $B$  if and only if, whenever  $A$  is true,  $B$  is also true.<sup>3</sup>
- (20)     $A$  presupposes  $B$  if and only if
  - a.     $A$  necessitates  $B$
  - b.    *not* –  $A$  necessitates  $B$

---

<sup>3</sup>As van Fraassen notes, this is equivalent to saying that  $A$  (semantically) entails  $B$ . I will freely use *necessitate* to express the relation between statements and the relation between propositions.

As it appears in (20), that we infer the truth of the presupposition of a statement  $A$  is *part of the definition* that van Fraassen gives for the notion of presupposition (at least under the assumption that we infer entailed statements). In addition, given this definition, for any statement  $A$  presupposing  $B$ , we are to infer  $B$  from  $\text{not} - A$ . To this extent, van Fraassen views the projection of presuppositions over negation as primitive: no special mechanism is responsible for the inference of the presupposition from the negative counterpart of a presuppositional statement.

We can show that both an atomic statement modeled as  $Pa$ , containing a name  $a$ , and its negative counterpart modeled as  $\neg Pa$ , necessitate that the name  $a$  have a reference: by definition the proposition  $Pa$  is *neither true nor false* in models where  $a$  does not refer, and neither is  $\neg Pa$ . Consequently, for  $Pa$  to be true in a model,  $a$  has to refer in the model, and for  $\neg Pa$  to be true in a model  $a$  also has to refer in the model. This means that both the affirmative statement and its negative counterparts necessitate that the name refer. Therefore, it is legitimate to say that van Fraassen adopts Strawson's position in considering that atomic statements containing names presuppose these names have references.

Conversely, to posit that (21-a) presupposes that there is an integer between 2 and 3 is to say that (21-a) as well as its negative counterpart necessitate that there is an integer between 2 and 3. This means that, if we translate (21-a) as  $\phi$  and its presupposition that there is an integer between 2 and 3 as  $\psi$ , there is no model where either  $\phi$  or  $\neg\phi$  is true and  $\psi$  is not true. This is equivalent to saying that every model where  $\psi$  is not true is a model where neither  $\phi$  nor  $\neg\phi$  are true. And given the treatment of negation in supervaluationism, models where neither  $\phi$  nor  $\neg\phi$  are true are models where  $\phi$ , and consequently  $\neg\phi$ , are undefined. We can generalize these considerations and conclude that any statement whose presupposition is not fulfilled lacks a truth-value.

Any statement presupposes that we are not in a situation where it lacks a truth-value, and any statement whose presupposition is not fulfilled lacks a truth-value: therefore, a statement that lacks a truth-value is a statement whose presupposition is unfulfilled. It is no surprise then that van Fraassen treats the argument in (21) the same way as he treats the argument in (18).<sup>4</sup>

(21) a. The integer between 2 and 3 is even.

---

<sup>4</sup>Van Fraassen uses "The King of France (in 1967) is bald" as an example of a presuppositional sentence. I prefer to base my examples on (i-a) for two reasons: first there is no model where the definite description would refer (as long as predicates are interpreted in any sensible way), thus making the possibility of treating (i) as valid even more challenging; and second it does not involve the predicate "bald" which could be described as vague.

- b. If the integer between 2 and 3 is even,  
then the integer between 2 and 3 is not prime.
- c. The integer between 2 and 3 is not prime.

First note that while (21-a) and (21-c) are necessarily to be evaluated as atomic propositions of value  $\#$  (given there is no integer between 2 and 3), it is not clear whether (21-b) would be translated as a tautological proposition or not, i.e. as a proposition which would be supertrue in every model. On the one hand, supervaluationism seems to make (21-b) non tautological, for its antecedent and its consequent do not involve the *same* propositions, so the preservation-of-identity property is no help here. But given that supervaluationism nonetheless has to be implemented in a way to access the *identity* of propositions, it is little step to imagine an implementation which would access logical relations *between propositions* and which would relatedly restrict the set of classical valuations to consider in building a supervaluation over a model (the set of *admissible classical valuations*). Such an implementation could make (21-b) tautological, because anything which is even is not prime, given that by definition it can be divided by 2. Restricting the set of admissible classical valuations on the basis of the logical relations existing between predicates has for instance been proposed in (Fine 1975), where these restrictions obtain from what is called “external penumbral connections” (see below for a discussion of his proposal).

Regardless of the value which supervaluations assign to (21-b), (21) is still a *supervalid* argument.<sup>5</sup> Indeed, even though every model that we consider makes (21-a) and (21-c) undefined (assuming that no model makes the definite description actually refer to something in the domains of *even* and *prime*), every classical valuation which makes (21-a) and (21-b) true also makes (21-c) true, for (21-c) is the consequent in (21-b) which has (21-a) as an antecedent, which is itself true in the valuation under consideration. Van Fraassen motivates this result on the ground that validity can be defined as preservation of truth:<sup>6</sup>

An argument is valid if and only if, were its premises true, its conclusion would be true also.

At this point, one might wonder how van Fraassen’s system deals with the projection of presuppositions. We saw that the negative counterpart *not* – *A* of a statement

---

<sup>5</sup>It seems that van Fraassen never employed this term, which was used by Fine (Fine 1975). For van Fraassen, arguments of the form of (21), involving non-referring terms, are simply *valid*.

<sup>6</sup>Although he notes the existence and interest of an understanding of validity as preservation of non falsity.

$A$  which presupposes  $B$  necessitates  $B$ , and that it directly accounts for the fact that we infer  $B$  from *not*  $-A$ . But this is not to say that *not*  $-A$  *presupposes*  $B$ . However, given that any statement presupposes that we are not in a situation where it lacks a truth-value, and given that in supervaluationism negative counterparts lack a truth-value in the exact same situations as their positive counterparts, we conclude that for any  $A$  presupposing  $B$ , *not*  $-A$  also presupposes  $B$ . Van Fraassen says nothing about the projection of presuppositions in connectives (though we know that we can have tautologies and contradictions built on presuppositional propositions). As a matter of fact, the statement in (22) does *not* presuppose that there are wooden planets: we do not infer this information and even if one thinks there is no wooden planet, one does not feel, when confronted to (22), the squeamishness characteristic of presupposition failure. We want to see if van Fraassen's account is consistent with this observation.

(22) If there are wooden planets, then the wooden planets are flammable.

Let us suppose that (22) is translated as a proposition of the form  $\phi \rightarrow \psi$ . In the present framework, that (22) not be presuppositional would mean that no model makes  $\phi \rightarrow \psi$  undefined. In regard of the evaluation of (22), models can be divided in two categories: models where  $\phi$  gets 1 and models where  $\phi$  gets 0 (we assume that no model assigns  $\#$  to  $\phi$ ). Because  $\phi$  is considered to be the only presupposition of  $\psi$ , models where  $\phi$  gets 1 are subdivided in models where  $\psi$  gets 0 and models where  $\psi$  gets 1 and there is no model where  $\psi$  gets  $\#$ . Because neither  $\phi$  nor  $\psi$  gets  $\#$  in these models, neither does  $\phi \rightarrow \psi$ . In models where  $\phi$  would get 0 though,  $\psi$  would get  $\#$  and we would have to consider the possible classical valuations over these models. There are classical valuations where  $\psi$  is assigned 1 and classical valuations where  $\psi$  is assigned 0. In all these valuations,  $\phi \rightarrow \psi$  would be assigned 1 because the value of the antecedent,  $\phi$ , would be 0. Therefore, even if we imagine models where  $\phi$  would get 0, these models would make  $\phi \rightarrow \psi$  supertrue. In the end, there is no model where  $\phi \rightarrow \psi$  would get  $\#$ , which means that supervaluationism correctly predicts (22) not to be presuppositional. The same reasoning could be applied to (23-a) and (23-b) and we would see that they don't presuppose anything either.

- (23) a. There are wooden planets and the wooden planets are flammable.  
b. Either there is no wooden planet, or the wooden planets are flammable.

Importantly, things are different when the complex statement we consider do not connect a statement and its presupposition. Consider (24), again to be translated as  $\phi \rightarrow \psi$ , but where the antecedent is no longer the presupposition of the consequent.

(24) If there are handcrafted planets, then the wooden planets are flammable.

In models where both  $\phi$  and  $\psi$  receive a bivalent truth-value,  $\phi \rightarrow \psi$  receives a bivalent truth-value too. Now there are two types of models where  $\psi$  is undefined: models where  $\phi$  gets 0 and models where  $\phi$  gets 1 (we assume that no model makes  $\phi$  undefined). As we saw earlier, in models where  $\phi$  gets 0,  $\phi \rightarrow \psi$  is supertrue because the antecedent is false in all the classical valuations. But in models where  $\phi$  gets 1, you have classical valuations where  $\phi \rightarrow \psi$  is assigned 1 (classical valuations where  $\psi$  is assigned 1) but also classical valuations where  $\phi \rightarrow \psi$  is assigned 0 (classical valuations where  $\psi$  is assigned 0). In these models,  $\phi \rightarrow \psi$  is undefined. This means that (24) presupposes that we are not in a situation described by these latter models. Again, these models are models where  $\phi$  gets 1 and  $\psi$  gets  $\#$ . These models thus describe situations where there are handcrafted planets but where there is no wooden planet. (24) presupposes that we are not in these situations, therefore (24) presupposes that if there are handcrafted planet, then there are wooden planets. And this seems to fit rather well with our intuitions about (24). Again, the same reasoning applied to (25-a) and (25-b) would predict the same presuppositions.

- (25) a. There are handcrafted planets and the wooden planets are flammable.  
 b. Either there is no handcrafted planets or the wooden planets are flammable.

However this supervaluationist approach comes with no consideration of linearity, therefore it predicts (26) to presuppose exactly the same as (25-a) and this seems to conflict with our intuitions on (26), which seems to unconditionally presuppose that there are wooden planets.

(26) The wooden planets are flammable and there are handcrafted planets.

In addition, the supervaluationist approach predicts both (27-a) and (27-b) to yield conditional presuppositions. As was said earlier, our intuitions seem to confirm that (27-a) presupposes that there are wooden planets if there are handcrafted planets, but the system predicts (27-b) to presuppose that *if we can build giant flamethrowers*, then there are wooden planets, and this conflicts with our intuitions that it unconditionally presupposes that there are wooden planets.

- (27) a. If there are handcrafted planets, then the wooden planets are flammable.  
 b. If we can build giant flamethrowers, then the wooden planets are flammable.



The contrast that we observe between (27-a) and (27-b), and that supervaluationism fails to capture, is characteristic of what has been called the *Proviso Problem*. This problem concerns the fact that some complex sentences containing a presuppositional subsentence unconditionally inherit the presupposition of this subsentence while others come with a conditional presupposition built on that of the subsentence. Because, as we saw, supervaluationism derives conditional presuppositions, in this framework the *Proviso Problem* would be approached as a problem of *presupposition strengthening*. Some authors, such as Danny Fox (Fox 2012), explored this solution for other trivalent systems.

### 2.1.2 Vagueness: Fine's implementation

Van Fraassen proposed supervaluationism to account for what he felt to be valid arguments even though they involve non-referring or presuppositional terms. To this extent, the notion of classical valuations that one has to consider in building a supervaluation were mostly formal tools to stick to classical logic in determining whether a sentence is supertrue or superfalse in a model or whether an argument is or is not supervalid. In Fine's (Fine 1975) adaptation of supervaluationism to vagueness, these classical valuations gain some substantivity: they correspond to as many ways of making a vague predicate *more precise*. Under his approach, vague predicates divide their arguments in (at least) three categories: arguments of which the predicate clearly hold, arguments of which the negation of the predicate clearly holds, and borderlines cases, of which neither the predicate nor its negation clearly hold. The predication of a borderline case is therefore initially undefined in the model, because of the very underspecified nature of the predicate.<sup>7</sup> But vague predicates also have this property of sorting their arguments along some dimension, as this is a crucial property for triggering sorites paradoxes. So, if one decided to use vague predicates in a bivalent manner, that is to say if one made a decision for each borderline case to treat it as a clear instance or as a clear counterinstance of the vague predicate, one would have to decide where to place the threshold separating cases of which the predicate holds from cases of which the predicate does not hold. Of course one would have many ways to do it, but for the sake of consistency, one can already exclude all bivalent extensions of the vague predicate which go against the order associated with it. Considering such a bivalent use of a vague predicate is considering what Fine calls a *precisification* of the predicate. Precisifications are in fact classical valuations, where vague predicates now

---

<sup>7</sup>In this paper, Kit Fine endorses a view of vagueness as *underspecification*, which is perfectly consistent with his vision and use of supervaluationism.

sharply divide their arguments in *two* extensions, a positive and a negative one. But Fine adds a constraint on these precisifications, based on the scalar nature of vague predicates, which he qualifies as *internal penumbral connections* on vague predicates. This constraint has the effect that for any  $Blob_b$  redder than  $Blob_a$ , no precisification over a model where  $Blob_a$  and  $Blob_b$  are borderline cases of *red* should make  $Blob_a$  in the positive extension of *red* while making  $Blob_b$  in its negative extension.<sup>8</sup> As a consequence, (28) will be supertrue, to the extent that  $Blob_b$  is closer to clear instances of red than is  $Blob_a$ .

(28) If  $Blob_a$  is red, then  $Blob_b$  is red too.

Importantly however, things are very different for (29).

(29) If  $Blob_b$  is red, then  $Blob_a$  is red too.

Placing the line separating red objects from not-red objects between  $Blob_a$  and  $Blob_b$  is perfectly compatible with the constraint above *even if  $Blob_a$ 's and  $Blob_b$ 's colors are very similar*. All that the constraint says is that in a precisification where  $Blob_a$  and  $Blob_b$  are not in the same extension of *red*, it is borderline-red  $Blob_b$  that should be in the positive extension and borderline-red  $Blob_a$  should be in the negative extension. Of course there will also be precisifications over this model where both are red, but *is not the case in all the precisifications* and therefore (29) will not be supertrue (nor superfalse) in this model (recall that precisifications are classical valuations). This means that if one builds a sorites argument based on multiple iterations of inductive premises (of the form of (29)), some of these premises will in fact be neither supertrue nor superfalse, preventing one from successfully applying the argument. Regarding the universal premise in (30), things are even clearer.

(30) For all pairs of objects  $A$  and  $B$  whose colors vary insensibly, if  $B$  is red then  $A$  is red too.

(30) is not just neither true nor false, it is superfalse. To see this, just note that when you evaluate (30) by examining every possible precisifications, you consider classical valuations where *red* is necessarily bivalent. So there is a precise line dividing red objects from not-red objects, even though this line varies depending on which precisification you consider. You can compare it to the case of *underage* which is associated with a different threshold-age depending on the country and the state you

---

<sup>8</sup>This constraint is called the *monotonicity principle* in (Egré to appear), after (Nouwen 2011). Note that (Fine 1975) uses *monotonicity* to refer to a different notion.

consider, but which sharply divides the population in two categories anyway. As a matter of fact, (31) is clearly false, and so is (30) according to Fine.

- (31) For all pairs of persons  $A$  and  $B$  whose ages are close enough, if  $B$  is underage then  $A$  is underage too.

Fine’s supervaluationist approach to vagueness thus makes both versions of the sorites argument valid, but both are inapplicable: the iterated inductive premise version is inapplicable because instances of its premises concerning borderline cases, like (29), are neither supertrue nor superfalse; and the universal inductive premise version is inapplicable because this very premise is superfalse. The inapplicability that Fine predicts for the sorites arguments is welcome to the extent that it prevents people from accepting the obviously false conclusions, but one still has to explain why we feel the inductive premises so compelling whereas supervaluationism makes them either superfalse or neither true nor false, and why we reject (32) (which is equivalent to the negation of superfalse (30)) as false whereas it is supertrue:

- (32) There is a pair of objects  $A$  and  $B$  whose colors vary insensibly such that  $B$  is red while  $A$  is not.

## 2.2 Strong Kleene

Stephen C. Kleene first proposed what is now referred to as “Strong Kleene’s truth-tables” in a paper on partial recursive functions and the inability of Turing machines to decide the value of a class of propositions (Kleene 1938). The truth-tables he proposed are represented in Figure 2.2, where  $\#$  stands for the third (“undefined”) value.

$\phi$	$\psi$	$\neg\phi$	$\neg\psi$	$\phi \wedge \psi$	$\phi \vee \psi$	$\phi \rightarrow \psi$
0	0	1	1	0	0	1
0	#	1	#	0	#	1
0	1	1	0	0	1	1
#	0	#	1	0	#	#
#	#	#	#	#	#	#
#	1	#	0	#	1	1
1	0	0	1	0	1	0
1	#	0	#	#	1	#
1	1	0	0	1	1	1

Table 2.2: Strong Kleene’s truth-tables

An important property of this set of connectives, Kleene notes, is the fact that all equivalences of classical calculus of propositions hold.<sup>9</sup> In particular, they respect De Morgan laws ( $\phi \wedge \psi \equiv \neg(\neg\phi \vee \neg\psi)$  and  $\phi \vee \psi \equiv \neg(\neg\phi \wedge \neg\psi)$ ), thus letting us define any connective in terms of another. (Cobreros et al. forthcoming<sup>a</sup>) express one manner of deriving these truth-tables the following way:

On his account, a connective takes a classical value when all ways of completing the assignment of the undefined value by a classical value converge to the same value; in all other cases, the function stays undefined

SK is thus a motivated system (the truth-tables can be derived from purely bivalent considerations) which maintains the equivalences established in classical bivalent logic. This surely explains why its use is so widespread in the literature.

A recurring critic against this system is that even though it leaves the classical *equivalences* unchanged, there are classical *laws* which can't be satisfied with these definitions of connectives. To see this, consider the formulas  $A \vee \neg A$  and  $A \wedge \neg A$ . In classical logic, the former is a tautology for it gets the value 1 in every model (i.e. classical logic respects the law of the excluded middle); and the latter is a contradiction for there is no model in which it gets the value 1 (i.e. classical logic respects the law of non-contradiction). Now, with the truth-tables in Table 2.2, we have models where the former and models where the latter get the value  $\#$ . If one regards  $\#$  as a designated truth-value (this is the position defended in (Priest 2006) for instance), then the law of the excluded middle is still satisfied ( $\phi \vee \neg\phi$  gets a designated truth-value in *every model*) but the law of non-contradiction is violated (*there are* models where  $\phi \wedge \neg\phi$  gets a designated truth-value). On the contrary, if one regards 1 as the only designated truth-value (this is the position presented in (Kleene 1952)), the law of non-contradiction holds (there is *no* model where  $\phi \wedge \neg\phi$  gets a designated truth-value) but the law of the excluded middle fails (*there are* models where  $\phi \vee \neg\phi$  does *not* get a designated value).

Some authors such as Priest readily endorse these consequences whereas some of them try to come up with workarounds to preserve the apparent validity of these laws (for instance, (Tye 1994) defines two notions of *quasi-tautology* and *quasi-contradiction* which are in fact reminiscent of Priest's notion of quasi-validity). Some

---

<sup>9</sup>This is not entirely true. For instance, (Spector 2012) notes:

[In SK,]  $p \wedge (q \vee \neg q)$  is not equivalent to  $p$ . If  $p$  is true or false but  $q$  is undefined,  $p \wedge (q \vee \neg q)$  is undefined as well

systems also implement both Priest's and Kleene's notions of satisfaction in a unique logical system. This is the case of (Cobreros, Egré, Ripley & van Rooij forthcoming<sup>b</sup>)'s trivalent system which correspondingly defines two dual notions of satisfaction. This is also a way of approaching Dunn-Belnap's *four-valued* system: indeed, as (Muskens 1999) notes, we obtain the truth-tables in Table 2.2 once we remove any of the two extra truth-values from this system, but we will reject either the law of non-contradiction or the law of the excluded middle depending on what extra truth-value we choose to put aside.

### 2.2.1 Presupposition: George's Implementation

(George 2008) draws on (Kleene 1952)'s method to derive the truth-tables, which he formalizes with the help of what he calls a *repair function*. George's repair function takes bivalent functions as inputs (more precisely, functions going from a vector of truth-values in  $\{0,1\}$  – a *boolean* vector – to a single truth-value in  $\{0,1\}$ ) and outputs trivalent functions (more precisely, functions going from a vector of truth-values in  $\{0,\#,1\}$  to a single truth-value in  $\{0,\#,1\}$ ). As a matter of fact, the method that Kleene uses to present his truth-tables is formally very similar to the way that supervaluations are computed. The crucial difference is that contrary to supervaluationism that considers alternatives (classical valuations) by replacing each undefined *proposition* by a bivalent proposition, SK considers alternatives by replacing each indefinite *truth-value* by a bivalent truth-value, no matter whether two occurrences of an indefinite truth-value correspond to the same proposition. More formally, George's repair function can be stated in the following way :

**Definition 2.2.1** (George's SK Repair Function  $\mathcal{R}$ ). *For any function  $f : \overrightarrow{Bool} \rightarrow Bool$ ,  $\mathcal{R}(f)$  is the function such that, for any vector  $\vec{v}$  of truth-values in  $\{0,\#,1\}$  and of the same length as vectors in the domain of  $f$ :*

- $\mathcal{R}(f)(\vec{v}) = \#$  if and only if there exist two bivalently repaired vectors  $\vec{v}'_1$  and  $\vec{v}'_2$ , i.e. vectors whose elements are the same as those of  $\vec{v}$  except that each  $\#$  has been arbitrarily replaced either by 0 or by 1, and such that  $f(\vec{v}'_1) \neq f(\vec{v}'_2)$ ,
- $\mathcal{R}(f)(\vec{v})$  is the unique value that  $f$  returns for any pair of such bivalently repaired vectors otherwise.

As an example, let us derive negation and the conjunction operators defined in the truth-tables in Table 2.2. To do so, we will consider the function  $f_- : \overrightarrow{Bool} \rightarrow Bool$

and  $f_{\wedge} : \overrightarrow{Bool} \rightarrow Bool$  corresponding to the interpretations of  $\neg$  and  $\wedge$ :  $f_{\neg}$  takes a vector of one truth-value in  $\{0,1\}$  as its argument and outputs the other truth-value in the set;  $f_{\wedge}$  takes a vector of two truth-values in  $\{0,1\}$  and returns 1 if both truth-values are 1, 0 otherwise.  $f_{\neg}$  and  $f_{\wedge}$  are not defined over vectors of truth-values in  $\{0,\#,1\}$ , but  $\mathcal{R}(f_{\neg})$  and  $\mathcal{R}(f_{\wedge})$  are. Given Definition 2.2.1, when  $\mathcal{R}(f_{\neg})$  and  $\mathcal{R}(f_{\wedge})$  take boolean vectors as arguments, they will respectively output the same values as  $f_{\neg}$  and  $f_{\wedge}$  do for these boolean vectors. We straightforwardly obtain the  $\#$ -free lines of Table 2.2.

- $\mathcal{R}(f_{\neg})$

Given that  $f_{\neg}$  takes vectors of length 1,  $\mathcal{R}(f_{\neg})$  subsequently does too. Therefore, there is only one non-boolean vector in the domain of  $\mathcal{R}(f_{\neg})$ :  $\langle \# \rangle$ .

- $\langle \# \rangle$  is associated with two bivalently repaired vectors:  $\langle 0 \rangle$  and  $\langle 1 \rangle$ . We have  $f_{\neg}(\langle 0 \rangle) = 1$  and  $f_{\neg}(\langle 1 \rangle) = 0$ , thus  $f_{\neg}(\langle 0 \rangle) \neq f_{\neg}(\langle 1 \rangle)$ . Hence,  $\mathcal{R}(f)(\langle \# \rangle) = \#$ .

- $\mathcal{R}(f_{\wedge})$

Given that  $f_{\wedge}$  takes vectors of length 2,  $\mathcal{R}(f_{\wedge})$  subsequently does too. Therefore, there are five non-boolean vectors in the domain of  $\mathcal{R}(f_{\wedge})$ :  $\langle 0, \# \rangle$ ,  $\langle 1, \# \rangle$ ,  $\langle \#, 0 \rangle$ ,  $\langle \#, 1 \rangle$  and  $\langle \#, \# \rangle$ . Given that  $f_{\wedge}$  is symmetric over the values in its argument vector and that the repaired functions inherit this property,  $\mathcal{R}(f_{\wedge})$  too is symmetric over the values in its argument vector. We thus have  $\mathcal{R}(f_{\wedge})(\langle 0, \# \rangle) = \mathcal{R}(f_{\wedge})(\langle \#, 0 \rangle)$  and  $\mathcal{R}(f_{\wedge})(\langle 1, \# \rangle) = \mathcal{R}(f_{\wedge})(\langle \#, 1 \rangle)$ .

- $\langle 0, \# \rangle$  is associated with two bivalently repaired vectors:  $\langle 0, 1 \rangle$  and  $\langle 0, 0 \rangle$ . We have  $f_{\wedge}(\langle 0, 1 \rangle) = f_{\wedge}(\langle 0, 0 \rangle) = 0$ . Hence  $\mathcal{R}(f_{\wedge})(\langle 0, \# \rangle) = \mathcal{R}(f_{\wedge})(\langle \#, 0 \rangle) = 0$ .
- $\langle 1, \# \rangle$  is associated with two bivalently repaired vectors:  $\langle 1, 0 \rangle$  and  $\langle 1, 1 \rangle$ . We have  $f_{\wedge}(\langle 1, 0 \rangle) = 0$  and  $f_{\wedge}(\langle 1, 1 \rangle) = 1$ , thus  $f_{\wedge}(\langle 1, 0 \rangle) \neq f_{\wedge}(\langle 1, 1 \rangle)$ . Hence,  $\mathcal{R}(f)(\langle 1, \# \rangle) = \mathcal{R}(f)(\langle \#, 1 \rangle) = \#$ .
- $\langle \#, \# \rangle$  is associated with four bivalently repaired vectors:  $\langle 0, 0 \rangle$ ,  $\langle 0, 1 \rangle$ ,  $\langle 1, 0 \rangle$  and  $\langle 1, 1 \rangle$ . We have  $f_{\wedge}(\langle 0, 0 \rangle) = f_{\wedge}(\langle 0, 1 \rangle) = f_{\wedge}(\langle 1, 0 \rangle) = 0$  and  $f_{\wedge}(\langle 1, 1 \rangle) = 1$ , thus in particular  $f_{\wedge}(\langle 1, 1 \rangle) \neq f_{\wedge}(\langle 0, 0 \rangle)$ . Hence,  $\mathcal{R}(f)(\langle \#, \# \rangle) = \#$ .

The truth-functional aspect of the repair function appears clearly in these derivations: what  $\mathcal{R}$  ultimately operates on are *truth-values*, not propositions. To this

extent, it has no access to the content nor to the form of the propositions that the operators take as arguments, contrary to what happens in supervaluationism. A consequence of this property of such an implementation of SK is that there is no way of expressing the preservation-of-identity constraint. In particular, in models where any  $\phi$  and  $\psi$  get  $\#$ , SK has no way to distinguish between the conjunction  $\phi \wedge \neg\phi$  and the conjunction  $\phi \wedge \psi$  and *will not* assign 1 to the former but  $\#$  to the latter. The same reasoning holds for  $\phi \vee \neg\phi$  as compared to  $\phi \vee \psi$ . In the end, in such models, we have  $\phi \wedge \neg\phi = \phi \vee \neg\phi$ . This is why the law of the excluded middle and the law of non-contradiction cannot be simultaneously valid in the system, whatever status we give to the truth-value  $\#$ . This very truth-functional aspect of SK is what allows us to dress the truth-tables in Table 2.2, whereas such an enterprise would be hopeless for supervaluationism.

In (George 2008)’s approach, a proposition receives the third-value if and only if it is associated with a presupposition failure:

The presuppositions of a sentence are just the logical complement of its failure conditions. The discourse significance of presupposition and presupposition failure is left to the discourse model, presumably with a rule that it is inappropriate to utter a sentence the presuppositions of which you think another conversational participant might reasonably dispute.

With this understanding of *presuppositions* in mind, we can say that any statement presupposes that we are not in a situation where it lacks a bivalent truth-value. We can therefore compute the presuppositions of the sentences in (33) simply by looking at Table 2.2.

- (33)
- a. The wooden planets are flammable.
  - b. The wooden planets are not flammable.
  - c. If there are handcrafted planets, then the wooden planets are flammable.
  - d. There are handcrafted planets and the wooden planets are flammable.
  - e. Either there is no handcrafted planet or the wooden planets are flammable.

To determine the presupposition that the system associates with each sentence, we simply look at the lines where the corresponding proposition gets the value  $\#$  in Table 2.2: the presupposition is that we are not in situations compatible with the distribution of truth-values in these lines. In the case of (33-a) (to be translated as  $\phi$ ) and (33-b) (to be translated as  $\neg\phi$ ), things are very simple: there is only one

line where they get #, and this line corresponds to situations in which there is no wooden planet. Therefore, both (33-a) and (33-b) presuppose that we are not in such situations, that is to say, they presuppose that there are wooden planets. Now we assume no model assigns # to the proposition on the left of the connectives in (33-c) (to be translated as  $\phi \rightarrow \psi$ ), (33-d) (to be translated as  $\phi \wedge \psi$ ) and (33-e) (to be translated as  $\neg\phi \vee \psi$ ). For this reason, we don't need to look at the lines of Table 2.2 where  $\phi$  gets the value #. Among the lines where  $\phi$  does not get #, the only one where  $\phi \rightarrow \psi$  gets # is the line where  $\phi$  gets 1 and  $\psi$  gets #, which in the case of (33-c) represents situations where *there are* handcrafted planets and where there is no wooden planet. (33-c) thus presupposes that we are not in such a situation, that is to say, it presupposes that if there are handcrafted planets, then there are wooden planets. Among the lines where  $\phi$  does not get #, the only one where  $\phi \wedge \psi$  gets # is the same as before, so (33-d) also presupposes that if there are handcrafted planets, then there are wooden planets. Finally, among the lines where  $\phi$  does not get #, the only one where  $\phi \vee \psi$  gets # is the line where  $\phi$  gets 0 and  $\psi$  gets #. Because (33-e) is to be translated as  $\neg\phi \vee \psi$ , this line represents situations in which it is *not* the case that there is *no* handcrafted planet and in which there is no wooden planet. Presupposing that we are not in such a situation is again presupposing that if there are handcrafted planets, then there are wooden planets.

Now if you replace the left-parts of (33-c) and (33-d) with the presupposition of their right-parts, and if you replace the left-part of (33-e) with the negation of the presupposition of its right-part, you get the sentences in (34).

- (34)    a.    If there are wooden planets, then the wooden planets are flammable.  
           b.    There are wooden planets and the wooden planets are flammable.  
           c.    Either there is no wooden planet or the wooden planets are flammable.

The same method as above will lead us to associate these sentences with the conditional presupposition that if there are wooden planets, then there are wooden planets, which is tautological. Therefore, these sentences presuppose nothing (but tautologies).

But, as is clear from the truth-tables in Table 2.2, SK connectives are symmetric: they are not sensitive to the order in which their arguments are passed. This means that SK, as well as supervaluationism, predicts that (35-a) and (35-b) presuppose the same conditional.

- (35)    a.    The wooden planets are flammable and there are wooden planets.



- b. There are wooden planets and the wooden planets are flammable.

Noting that contrary to sentences like (35-b) which seems felicitous in most situations, sentences like (35-a) are felt to be generally infelicitous, George concludes that sentences of the form of (35-b) should indeed be associated with conditional presuppositions (which in the case of (35-a) is tautological) but that sentences of the form of (35-a) should unconditionally inherit the presupposition of the left conjunct. Therefore, he proposes another implementation of the repair function which derives asymmetric connectives. This leads to the truth-tables in Figure 2.3 which have also been proposed by Peters (Peters 1977) to deal with the same considerations.

$\phi$	$\psi$	$\neg\phi$	$\neg\psi$	$\phi \wedge \psi$	$\phi \vee \psi$	$\phi \rightarrow \psi$
0	0	1	1	0	0	1
0	#	1	#	0	#	1
0	1	1	0	0	1	1
#	0	#	1	#	#	#
#	#	#	#	#	#	#
#	1	#	0	#	#	#
1	0	0	1	0	1	0
1	#	0	#	#	1	#
1	1	0	0	1	1	1

Table 2.3: Peters' truth-tables

It is clear that a complex proposition built with such connectives will inherit any presupposition associated with the proposition appearing on its left. The new repair function that George proposes is thus *incremental*: it evaluates the left-most arguments of the connectives before evaluating those appearing on the right.<sup>10</sup> The function can be reformulated the following way.<sup>11</sup>

**Definition 2.2.2** (George's Peters Repair Function). *For any function  $f : \overrightarrow{Bool} \rightarrow Bool$ ,  $\mathcal{R}(f)$  is the function such that, for any vector  $\vec{v}$  of truth-values in  $\{0, \#, 1\}$  and of the same length as vectors in the domain of  $f$ :*

- $\mathcal{R}(f)(\vec{v}) = \#$  if and only if there exist two incrementally bivalently repaired vectors  $\vec{v}_1$  and  $\vec{v}_2$ , i.e. vectors whose elements are the same as those of  $\vec{v}$  until

<sup>10</sup>Although George notes that the linear order is one possible criterion among others, such as syntactic constituency for instance.

<sup>11</sup>Again, George's approach is finer and in the end, he proposes a general function which repairs any function on bivalent truth-values associated with a linguistic expression, be it an operator or a quantifier.

the first  $\#$ , starting from which each element has been arbitrarily replaced by either 0 or 1, and such that  $f(\vec{v}_1') \neq f(\vec{v}_2')$ ,

- $\mathcal{R}(f)(\vec{v})$  is the unique value that  $f$  returns for any pair of such incrementally bivalently repaired vectors otherwise.

Because contrary to the previous repair function, the new repair function arbitrarily replaces even bivalent truth-values *occurring after* a  $\#$ , when it evaluates a connective which takes a proposition of value  $\#$  on its left, it returns  $\#$ , regardless of the value of the proposition on its right. To this extent, when it repairs a proposition of value  $\#$ , it yields the same value that the previous repair function would have yielded if *every* proposition occurring later had been the value  $\#$ . However, to the extent that it still has access to the value of the propositions occurring *earlier*, it derives the same lines as in Table 2.2 when only  $\psi$  gets the value  $\#$ .

### 2.2.2 Vagueness: Tye's implementation

Surprisingly enough, not many authors have clearly endorsed the Strong Kleene definition of connectives to account for vagueness, but the clearest case may be (Tye 1994). Under his view, a predication of a borderline case yields a proposition with the third value. Interestingly, for Tye, the third value is actually not a proper *truth-value* but rather as a *truth-value gap*, and he says:

In my view, there are gaps due to failure of reference or presupposition and gaps due to vagueness.

He does not say however if all gaps should be formalized with the same third-value, but he claims in a note that when we observe truth-value gaps due to vagueness, “*something is said*” about a borderline case whereas he doubts that “*anything is said*” in the case of a presupposition failure. Unfortunately this is his only comparison between vagueness and presupposition. As will become clear in Chapter 4, a position assigning the same third truth-value to predications of borderline-cases and to proposition with an unfulfilled presupposition appears unrealistic when we look at the differences between the two phenomena.

As a reply to the critics mentioned above concerning the law of the excluded middle and the law of non-contradiction, Tye proposes two dual notions: quasi-tautologies and quasi-contradictions. A formula is a quasi-tautology if and only if there is no model in which it takes the value 0; and a formula is a quasi-contradiction if and only

if there is no model in which it takes the value 1. A quick look at the truth-tables in Fig.2.2 makes it obvious that  $A \vee \neg A$  and  $A \wedge \neg A$  are respectively a quasi-tautology and a quasi-contradiction.

Tye considers two versions of the sorites: a version with a universal premise, and a version with iterations of intermediate premises all of the same form. In both cases, his explanation hinges on the following reasoning: in a sorites series built from a vague predicate  $P$ , there are borderline cases for  $P$ , i.e. cases of which the predication of  $P$  will yield a truth-value gap, modeled by the third value  $\#$ . As such, it is not the case that all instances of the universal premise have the value 1, therefore the universal premise gets the value  $\#$  (given his definition of universal quantification, which can be seen as a big conjunction). For the same reason, when one tries to apply a sequence of vague conditionals one after the other, one will necessarily apply *indefinite* conditionals sooner or later, thus preventing one from continuing the application of the argument. However, Tye emphasizes that his position does not commit him to asserting that there is a pair of sentences like (36-a) and (36-b) such that (36-a) would be true and (36-b) *not true* (where *not true* means associated either with 0 or with  $\#$ ).

- (36)    a.    A man with  $N$  hairs on his head is bald.  
           b.    A man with  $N + 1$  hairs on his head is bald.

To him, given that *bald* is a vague predicate, there is just *no fact of the matter* whether such a pair of true and not true sentences exists. Furthermore, for Tye, there is no fact of the matter whether the predicate *true* is vague or not, i.e. if there are sentences like “(36) is true” which are indefinite. He concludes that under this view, *true* is a *vaguely vague* predicate, and that the sentence “There is a pair of sentences like those in (36) such that (36-a) is true and (36-b) is not true” is itself indefinite. In the end, Tye’s position leads to the conclusion that no formulation of the sorites argument is applicable, for vagueness makes us unable to designate two successive statements in a sorites series that sharply differ in their truth-status.

However, there is still a major problem with this analysis. We saw that there is no pair of sentences like those in (36) such that a model would assign 1 to (36-a) and 0 to (36-b). For this reason, certainly no model would assign 0 to (37-a), and certainly no model would assign 1 to (37-b). But no reasonable model would either assign both to any (36-a) and (36-b) the truth-value 1 or the truth-value 0 (this would consider every individual bald or every individual not bald), and no reasonable model would assign 0 to any (36-a) while assigning 1 to any (36-b) (this would go against the scalar

aspect of *bald*). For this reason, no model would assign 1 to (37-a) and no model would assign 0 to (37-b). In conclusion, both (37-a) and (37-b) should receive # in every model. The problem with this conclusion that (37-a) and (37-b) are equally indefinite is that it does not explain why speakers tend to feel the former true while they tend to feel the latter false.

- (37)    a.    For any definite number,  $N$ , if a man with  $N$  hairs on his head is bald then a man with  $N + 1$  hairs on his head is also bald.  
           b.    There is a definite number,  $N$ , such that a man with  $N$  hairs on his head is bald and a man with  $N + 1$  hairs on his head is not.

More generally, if Tye's position has the merit to let us escape the sorites paradoxes, it does not account for why the sorites argument is so compelling.

## 2.3 Predictions and Comparisons

As they have been presented here, these two systems mainly differ regarding how they deal with the law of the excluded middle and with the law of non contradiction. The status they attribute to the formulas  $\phi \vee \neg\phi$  and  $\phi \wedge \neg\phi$  are determined by the core properties of the systems, and therefore SK and supervaluationism make specific predictions for these sorts of disjunctions and conjunctions, regardless of whether  $\phi$  stands for a proposition about a borderline case (as exemplified in (38) with Blob a borderline-red patch) or for a proposition associated with a presupposition failure (as exemplified in (39)).

- (38)    a.    Blob is red or Blob is not red  
           b.    Blob is red and Blob is not red
- (39)    a.    The wooden planets are flammable or the wooden planets are not flammable  
           b.    The wooden planets are flammable and the wooden planets are not flammable
- (40)    a.    Alien life exists or alien life does not exist  
           b.    Alien life exists and alien life does not exist

As mentioned above, SK has usually been attacked on the grounds that it does not make both  $\phi \vee \neg\phi$  a tautology and  $\phi \wedge \neg\phi$  a contradiction (despite Tye's defense of this result). But we can question whether we actually feel (38-a) and (39-a) true and

(38-b) and (39-b) false. Concerning vagueness, it seems that (38-a) is at least not as readily accepted as true as (40-a) is; and it seems that (38-b) is at least not as readily rejected as false as (40-b) is. In fact, it seems that we could even reject (38-a) as false or accept (38-b) as true on the grounds that Blob is borderline-red. This suggests that the alleged superiority of supervaluationism over SK is not as grounded as it is sometimes claimed. Things are slightly different concerning presupposition. It seems to me that we are also reluctant to accept (39-a) as true (in comparison with accepting (40-a) as true) and we might in fact reject it as false; but even if we have some reluctance to reject (39-b) as false, I hardly imagine accepting it as true (in contrast with (38-b)). Table 2.4 summarizes these judgments. They are meant to correspond to possible judgments, even though we would probably not give them in a row when asked to judge a sentence.

	False	Neither	True
(38-a) (Blob red or not red)	Yes	Yes	Yes
(38-b) (Blob red and not red)	Yes	Yes	Yes
(39-a) (The ... flammable or ... not flammable)	Yes	Yes	No?
(39-b) (The ... flammable and ... not flammable)	Yes	Yes	No

Table 2.4: Possible judgments for (38) and (39)

To this extent, supervaluationism is threatened by our intuitive judgments in situations of borderline cases and in situations of presupposition failure. We may therefore want to turn to SK, and see how this system can deal with these intuitive judgments. In SK as initially in supervaluationism, all the sentences in (38) and (39) correspond to propositions which connect two propositions of intermediate value. As a consequence, SK assigns the third value to all the resulting propositions, which are therefore considered neither true nor false. This would account for why we are reluctant to give a *true* judgment for (38-a) and (39-a) and to give a *false* judgment for (38-b) and (39-b). This however does not tell us why we also have the possibility to give bivalent judgments for them at all.

One possibility is to resort to an accommodation operator which has the effect of turning any proposition of the intermediate value into a proposition of value 0. (Beaver 2001) traces this operator back to (Bochvar 1939) and discusses it in terms of *meta-assertion*: meta-asserting a proposition is asserting that this proposition is true *simpliciter*. As a consequence, meta-asserting simple sentences like (41-a) or (41-b) is asserting that they are true *simpliciter*. But in models consistent with Blob

being borderline-red and there being no wooden planet, the propositions corresponding to (41-a) and (41-b) would normally receive an intermediate value. Therefore, their meta-assertion asserts something false, because (41-a) and (41-b) are *not* true *simpliciter*. (41-c) sketches the semantics of this accommodation operator.

- (41) a. Blob is red.
- b. The wooden planets are flammable.
- c.  $\mathcal{I}(\mathcal{A}(\phi)) = 1$  iff  $\mathcal{I}(\phi) = 1$ , 0 otherwise.

Things get very interesting when one allows this operator to be freely embedded under other operators. Typically, the negations of (41-a) and (41-b) represented in (42-a) and (42-b), can now be translated by three different types of propositions listed in (42-c), (42-d) and (42-e).

- (42) a. Blob is not red.
- b. The wooden planets are not flammable.
- c.  $\neg\phi$
- d.  $\mathcal{A}(\neg\phi)$
- e.  $\neg\mathcal{A}(\phi)$

If  $\phi$  receives an intermediate value, then negation in SK gives (42-c) an intermediate value too, and (42-d) gets the value 0 for it is the meta-assertion of (42-c) which has not gotten the value 1, whereas (42-e) gets the value 1 for it is the negation of the meta-assertion of  $\phi$ , that is to say the negation of a *false* meta-assertion. This last possibility, where the accommodation operator takes scope under the negation and which results in a true negative sentence, is often discussed under the expression of *local accommodation* in the literature (this is precisely how (Beaver 2001) introduces it), but this expression is also sometimes used to refer to the resulting interpretation only, with no further commitment to the existence of an accommodation operator. Note that the truth-values that SK derives for these simple sentences when it is augmented with an accommodation operator are welcome. Indeed, let us imagine that this operator expresses something of the form *It is plainly true that....* The sentences *It is plainly true that Blob is red* and *It is plainly true that the wooden planets are flammable* are clearly false when we believe Blob to be borderline-red and there to be no wooden planet; and so are the sentences *It is plainly true that Blob is not red* and *It is plainly true that the wooden planets are flammable*. And because we can embed this operator under negation, we have to consider the sentences *It is not plainly true*

*that Blob is red* and *It is not plainly true that the wooden planets are flammable*, which sounds true. Looking back at (38-a) and (39-a), embedding this accommodation operator under the negation as we just considered has the effect of yielding a true disjunct in each case and thus to make these sentences true, accounting for their tautological flavor (see (43-a) and (43-b) for a possible translation of these interpretations). When the accommodation operator directly applies to any conjunct of the conjunctions in (38-b) and (39-b), this has the effect of yielding a false conjunct in each case and thus to make these sentences false, accounting for their contradictory flavor (see (43-c) and (43-d) for a possible translation of these interpretations). In addition, when the accommodation operator directly applies to both *disjuncts* in (38-a) and (39-a) (and importantly, when the accommodation operator does not appear under the negation), these sentences end up false, in accordance with a tendency in our intuitions (see (43-e) and (43-f) for a possible translation of these interpretations).

- (43)
- a. It is plainly true that Blob is red or it is not plainly true that Blob is red.
  - b. It is plainly true that the wooden planets are flammable or it is not plainly true that the wooden planets are flammable.
  - c. It is plainly true that Blob is red and it is not plainly true that Blob is red.
  - d. It is plainly true that the wooden planets are flammable and it is not plainly true that the wooden planets are flammable.
  - e. It is plainly true that Blob is red or it is plainly true that Blob is not red.
  - f. It is plainly true that the wooden planets are flammable or it is plainly true that the wooden planets are not flammable.

SK augmented with an accommodation operator seems to accommodate our intuitions quite well in terms of truth-value judgments and derives them in a very reasonable manner. However it remains to explain why the vague conjunction (38-b) is sometimes felt to be true, whereas the presuppositional conjunction (39-b) is not. This raises the problem of the asymmetry between vagueness and presupposition, which can even be observed at the atomic level: it seems possible to judge (41-a), “Blob is red”, true to the extent that Blob is borderline-red, but it seems totally impossible to judge (41-b), “The integer between 2 and 3 is even”, true given that there is no wooden planet.

This suggests that, at some level, vagueness and presupposition have to be approached differently. In the context of SK, one possibility is to imagine a “*tolerant*” operator that would be the vagueness-specific dual of the accommodation operator: it would only turn vague propositions with an intermediate value into propositions of value 1 and would leave any other proposition unchanged. Its semantics is sketched in (44).

$$(44) \quad \mathcal{I}(\mathcal{T}(\phi)) = 0 \text{ iff } \mathcal{I}(\phi) = 0, 1 \text{ otherwise.}$$

These two operators can be regarded as syntactically implementing the two positions on the designated status of  $\#$  discussed above, where  $\mathcal{A}$  is to be associated with (Kleene 1952)’s position (resulting in a full logical system usually referred to as K3) and  $\mathcal{T}$  is to be associated with (Priest 2006)’s position (resulting in a full logical system usually referred to as LP). Interestingly, they echo two dual notions of negation discussed in the literature on trivalence which differ in which bivalent value they choose to map  $\#$  to. (Alxatib & Pelletier 2011) call them the *intuitionistic* negation (noted  $-$ , also called *Gödel’s* negation in the literature) and the *exclusion* negation (which they note  $\neg$  but which we will note  $\sim$ ). All these notions then form a square of opposition represented in Fig. 2.1 (labels have been added for ease of interpretation, but one should keep in mind that each notion is defined as primitive here).

Because they enter in such a relationship and because the initial negation  $\neg$  leaves the intermediate value unchanged, any of these operators taken in par with  $\neg$  can be used to define the three others, as exemplified in (45).

$$(45) \quad \begin{array}{ll} \text{a. } \mathcal{A}: & -\phi \equiv \neg \mathcal{A}(\phi), \sim \phi \equiv \mathcal{A}(\neg \phi), \mathcal{T}(\phi) \equiv \neg \mathcal{A}(\neg \phi) \\ \text{b. } \mathcal{T}: & \sim \phi \equiv \neg \mathcal{T}(\phi), -\phi \equiv \mathcal{T}(\neg \phi), \mathcal{A}(\phi) \equiv \neg \mathcal{T}(\neg \phi) \\ \text{c. } \sim: & \mathcal{T}(\phi) \equiv \neg \sim \phi, \mathcal{A}(\phi) \equiv \sim \neg \phi, -\phi \equiv \neg \sim \neg \phi, \\ \text{d. } -: & \mathcal{A}(\phi) \equiv \neg - \phi, \mathcal{T}(\phi) \equiv - \neg \phi, \sim \phi \equiv \neg - \neg \phi, \end{array}$$

This suggests that one of these operators could have a primitive role in our intuitive understanding of natural language. The labels we naively used in Fig. 2.1 suggest  $\mathcal{A}$  and  $\sim$  as good primitives, but falsity is also sometimes felt as being built on truth. In the end, these considerations suggest that  $\mathcal{A}$  might play a primitive role in the derivation of our truth-value judgments, and that embedding negation below or above this operator might be costlier and that the resulting truth-value judgments should be observed less often or given with a more important delay than the truth-value judgments resulting from non-embedded structures. Leaving these considerations aside, the  $\mathcal{T}$  operator would make (41-a) true and would therefore also make the



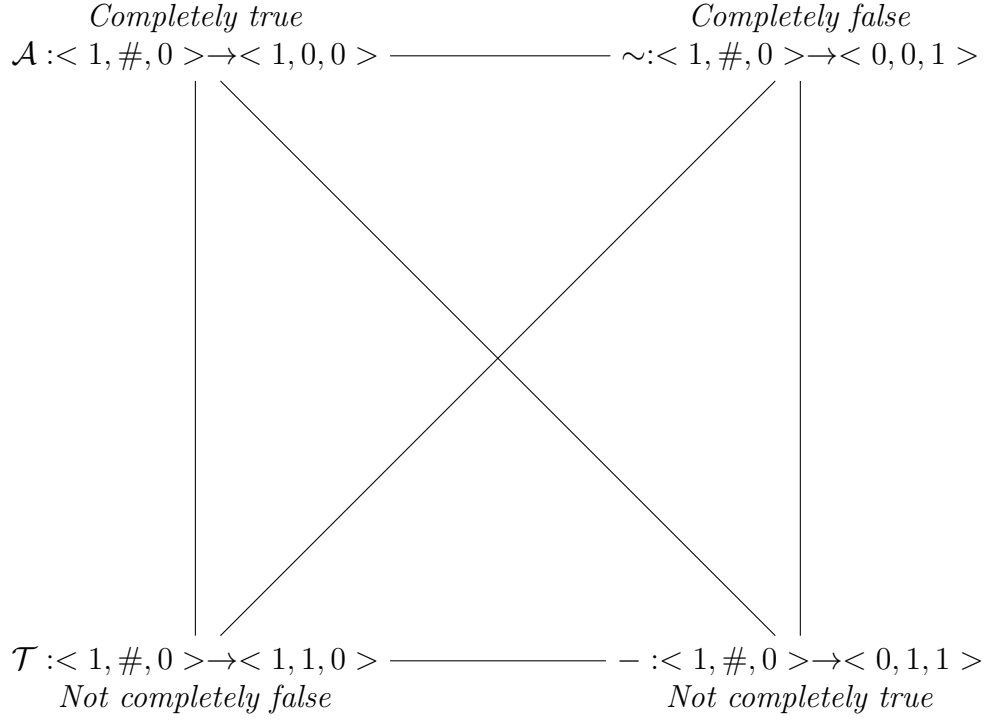


Figure 2.1: Operators' and negations' square of opposition

conjunction (38-b) true either when applied globally or when applied to each conjunct that initially received an intermediate value. However, to make this operator sensitive to the distinction between vague and presuppositional propositions, one either has to include a non truth-functional mechanism, but one would thus loose the specificity of SK over supervaluationism, or to define different intermediate values for vagueness and for presupposition. This is this latter possibility that I will pursue in Chapter 4 and 6.

A different option from resorting to a “tolerant” operator is to consider that the intermediate value associated with vagueness should be part of the set of logical values designated as representing the truth, as is considered the intermediate value in (Priest 2006)’s Logic of Paradox and as it is advocated for by (Ripley 2013) in his own comparison between supervaluationism and truth-functional systems for vagueness. Yet another option is to resort to one system for one phenomenon and to the other system for the other phenomenon. Let us see what tools supervaluationism has to offer to deal with our intuitive judgments on the present sentences. As sentences with a tautological form, supervaluationism predicts *true* judgments for (38-a) and (39-a); and as sentences with a contradictory form, supervaluationism predicts *false* judgments for (38-b) and (39-b). But supervaluationism has more to say about the

*true* judgment that we can give for (38-a), and this has to do with the intrinsically non truth-functional notion of *penumbral connections* that Fine introduces. Penumbral connections put constraints over the classical valuations that one should consider when evaluating a proposition of intermediate value. Fine distinguishes two sorts of penumbral connections: *internal* and *external* penumbral connections. Internal penumbral connections deal with one predicate at a time. Fine proposes one type of internal penumbral connection that ensures that the only classical valuations that we consider are consistent with the order of the borderline cases on the scale associated with the adjective: any borderline case that ranks higher on the adjectival scale than a borderline case in the considered extension of the adjective must be in the extension too.<sup>12</sup> External penumbral connections concern the logical relations that hold *between* the predicates of the language. As a typical external penumbral connection, Fine claims that any assignment of a specific bivalent extension to *red* should come on a par with a bivalent extension for *pink* so that the two do not intersect. In other words, because these predicates are contradictories according to Fine, the classical valuations we consider must never make an entity both red and pink. Interestingly, the tautological flavor of (38-a) can be retrieved indirectly from this external penumbral connection. To see this, let us accept that our borderline-red Blob is either red or pink. The external penumbral connection tells us that Blob cannot be both red and pink, so if it is pink, then it is not red. Therefore, Blob is either red or not red.<sup>13</sup>

But supervaluationism still has to explain why we are quite reluctant to give these judgments, and why we even sometimes judge (38-a) and (39-a) false and (38-b) true. This is a difficult task, for even though there is an initial level where the propositions corresponding to the simple sentences (41-a) and (41-b) lack a truth-value, there is no such level for the disjunctions in (38-a) and (39-a) and for the conjunctions in (38-b) and (39-b). Indeed, the main idea of supervaluationism is to stick to a classical semantics for the connectives: as such, it assigns no logical truth-value to disjunctions and conjunctions that contain propositions with an intermediate value, except for their derived “super-value”. However, if one had to give a logical truth-value for these sentences *before* computing it over their classical valuations (and assuming that the algorithm underlying supervaluationism actually consists of some repair strategy), it would be reasonable to imagine that each of them receive an intermediate truth-value

---

<sup>12</sup> Note that this monotonicity constraint (see fn. 8) proves useful in modeling the comparative use of gradable adjectives. Trivalent truth-functional systems like SK cannot resort to penumbral connections, but see (Burnett 2012) who bases on (Cobreros et al. 2012) to model gradable adjectives and their monotonicity aspect within a trivalent framework.

<sup>13</sup> Thanks to Paul Egré who presented this argument to me in a personal communication.

to the extent that they all contain propositions of intermediate truth-value. This amounts to what is usually called Weak Kleene: whatever the form of a proposition, it gets the intermediate truth-value as soon as it contains a proposition which fails to get a bivalent truth-value. Imagining that this failure is thus initially contagious (and then possibly resolved by the supervaluationist algorithm) would account for our reluctance toward giving a bivalent truth-value judgment for the present sentences. And if we further imagine meta-assertion as a possible global mode of assertion (but not anymore as an *operator* which could be embedded) then we would account for the *false* judgments that we can give for (38-a) and for (39-a) given their initial truth-value would be different from 1. I cannot see, however, any way how supervaluationism could account for the *true* judgments that we can give for (38-b). As a last remark on how supervaluationism faces our judgments, note that the simple vague sentence (41-a) as well as its negation can trigger *true* and *false* judgments to the extent that Blob is borderline-red, in addition to the *neither true nor false* judgment. These exhaustive bivalent truth-value judgments might reflect the diverse classical valuations under consideration when we evaluate the sentence. The simple presuppositional sentence (41-b) does not trigger such a variety of judgments though: only *false* judgments are observed in addition to *neither true nor false*, to the extent that there is no wooden planet. If we still want to say that speakers can give all the truth-value judgments associated with each classical valuation they consider, then one has to say that (41-b) is superfalse: it would initially lack a truth-value, but there would be a constraint forcing us to consider only classical valuations where (41-b) gets the value 0. For instance, we could imagine a penumbral connection which would have the effect that whenever the proposition corresponding to (41-b) gets the value 1, then the proposition expressing that there are wooden planets should receive the value 1 too, and this last assignment would go against our beliefs in situations of presupposition failure. This would still make (39-a) supertrue and (39-b) superfalse while initially lacking a bivalent truth-value under a Weak Kleene pre-supervaluationist approach. Penumbral connections make it rather natural to make (41-b) superfalse while leaving (41-a) as lacking a bivalent truth-value: the non-truth-functional aspect of supervaluationism here is a clear advantage over SK for specific treatments of vagueness and presupposition.

In sum, SK and supervaluationism can be enriched so as to make similar predictions, in line with our judgments. They differ however in how they derive these truth-value judgments. SK favors judgments in line with a lack of truth-value and it needs additional mechanisms to derive bivalent truth-value judgments: these could

therefore be expected to appear less often or to take a little more processing time to obtain, with a possible observable scope effect regarding the operators (eg. with the *true* reading of the disjunctions being more demanding given the need of placing the accommodation operator under the scope of negation). In contrast, if we take the value that supervaluationism computes for propositions as corresponding to our most natural judgments, this system favors bivalent truth-value judgments for the disjunctions and the conjunctions in (38) and in (39). We could also imagine that there is an initial representation where these sentences lack a truth-value and that we should therefore expect a lack of bivalent truth-value judgment, as does SK. But importantly, SK assigns the intermediate truth-value to the simple sentences in (41-a) and (41-b) as well as to the corresponding disjunctions and conjunctions. In contrast, supervaluationism naturally derives bivalent truth-values for the latter but sticks to a lack of bivalence for the simple sentences (or at least for the vague simple sentence, if we take propositions with unfulfilled presuppositions to be superfalse). A review of the experimental literature on truth-value judgments for vagueness and for presupposition will prove useful in deciding what kind of system to use to deal with both of these phenomena.

## Chapter 3

# Experimental Literature on Vagueness and on Presupposition

As put forward in the previous section, the non-bivalence of vagueness and the non-bivalence of presupposition have long been discussed in the theoretical literature. Recently, some authors have started to experimentally investigate these phenomena. Some of these experiments partly or totally address the question of the possible truth-value judgments that speaker access in situations of borderline cases or in situations of presupposition failure. In the end, the results overall confirm that the judgments that speakers give for vague and presuppositional sentences in critical situations is different from those that they give for classical bivalent sentences. In addition, they suggest that these two types of sentences might trigger specific truth-value judgments, arguing for a specific treatment of each phenomenon.

I will review four experiments on vagueness. First, I will present two studies out of four by (Serchuk, Hargreaves & Zach 2011), which asked subjects to judge sentences of different forms about borderline cases by choosing among a variety of truth-value judgments. Second, I will discuss a study by (Alxatib & Pelletier 2011) which focused on affirmative and negative counterparts of vague sentences and on conjunctions and disjunctions like those in (38). Third, I will turn to (Ripley 2011)'s study, which can be seen as a variation of (Alxatib & Pelletier 2011)'s experiment. Finally, I will present (Egré, de Gardelle & Ripley 2013)'s series of experiments on color adjectives, which also focused on negation, conjunctions and disjunctions.

There seem to be very few experiments investigating the non-bivalent status of sentences whose presuppositions are unfulfilled, in comparison to the importance of this topic in the theoretical literature. In the following, I will review two experiments on presupposition. First, I will discuss (Schwarz to appear)'s study, which compared delays in giving a *false* judgment for sentences with non-referring definite

noun-phrases and for their existential, non-presuppositional counterparts. Then, I will turn to (Abrusán & Szendrői 2013)’s experiment where participants faced a threefold choice when evaluating a variety of sentences with unfulfilled presuppositions, both in their positive and negative forms.

### 3.1 Vagueness

(Serchuk et al. 2011) conducted four different experiments, all investigating questions about gradable adjectives and borderline cases. Because their second and third experiments did not investigate truth-value judgments, we will not discuss them here.<sup>1</sup> In their first and fourth experiment, though, participants were asked to choose their answer among a variety of truth-value judgments. These two experiments involved conjunctive descriptions like (38-b) and one of them also involved disjunctive descriptions like (38-a). In their first experiment, they explicitly asked participants to imagine a woman, Susan, who was described as somewhere between clear instances of the adjective and clear counter-instances of the adjective. The adjective was either *rich* or *heavy*, and the description to be evaluated was either of the form *Susan is ADJECTIVE* or *Susan is definitely ADJECTIVE*. Then participants had to check one truth-value judgment among *true*, *false*, *both true and false*, *partially true and partially false*, *not true, but also not false* and *true or false, but I don’t know which*. The presence of *definitely* had the effect that a vast majority of participants checked *false* for either adjective (73.3% for *rich*, 73.6% for *heavy*). When *definitely* was absent though, the answers for the description containing *heavy* distributed over all the judgments, with a tendency toward *not true, but also not false* (30.7%) and *partially true and partially false* (25.6%), and against *true* (10.2%) and *both true and false* (8.5%). The answers for the description containing *rich* were similar, except that participants preferred *false* (27%) over *partially true and partially false* (17.8%). Since they were only interested in the effect of the presence/absence of *definitely*, they did not discuss the specific truth-value judgments nor provide any statistical analysis for them. Their fourth experiment was a variant of the first one, where they simply replaced the affirmative descriptions with several descriptions involving negations. One form of these descriptions was *Susan is not ADJECTIVE*, the negative counterpart of the *definitely*-free sentence of their first experiment. They basically observed the same judgments with this polarity, with a flatter distribution over *not true, but also*

<sup>1</sup>The second experiment was a variant of an experiment by (Bonini, Osherson, Viale & Williamson 1999) asking subjects for the limit heights where they could apply *tall* or *not tall*. The third experiment investigated speakers’ inferences when facing sorites paradoxes.

*not false, partially true and partially false, true and false with rich*. Another form of these descriptions was *Susan is ADJECTIVE* and *Susan is not ADJECTIVE*, and yet another one was *Either Susan is ADJECTIVE or Susan is not ADJECTIVE*. A majority of participants judged the former descriptions *false* (55.7% cross-adjectives), with three times less people judging it *true* (18.9%, the other judgments proved to be very low). Interestingly, for the latter descriptions, the judgments were equally divided between *true* (32.3%) and *false* (39.1%, the other judgments were again very low). Even though these results are hard to interpret, in part because the authors did not have the same questions in mind as ours when creating the design and conducting the analyses, they seem compatible with a view where glutty judgments are dispreferred in favor of gappy judgments.

(Alxatib & Pelletier 2011) presented their subjects with a picture representing five men of various heights and for each of these men participants had to indicate their judgments on a hard-copy for 4 descriptions: *the man is tall*, *the man is not tall*, *the man is tall and not tall* and *the man is neither tall nor not tall*. To this end, they could check *true*, *false* or *can't tell*. The rate of *can't tell* responses proved to be very low for every condition. For simple affirmative and negative sentences, participants behaved classically: when they accepted one as true for a man, they rejected the other as false for the same man. About half of the participants chose to qualify the man of average height as *tall* and about half of them too chose to qualify him as *not tall*. It turned out that 44.7% of participants accepted the glutty description as true for the same man (contra 14.5% for the smallest man and 5.3% for the tallest man) and 53.9% of participants accepted the gappy description as true for him (contra 27.6% for the smallest man and 6.6% for the tallest man). In addition, both for participants who accepted the glutty description and for participants who accepted the gappy description, more than half of them also accepted the other complex description. Therefore, even though simple sentences provide no strong evidence for a non-bivalent theory of vagueness, the high acceptance of glutty and gappy descriptions strongly argues in favor of such an approach. Importantly, these results ask for a system where the logical value of vague sentences about borderline cases derives both gappy and glutty judgments.

(Ripley 2011) also investigated gappy and glutty descriptions, and found results similar to the ones just discussed. All the participants were presented with the same picture (projected onto a screen) representing seven pairs of a square and a circle aligned horizontally, with the distance between them varying across the pairs. The pair presenting the highest distance between the square and the circle was at the top

of the picture, and the distance decreased with each pair so as to have side by side figures at the bottom of the picture. Depending on the group they belonged, they were asked to rate a gappy or a glutty description constructed with *near* on a seven point scale for each pair (there were four groups, with two different gappy descriptions and two different glutty descriptions). In each group, there was a pair for which the mean rate that participants gave was important. Individually, over half of the participants rated the description 6/7 or 7/7 for one of the seven pairs. All of the four highest mean scores (5.2/7, 5.3/7, 5.7/7, 5.1/1=7) were significantly above 4, suggesting that participants did not answer by chance for borderline-near figures but actually felt the description to be admissible. In addition, the statistical analyses revealed no significant difference between the groups' answers, which means that both glutty and gappy descriptions were as readily accepted.

Finally, (Egré et al. 2013) conducted an experiment on color terms. Participants had to agree or disagree with descriptions containing a color term *COLOR* against a series of patches whose color varied along a scale from the color *COLOR* to another color. Participants were divided in three groups: one group saw the patches in random order, one saw them in ascending order and another one saw them in descending order. In the end, participants had given 8 judgments for each patch. Some of the descriptions that they had to judge was of the form *the square is COLOR and not COLOR*. Whether the patches were presented in random, ascending or descending order proved to have no crucial effect for these descriptions. They found that participants agreed with these glutty descriptions more than half the time for patches in the central region, and even significantly more often than simple descriptions of the form *the square is COLOR* or *the square is not COLOR* judged in isolation for the same patches. These are very compelling results.

(Serchuk et al. 2011)'s and (Alxatib & Pelletier 2011)'s experiments are the only ones among those discussed here where participants did not face a bivalent forced-choice, and only in (Serchuk et al. 2011) did they make use of them. In this study though, they seem to have preferred glutty and gappy truth-value judgments over the others for either polarity. Even though this is in contradiction though with the low acceptance of glutty descriptions that they observed in their fourth experiment, this is in accordance with the even repartition of *true* and *false* responses of subjects to simple affirmative and negative vague descriptions of borderline cases observed in the three other studies, and with the high acceptance rate of gappy and glutty descriptions that they report. In addition, (Serchuk et al. 2011)'s fourth experiment is the only one to have tested disjunctive sentences of the form of (38-a). Despite their tautological



form, half of the participants rejected them as false. On the basis of these results, Table 3.1 summarizes the bivalent and non-bivalent truth-value judgments that simple affirmative and negative vague descriptions of borderline cases can trigger. As a matter of fact, these descriptions seem to yield extremely liberal evaluations. Table 3.2 summarizes the bivalent truth-value judgments that conjunctive and disjunctive vague descriptions of borderline cases can trigger. Once again, they appear to favor much tolerance in their evaluation.

	Both true and false	True	False	Neither true nor false
X is ADJ	✓	✓	✓	✓
X is not ADJ	✓	✓	✓	✓

Table 3.1: Available truth-value judgments for simple affirmative and negative vague sentences describing borderline cases after experimental observations

	True	False
X is ADJ or not ADJ	✓	✓
X is neither ADJ nor not ADJ	✓	✓
X is both ADJ and not ADJ	✓	✓

Table 3.2: Available truth-value judgments for complex vague sentences describing borderline cases after experimental observations

All the trivalent approaches discussed above account for the absence of effect of polarity, for negation does not change the intermediate truth-value in any system. The observation of glutty truth-value judgments and the acceptance of glutty descriptions from argues in favor of systems such as LP where vagueness is associated with a *designated* intermediate truth-value. But the observation of gappy truth-value judgments and the rejection of vague sentences with a tautological form argues in favor of systems such as K3 where vagueness is associated with a *non-designated* intermediate truth-value. The more theoretical considerations of (Ripley 2011) echo these observations, and as we will see in Chapter 4, Cobreros et al. take them results in account and consequently propose a logical system which, as they note, can be seen as incorporating the assets of both LP and SK in (Cobreros et al. 2012) and (Cobreros et al. forthcoming).

## 3.2 Presupposition

As pointed out earlier, there are, to my knowledge, very few experiments directly addressing the question of non-bivalent *truth-value* judgments for presuppositional sen-

tences. Yet authors usually distinguish two main competing views of presupposition differing in what truth value they assign to these sentences in situations of presupposition failure. On the one hand is a view which can be called a *Frege-Strawsonian* view and which claims that a simple sentence whose presupposition is unfulfilled lacks a truth-value. On the other hand is a view which is commonly designated as *Russellian* and which claims that a simple sentence whose presupposition is unfulfilled is false *simpliciter*.<sup>2</sup> But as a matter of fact, most of the truth-value judgment tasks on presupposition involve forced bivalent choice. The reason for this may lie in the fact that the two views just mentioned make different predictions on the projection of presuppositions in complex sentences, which can usually be investigated with a bivalent task by collecting various measures.<sup>3</sup> But discarding one view on the basis of its predictions concerning presupposition projection does not suffice *per se* to establish that the other view is correct and more importantly for the present purpose, the vast majority of these studies do not inform us on the nature of the possibly non-bivalent truth-value judgments of speakers in situations of presupposition failure.

We can nonetheless find clues about them in some experiments resorting to a forced choice bivalent truth-value judgment task. This is the case of the two experiments in (Schwarz to appear). Because the second experiment was a refinement of the first experiment and provided consistent, clearer results, I will not discuss the first one. In the critical condition, participants were asked to judge a simple affirmative presuppositional sentence like (46-a) against a picture describing two boys' week schedules where no boy had an outing on Tuesday. (46-b) served as a baseline false sentence.

- (46) a. The boy with an outing on Tuesday is going to play golf.
- b. There's a boy with an outing on Tuesday who's going to play golf.

When evaluated against the sort of pictures presented in the critical condition, (46-a) has an unfulfilled presupposition, namely that there is a boy with an outing on Tuesday. In contrast, in the same situation, (46-b) simply *asserts* this statement, thus

---

<sup>2</sup>These views originally discuss non-referring singular terms. The most discussed such terms are names of fictive characters like *Pegasus* and definite expressions who lack a referent, like *the king of France in 2014*.

<sup>3</sup>For instance, (Chemla & Bott 2013) found that a *false* reading of negative sentences with an unfulfilled presupposition like (i) was faster to compute than a *true* reading.

- (i) The zoologists don't realize that the birds are mammals.

These results argue against an implementation of the Russellian view where (i) is *true simpliciter* and which needs a subsequent, time-consuming process to derive the *false* judgment.

making the sentence false.<sup>4</sup> Schwarz collected and compared the time participants took to give a *false* judgment for (46-a) and the time participants took to give a *false* judgment for (46-b). Participants showed a significantly higher delay in giving a *false* judgment for (46-a) than they did in giving a *false* judgment for (46-b). To the extent that the *false* judgment that participants gave for (46-b) is a regular, typical case of a *false* judgment, the delay we observe with (46-a) indicates that the associated *false* judgment is not straightforward. Schwarz lists several possible explanations for this delay, one of them being that the *false* judgment is not immediately available to the speaker but needs to be derived. If this explanation is correct, then it means that in the evaluation of a simple sentence with an unfulfilled presupposition (at least for presuppositions associated with the definite article), speakers go through a primordial mental state where the sentence is neither assigned *true* nor *false*. Even though there are other possible explanations for them, it should be noted that these results are expected from a Frege-Strawsonian perspective. I therefore take them as partial evidence that in situations of presupposition failure, simple propositions should be assigned a non-bivalent logical truth-value.

Further evidence come from the only experiment I know of which presented subjects with a threefold choice. (Abrusán & Szendrői 2013) investigated the judgments of speakers for the positive and negative counterparts of presuppositional sentences of various forms. All the test sentences contained a definite description lacking a referent, such as *the king of France*. Participants could judge the sentences they evaluated by choosing the option *true*, the option *false* or the option *can't say*. For all the simple affirmative sentences containing these non-referring definite descriptions, such as (47-a), participants chose the *false* option at a non significantly different rate from the control false sentences. Interestingly though, participants' answers for their negative counterparts varied depending on the form of the sentence. Importantly, they mostly judged sentences such as (47-b) (the negative counterpart of (47-a)) false.

- (47)    a.    The king of France is bald  
           b.    The king of France is not bald

However, their rate of *false* judgments for sentences such as (47-b) was significantly different from their rate of *false* judgments for control sentences. Abrusán & Szendrői consider this significant difference hard to explain but indicate in a footnote that a 4-valued logic for presupposition may be able to explain this difference. This is of

---

<sup>4</sup>In addition, there is no expression triggering an unfulfilled presupposition in the rest of the sentence.

importance, for the 5-valued system presented in the next chapter can be seen as merging a trivalent logic for vagueness and a *4-valued logic for presupposition*. They also looked at four other types of negative presuppositional sentences exemplified in (48).

- (48)    a.    The King of France is not on a state visit to Australia this week.  
           b.    The King of France, he did not call Sarkozy last night.  
           c.    Sarkozy, he did not call the King of France last night.  
           d.    The King of France is not married to Carla Bruni.

Sentences such as (48-a) and (48-b) were judged *true* almost half of the time, whereas sentences such as (48-c) and (48-d) were judged true most of the time. Surprisingly enough, the authors did not report any analysis of the rate of *can't say* answers. Yet the table they furnish in their Appendix 2 exhibits a rather high rate of *can't say* answers for (47-a) (19.1%), (47-b) (33.9%), (48-a) (28%), (48-b) (26.8%) and (48-c) (19.5%) in comparison to what the one reported for the clear false sentences such as (49) with no presupposition failure (9.5%).

- (49)    France has a king, and he is bald.

Because no analysis is provided for these data (the authors only report to have analyzed the rate of *false* answers for affirmative presuppositional sentences and the rate of *true* answers for their negative counterpart), we can but suspect that the rate of *can't say* answers was higher for these sentences than for control sentences. The point of this experiment was to test three categories of theories of presupposition by investigating various sentences that would possibly call for fallback strategies in the evaluation of their truth-value judgment. In this perspective, the sentences such as (47-a) and (47-b) served as baselines, that is to say as sentences with unfulfilled presupposition where the observed truth-value judgments would not be affected by fallback strategies, if any. This is why I will take the judgments that participants gave for these sentences as aiming the “pure” judgments available for sentences with unfulfilled presuppositions. Therefore, in Figure 3.1, I report the rate of answers that Abrusán & Szendrői indicate in their appendix for the baseline presuppositional sentences of the form of (47-a) and (47-b) along with the results they report for the clear *false* sentences such as (49).

Once again, these results can be taken as indicative of a non-bivalent status of (47-a) and (47-b) when interpreted under the knowledge that France has no king. Indeed, even though Abrusán & Szendrői do not provide a statistical analysis of the

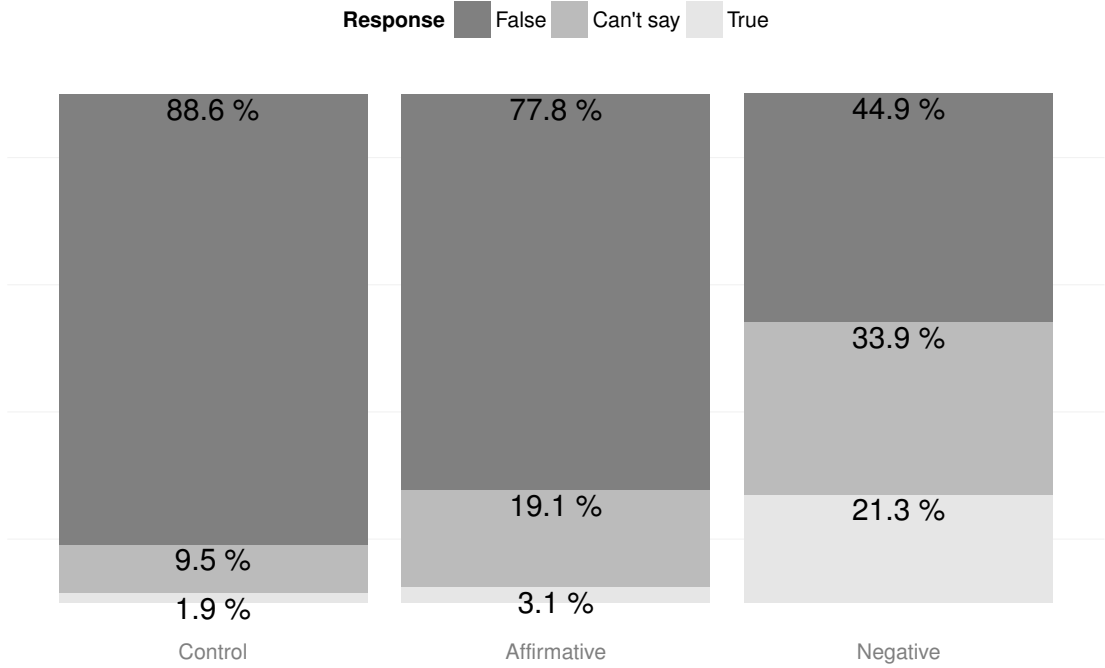


Figure 3.1: Proportions of responses in control and baseline conditions in Abrusán & Szendrői's experiment

*Can't say* responses, we can suspect a significant difference between the control condition and both baseline conditions. Even though we have a majority of *False* answers in each condition (a *relative* majority for baseline negative sentences) this does not mean that speakers who gave this answer did it straightforwardly. Keeping Schwarz's results in mind, we can imagine that the truth-value judgments for the baseline sentences are derived from an initially non-bivalent representation, contrary to the more homogeneous and more direct truth-value judgments for control sentences. In addition, the mean proportion of *True* responses for the baseline negative sentences suggests that this is an available truth-value judgment in this condition: the authors also indicate in a note that 4 participants chose the *true* option 84.4% of the time for these sentences. The existence of what is called *local accommodation* and which leads to a *true* reading of negative presuppositional sentences in situations of presupposition failure receives further evidence from (Chemla & Bott 2013)'s study, already mentioned in Footnote 3. I do not want to claim however that the *true* responses observed in (Abrusán & Szendrői 2013) necessarily result from a *process* of local accommodation. In discussing this question we should keep in mind the distinction between the *interpretation* which makes us judge negative presuppositional sentences

true in situations of presupposition failure and the *process* by which we arrive at this interpretation. In particular, local accommodation readings do not necessarily result from the insertion or movement of some linguistic material under the scope of negation, as is mentioned in (Beaver 2001)’s discussion of a meta-assertion operator. For instance, (Schlenker 2008)’s pragmatic account of presupposition deems *true* readings as primitive for sentences of this sort.<sup>5</sup>

As yet, no experimental study on presupposition seems to have presented its participants with non-bivalent truth-value judgments like *neither true nor false* or *false, but not simply false*. (Abrusán & Szendrői 2013) have grounded motivations for using *Can’t say* as a third option, but as a matter of fact it comes with a strong epistemic flavor which may have lowered their distribution, and the judgments of participants who chose this option are consequently hard to interpret. I do not know of any experiment looking at presuppositional conjunctions like (39-b) or presuppositional disjunctions like (39-a) either. Because the experiments discussed here exclusively looked at reference failures due to the definite article, one can but speculate that similar results would obtain for other presuppositional expression. Table 3.3 summarizes the results of these experiments, under the assumption that the *Can’t say* responses observed in the baseline conditions in (Abrusán & Szendrői 2013) were significant. “The NP VP” stands for affirmative sentences containing a definite descriptions which failed to refer in the target contexts, and “The NP not VP” stands for their negative counterparts.

	True	False	Else
The NP VP	×	✓	✓
The NP not VP	✓	✓	✓

Table 3.3: Available truth-value judgments for simple affirmative and negative sentences with unfulfilled presuppositions after experimental observations

These judgments are compatible with a trivalent treatment of presupposition that would not systematically predict non-bivalent truth-value judgments when evaluating a proposition of intermediate value. Affirmative sentences whose presuppositions are unfulfilled seem to trigger *False* and non-bivalent truth-value judgments, whereas negative sentences whose presuppositions are unfulfilled seem to trigger both *False*, *True* and non-bivalent truth-value judgments. This effect of polarity puts a constraint

---

<sup>5</sup>As the authors note, (Chemla & Bott 2013)’s results where the global accommodation readings were given faster than the local accommodation readings seem to argue against accounts that take *true* readings as primitive. See their discussion however for explanations making these accounts compatible with their results.

on the trivalent system that one would use to model presupposition, namely that it should make presuppositional propositions sensitive to negation. The absence of *true* truth-value judgments for simple sentences with unfulfilled presuppositions seems to exclude *glutty* accounts à la LP, and the presence of truth-value judgments which are neither *true* judgments nor *false* judgments seems to argue in favor of *gappy* accounts à la K3. Even though K3 alone does not account for the *true* judgments for negative sentences with unfulfilled presupposition, the previous chapter discussed the possibility of an accommodation operator that would make the right predictions.

### 3.3 Summary

The results from the experimental literature confirm two insights that were foreseen in the theoretical discussion of trivalent systems in the previous chapter. First, both vagueness and presupposition yield non-bivalent truth-value judgments. Second, the panel and maybe the nature of the various accessible truth-value judgments are different across the phenomena. Vagueness triggers a very broad range of truth-value judgments and calls for very “tolerant” trivalent systems, that would incorporate aspects of both LP and K3.

On the contrary, presupposition triggers a narrower range of truth-value judgments and calls for a “stricter” trivalent system which should in addition make use of the third value in a way sensitive to the presence of negation. This last contrast between vagueness and presupposition is important: the effect of negation seems to disappear with vagueness but not with presupposition. From a perspective of unification, this may be the strongest argument for a different treatment of each type of sentence. Non truth-functional systems such as supervaluationism may find the resources to do so while keeping a unique third value in the notion of penumbral connections, but truth-functional systems seem to have no choice but to assign specific non-bivalent truth-values to the propositions expressing vague and presuppositional sentences evaluated in critical situations.

But accounting both for the truth-value judgments associated with vagueness and for the truth-value judgments associated with presuppositions is not the only challenge of a unified account of these phenomena. A unified account should also make some predictions concerning hybrid sentences, that is to say sentences such as (50-a) or (50-b) that involve both vague and presuppositional expressions.

- (50) a. The amplifiers have stopped being loud
- b. The amplifiers are loud and they have stopped buzzing

To my knowledge, no theory considers such sentences and therefore no theory makes any prediction regarding the semantic status of (50-a) or (50-b). The next chapter presents a unifying truth-functional 5-valued system, drawing on (Cobreros et al. 2012)’s trivalent system for vagueness which precisely incorporates the assets of both LP and K3.



## Chapter 4

### ST5: a 5-Valued System

In order to deal with *truth-value judgments* concerning vagueness and presupposition, I developed a 5-valued system, ST5.<sup>1</sup> I started from the position that we observe *conflicting* judgments for vague sentences as well as for presuppositional sentences in specific situations. For instance, consider the presuppositional sentence (51):<sup>2</sup>

(51) The amplifiers have stopped buzzing.

If I'm told (51) and I know that, in fact, the amplifiers have never buzzed, I can say that (51) is *both false and not false*: it is *false* because the amplifiers were *not* buzzing before, and it is *not false* because if (51) were false, it would mean that the amplifiers *were* buzzing before. Similarly, consider the vague sentence (52), involving the vague adjective *loud*:

(52) The amplifiers are loud.

If I'm told (52) and I find the volume of the amplifiers to be neither clearly loud nor clearly not loud, I can say that (52) is *both true and false*: it is true to some extent, because the amplifiers are not clearly not loud, but it is false to some extent too, because they're not clearly loud either.<sup>3</sup>

My aim here will be to offer a semantics that assigns *logical truth values* to propositions involving vague and presuppositional expressions on the basis of which one could

---

<sup>1</sup>This chapter is partly based on an article that was published under the name “ST5: A 5-Valued Logic for Truth-Value Judgments Involving Vagueness and Presuppositions” in the journal *Pristine Perspectives on Logic, Language, and Computation*.

<sup>2</sup>Aspectual verbs such as *stop* are well-known to trigger a presupposition. See for instance the article “Presupposition” in the Stanford Encyclopedia of Philosophy (Beaver & Geurts 2013).

<sup>3</sup>Serchuk et al. (Serchuk et al. 2011) conducted several experiments revealing this apparent contradictory characteristic of truth-value judgments for vagueness.

correctly predict the *truth-value judgments* of speakers in regular *and* conflicting-judgment contexts. In Sect. 4.1, I begin by reviewing truth-value judgments that we find for positive and negative counterparts of sentences involving vague expressions and sentences involving presuppositional expressions. Section 4.2 presents the 3-valued ST system (Cobreros et al. forthcoming<sup>b</sup>), which has been developed for vagueness and which offers a natural way of accounting for the conflicting truth-value judgments to which vagueness gives rise. I then consider a 5-valued extension of this system, which I call ST5, in order to incorporate presuppositional expressions. Finally in Sect. 4.3, I consider the interactions between vagueness and presupposition, by looking at sentences that involve both vague and presuppositional expressions (hybrid sentences). I propose a semantics for presuppositional sentences in ST5 that makes predictions for hybrid sentences and for sentences with iteratively embedded presuppositional expressions.

## 4.1 Truth-Value Judgments

By a *truth-value judgment* I here mean any position that a speaker can have toward the *truth* or the *falsity* of a sentence. My use of this notion then refers to the set of combinations of *true* and *false* closed under *not*, *and*, *(n)or*, *both* and *(n)either*.<sup>4</sup>

Each element of this set is a truth-value judgment. It is clear that, as truth-value judgments, some of the elements in the set are so-to-speak “regular”: speakers often judge sentences *true*, *false*, *not true* or *not false*. But other elements are far less “regular” (*neither true nor false*) and some even sound contradictory: *both true and false*, *both true and not true*, *both false and not false* for instance.<sup>5</sup> Yet, I claim that speakers can use these elements to qualify some sentences. That is to say, I claim that speakers can exhibit apparently *conflicting* truth-value judgments. Even though some dialetheists, such as Priest (Priest 2006), endorse the view that there are true contradictions, Lewis (Lewis 1982) for instance proposed to see underlying ambiguity in judgments of this kind.<sup>6</sup>

In the next two subsections, I present some evidence that speakers have access to these kinds of judgments concerning vagueness and presupposition. The account I will eventually give for this relies on a notion of assertoric *ambiguity* developed in

---

<sup>4</sup>Importantly, the set of *truth-value judgments* is to be distinguished from the set of *logical values* that a system assigns to propositions. There is no necessary one-to-one correspondence between their elements; and the system I will eventually propose exhibits no such correspondence.

<sup>5</sup>Note the italics that distinguish between judging a sentence both *false* and *not false* and judging a sentence *both false and not false*.

<sup>6</sup>See Kooi & Tamminga (Kooi & Tamminga 2013) for support for Lewis’ view contra Priest.

the 3-valued logic ST (Cobreros et al. forthcoming<sup>b</sup>). So far, there have been few experiments exploring the *truth-value judgments* of speakers concerning vagueness or presupposition, I will therefore rely on indirect evidence that speakers have access to conflicting truth-value judgments in the cases of vagueness and of presupposition.

## 4.2 ST5

### 4.2.1 The Original 3-Valued ST System

ST is a trivalent logical system developed to deal with vague predicates (Cobreros et al. forthcoming<sup>b</sup>), and more specifically to account for conflicting judgments such as “X is tall and not tall”.<sup>7</sup> There are two reasons for which I base my 5-valued system on ST: first, ST already comes with an account for vagueness. Hence only half of the works remains to be done. Second, ST comes with a notion of *assertoric ambiguity* that leads to a nice explanation for our conflicting judgments.

#### 4.2.1.1 Two Notions of Satisfaction

Let’s consider as our language  $\mathcal{L}$  a non-quantified fragment of monadic first-order logic such that:

**Definition 4.2.1** (Syntax). *i. For any predicate  $P \in \mathcal{L}$  and any individual name  $a \in \mathcal{L}$ ,  $Pa$  is a well-formed formula (wff).*

*ii. For any wff  $\phi$ ,  $\neg\phi$  is a wff.*

*iii. For any  $\phi$  and  $\psi$  such that  $\phi$  and  $\psi$  are wff,  $[\phi \wedge \psi]$ ,  $[\phi \vee \psi]$  and  $[\phi \rightarrow \psi]$  are wff. Nothing else is a wff.*

$\mathcal{M}$  consists of a non-empty domain of individuals  $\mathcal{D}$  and an interpretation function  $\mathcal{I}$  such that:

**Definition 4.2.2** (Semantics). *i. For any predicate  $P \in \mathcal{L}$  and any individual name  $a \in \mathcal{L}$ ,  $\mathcal{I}(Pa) = \frac{1}{2}$  iff  $\mathcal{I}(a)$  is a borderline case for  $\mathcal{I}(P)$ ,  $\mathcal{I}(Pa) \in \{0, 1\}$  otherwise.*

---

<sup>7</sup>ST is a built-in 3-valued version of TCS (Cobreros et al. 2012), which assumed bivalent extensions for vague predicates on which it built their trivalent extensions. As I present it here, ST seems to be committed to the existence of a sharp boundary between eg. clearly tall men and borderline tall men, which might sound unrealistic. This point is related to the question of higher-order vagueness, which is much discussed in the literature on vagueness. A discussion of higher-order vagueness goes far beyond the scope of this paper. I will therefore just endorse the assumption that vagueness defines a well defined trivalent extension in the rest of the paper, with no further justification.

ii. For any wff  $\phi$ ,  $\mathcal{I}(\neg\phi) = 1 - \mathcal{I}(\phi)$ .

iii. For two wff  $\phi$  and  $\psi$ ,  $\mathcal{I}(\phi \wedge \psi) = \min(\mathcal{I}(\phi), \mathcal{I}(\psi))$ ,  
 $\mathcal{I}(\phi \vee \psi) = \max(\mathcal{I}(\phi), \mathcal{I}(\psi))$  and  $\mathcal{I}(\phi \rightarrow \psi) = \mathcal{I}(\neg\phi \vee \psi)$

The system ST owes its name to the definition of two notions of satisfaction:<sup>8</sup>

**Definition 4.2.3** (Strict and Tolerant Satisfaction). *For any model  $\mathcal{M}$  whose interpretation function is  $\mathcal{I}$ ,*

**Strict satisfaction:**  $\mathcal{M} \models^s \phi$  iff  $\mathcal{I}(\phi) = 1$

**Tolerant satisfaction:**  $\mathcal{M} \models^t \phi$  iff  $\mathcal{I}(\phi) \geq \frac{1}{2}$

Now, imagine  $a$  is the name of a borderline case for  $\mathcal{I}(P)$ . We have  $\mathcal{I}(Pa) = \frac{1}{2}$  and  $\mathcal{I}(\neg Pa) = 1 - \frac{1}{2} = \frac{1}{2}$ . Hence, we get  $\mathcal{I}(Pa \wedge \neg Pa) = \min(\frac{1}{2}, \frac{1}{2}) = \frac{1}{2}$  and  $\mathcal{I}(\neg(Pa \vee \neg Pa)) = 1 - \max(\frac{1}{2}, \frac{1}{2}) = 1 - \frac{1}{2} = \frac{1}{2}$ . This leads us to:

- i.  $\mathcal{M} \models^t Pa$  but  $\mathcal{M} \not\models^s Pa$
- ii.  $\mathcal{M} \models^t \neg Pa$  but  $\mathcal{M} \not\models^s \neg Pa$
- iii.  $\mathcal{M} \models^t Pa \wedge \neg Pa$  but  $\mathcal{M} \not\models^s Pa \wedge \neg Pa$
- iv.  $\mathcal{M} \models^t \neg(Pa \vee \neg Pa)$  but  $\mathcal{M} \not\models^s \neg(Pa \vee \neg Pa)$

With  $P$  standing for “is tall” and  $a$  standing for borderline-tall “John”, what we have is that none of “John is tall”, “John is not tall”, “John is tall and not tall” and “John is neither tall nor not tall” is *strictly* satisfied,<sup>9</sup> but all of them are *tolerantly* satisfied. Cobreros et al. propose to account for the results of Alxatib & Pelletier (Alxatib & Pelletier 2011) by assuming that speakers can assert vague sentences either strictly or tolerantly. To this, I add the following bridge principles:<sup>10</sup>

<sup>8</sup>See (Cobreros et al. forthcoming<sup>b</sup>) for a discussion of inference rules in this system.

<sup>9</sup>Here, I regard *neither... nor...* as the negation of a disjunction

<sup>10</sup>In formulating these bridge principles, I use  $\mathcal{M}$  as a free variable universally quantified over the models compatible with the beliefs of the speaker. In particular, a speaker knows a proposition if all the models compatible with his beliefs assign 1 to this proposition. The question of whether a speaker can be said to know a proposition if all the models compatible with his beliefs assign *either* 1 or  $\frac{1}{2}$  to this proposition finds some answers in Sect. 4.3.

**Principle 1** (Truth-Value Judgments). *One can judge a proposition  $\phi$ ...*

1. “true” if  $\mathcal{M} \models^t \phi$
2. “false” if  $\mathcal{M} \models^t \neg\phi$
3. “not true” if  $\mathcal{M} \not\models^s \phi$
4. “not false” if  $\mathcal{M} \not\models^s \neg\phi$
5. “both true and false” if 1 and 2.
6. “neither true nor false” if 3 and 4.
7. “both true and not true” if 1 and 3.
8. “both false and not false” if 2 and 4.

It is straightforward that, for borderline-tall John, “John is tall” as well as “John is not tall” can be judged *both true and false* and *neither true nor false*.

#### 4.2.1.2 No Room for Presupposition

Now, looking at the bridge principles, it would be ideal if we could add presuppositional propositions  $\phi$  to our language in such a way that, *when the presupposition of  $\phi$  is unfulfilled*:

1.  $\mathcal{M} \models^t \neg\phi$  (so that a speaker can judge  $\phi$  false)
2.  $\mathcal{M} \not\models^s \neg\phi$  (so that a speaker can judge  $\phi$  not false)
3.  $\mathcal{M} \not\models^t \phi$  (so that a speaker cannot judge  $\phi$  true)

But the only way in ST to have 1. and 2. is for  $\phi$  to get the value  $\frac{1}{2}$ , and then we would have  $\mathcal{M} \models^t \phi$  and a speaker could judge  $\phi$  true as well. More specifically, ST has the following property (see (Cobreros et al. forthcomingb)):

**Lemma 4.2.4** (Duality in ST). *For any wff  $\phi$ ,  $\mathcal{M} \models^{s/t} \phi$  iff  $\mathcal{M} \not\models^{t/s} \neg\phi$*

The solution I propose consists in breaking this duality by adding two logical values to the system: propositions that get one of these two extra values will obey the three constraints above, but propositions that get one of the three initial values will still present the equivalence noted in Lemma 4.2.4.

#### 4.2.2 Why Exclude Alternative Systems

One might be tempted of adding a single extra value to  $\{0, \frac{1}{2}, 1\}$ . There are two substantial ways of doing this: either making the four values totally ordered, or making them partially ordered.

Under the total order alternative, the semantics for  $\neg$  would force us to consider a system containing a value which would correspond to 1 minus the extra value.

But given that our initial three-valued set was  $\{0, \frac{1}{2}, 1\}$ , adding a fourth value would therefore lead us to add a fifth value, leaving us with the basis of the ST5 system. Another option would be to abandon the initial value  $\frac{1}{2}$  and to consider a set of four values  $\{0, \mathcal{P}, \mathcal{V}, 1\}$ , where  $\mathcal{P}$  would be a value assigned to propositions describing situations of presupposition failure and  $\mathcal{V}$  a value assigned to propositions describing borderline cases. In addition, we would have  $\mathcal{P} = 1 - \mathcal{V}$  in order to fit the semantics for  $\neg$ . But there is a problem with this solution, and it is precisely related to negation. Imagine you have a proposition  $\phi_P$  describing a case of presupposition failure and a proposition  $\psi_V$  describing a borderline case: as such,  $\phi_P$  gets the value  $\mathcal{P}$  and  $\psi_V$  gets the value  $\mathcal{V}$ . But now  $\neg\phi_P$  gets the value  $1 - \mathcal{P} = \mathcal{V}$ , which is the value of  $\psi_V$ . And conversely,  $\neg\psi_V$  gets the value  $1 - \mathcal{V} = \mathcal{P}$ , which is the value of  $\phi_P$ . This has two unwelcome effects: first it predicts that we should observe the same truth judgments for negative counterparts of presuppositional sentences used in case of presupposition failure and for vague sentences used to describe borderline cases; second it predicts that we should observe different truth judgments for affirmative and negative counterparts of vague sentences. These predictions seem unreasonable enough to exclude this solution.

Under the partial order alternative, we have a set of four values  $\{0, \mathcal{P}, \mathcal{V}, 1\}$  where  $0 < \mathcal{P} < 1$  and  $0 < \mathcal{V} < 1$ . We would then need to adapt the semantics of our connectives to a partial ordered lattice: negation could semantically contribute as a symmetric operator (ie. for  $\mathcal{I}(\phi) = 1$ ,  $\mathcal{I}(\neg\phi) = 0$ , for  $\mathcal{I}(\phi) = 0$ ,  $\mathcal{I}(\neg\phi) = 1$ , for  $\mathcal{I}(\phi) = \mathcal{V}$ ,  $\mathcal{I}(\neg\phi) = \mathcal{V}$  and for  $\mathcal{I}(\phi) = \mathcal{P}$ ,  $\mathcal{I}(\neg\phi) = \mathcal{P}$ ), and conjunction and disjunction could respectively semantically contribute as the greatest lower bound and as the least upper bound.<sup>11</sup> But note that in this system, a proposition describing a case of presupposition failure would receive the same value as its negation: we would therefore have to say something more to explain the asymmetry in our truth judgments for presupposition. One solution would be to consider a *partially ordered five-valued* set  $\{0, \mathcal{P}^0, \mathcal{P}^1, \mathcal{V}, 1\}$  such that  $0 < \mathcal{V} < 1$  and  $0 < \mathcal{P}^0 < \mathcal{P}^1 < 1$ : positive propositions describing situations of presupposition failure would have the value  $\mathcal{P}^0$  and their negation would have the value  $\mathcal{P}^1$ . Whether one readily adds a fifth value or not doesn't solve a major problem of such partially ordered systems. Consider the conjunction and the disjunction in (53).

- (53) a. The amplifiers are loud and they have stopped buzzing

---

<sup>11</sup>A reviewer argued that there are other ways of defining the connectives that might be as legitimate as the standard Dunn-Belnap definition. See Chapter 6 for an investigation of four-valued systems.

- b. The amplifiers are loud or they have stopped buzzing

With either the four-valued or the five-valued version of a partially ordered lattice, in situations where the amplifiers are borderline-loud and have never buzzed, (53-a) would express the conjunction of two propositions that would receive *non-ordered values* and (53-b) would express their disjunction. With conjunction being defined as the greatest lower bound and disjunction being defined as the least upper bound, the proposition expressed by (53-a) would get the value 0 and the proposition expressed by (53-b) would get the value 1. Such a system would therefore predict a pure *false* judgment for (53-a) and a pure *true* judgment for (53-b) in those situations, which clearly goes against our intuitions.

One might finally consider a system with still partially ordered values but such that the greatest lower bound and the least upper bound of the values for vagueness and presupposition are not 0 and 1. With  $\mathcal{E}^0$  and  $\mathcal{E}^1$  the new  $\mathcal{E}$ extra values, we would have a set of six values  $\{0, \mathcal{E}^0, \mathcal{V}, \mathcal{P}, \mathcal{E}^1, 1\}$  such that  $0 < \mathcal{E}^0 < \mathcal{V} < \mathcal{E}^1 < 1$  and  $0 < \mathcal{E}^0 < \mathcal{P} < \mathcal{E}^1 < 1$ . In this system, vagueness and presupposition seem ontologically well distinguished ( $\mathcal{P}$  and  $\mathcal{V}$  are not ordered with each other), and in critical situations, the conjunction expressed in (53-a) would get the value  $\mathcal{E}^0$  (the greatest lower bound of  $\mathcal{P}$  and  $\mathcal{V}$ ) and the disjunction expressed in (53-b) would get the value  $\mathcal{E}^1$  (the least upper bound of  $\mathcal{P}$  and  $\mathcal{V}$ ). But this brings the question of what  $\mathcal{E}^0$  and  $\mathcal{E}^1$  actually represent. If their existence is motivated only by the existence of conjunctions and disjunctions of propositions describing borderline cases and propositions describing cases of situation failure, this seems an important price to pay. In addition, the six-valued system I considered here is based on a partially ordered four-valued system which doesn't distinguish between affirmative and negative presuppositional sentences in cases of presupposition failure: a partially ordered *seven*-valued system might then be more adequate.

Eventually, one remains with the alternative of a five totally ordered values system where each value has a clear ontological status. This is what I will explore now in the next section.

### 4.2.3 The ST5 System

In ST, we had three values:  $\{0, \mathcal{V} = \frac{1}{2}, 1\}$ , and vague predications on borderline cases got the value  $\mathcal{V}$ . Now, in ST5, we add two more values,  $\mathcal{P}^0$  and  $\mathcal{P}^1$ , such that:  $0 < \mathcal{P}^0 < \mathcal{V} < \mathcal{P}^1 < 1$  and such that  $\mathcal{P}^0 = 1 - \mathcal{P}^1$ . The syntax and the semantics of

ST remain unchanged in this extended system, as well as Definition 4.2.3 of tolerant and strict satisfactions. By this simple addition, we obtain the following:

**Lemma 4.2.5** (Duality lost).

- For any proposition  $\phi$  such that  $\mathcal{I}(\phi) = \mathcal{P}^0$ :
  - i.  $\mathcal{M} \not\models^t \phi$  and  $\mathcal{M} \not\models^s \phi$  since  $\mathcal{P}^0 < \frac{1}{2} < 1$ .
  - ii.  $\mathcal{M} \models^t \neg\phi$  but  $\mathcal{M} \not\models^s \neg\phi$  since  $1 - \mathcal{P}^0 = \mathcal{P}^1$  and  $\mathcal{P}^1 \geq \frac{1}{2}$  but  $\mathcal{P}^1 < 1$ .
- For any proposition  $\phi$  such that  $\mathcal{I}(\phi) = \mathcal{P}^1$ :
  - i.  $\mathcal{M} \models^t \phi$  but  $\mathcal{M} \not\models^s \phi$  since  $\mathcal{P}^1 \geq \frac{1}{2}$  but  $\mathcal{P}^1 < 1$ .
  - ii.  $\mathcal{M} \not\models^t \neg\phi$  and  $\mathcal{M} \not\models^s \neg\phi$  since  $1 - \mathcal{P}^1 = \mathcal{P}^0$  and  $\mathcal{P}^0 < \frac{1}{2} < 1$ .

Given that we now have propositions  $\phi$  for which  $\mathcal{M} \not\models^s \neg\phi$  but  $\mathcal{M} \not\models^t \phi$  (propositions of value  $\mathcal{P}^0$ ), Lemma 4.2.4 no longer holds in ST5. Nonetheless, the following holds in ST as well as in ST5:

**Lemma 4.2.6** (Entailment). For any wff  $\phi$ ,  $\mathcal{M} \models^s \phi$  entails  $\mathcal{M} \models^t \phi$ .

Now let us stipulate (54).

- (54) Any simple positive proposition  $\phi$  whose presupposition is *unfulfilled* gets the value  $\mathcal{P}^0$ .

For instance, with  $\phi$  standing for (51), repeated in (55-a), its truth-value would be defined such that (55-b) holds.

- (55) a. The amplifiers have stopped buzzing.  
 b.  $\mathcal{I}(\phi) = \mathcal{P}^0$  iff the amplifiers have never buzzed, 0 iff they still, 1 iff they have stopped.

From the semantics of the negation operator in ST5, it follows that the negation of (51) would get the value  $\mathcal{P}^1$ . The bridge principles thus predict the following, as desired:<sup>12</sup>

- i. One can judge  $\phi$  *both false and not false* ( $\mathcal{M} \models^t \neg\phi$  but  $\mathcal{M} \not\models^s \neg\phi$ )
- ii. One can judge  $\phi$  *neither true nor false* ( $\mathcal{M} \not\models^s \phi$  and  $\mathcal{M} \not\models^s \neg\phi$ )

---

<sup>12</sup>Recall that we have  $\neg\neg\phi \equiv \phi$ .



- iii. One can judge  $\neg\phi$  both true and not true ( $\mathcal{M} \models^t \neg\phi$  but  $\mathcal{M} \not\models^s \neg\phi$ )
- iv. One can judge  $\neg\phi$  neither true nor false ( $\mathcal{M} \not\models^s \neg\phi$  and  $\mathcal{M} \not\models^s \neg\neg\phi$ )
- v. One cannot judge  $\phi$  true ( $\mathcal{M} \not\models^t \phi$ )

On the basis of the 5 possible truth-values that propositions can receive depending on the model in which they are interpreted, we are now able to formalize the notions of presuppositional, vague and hybrid propositions.

**Definition 4.2.7** (Bivalent Propositions in ST5). *A proposition  $\phi \in \mathcal{L}$  is bivalent with respect to a set of interpretation functions  $\mathcal{J}$  if, for all  $\mathcal{I} \in \mathcal{J}$ ,  $\mathcal{I}(\phi) \in \{0, 1\}$ .*

**Definition 4.2.8** (Vague Propositions in ST5). *A proposition  $\phi \in \mathcal{L}$  is vague with respect to a set of interpretation functions  $\mathcal{J}$  if there is a  $\mathcal{I} \in \mathcal{J}$  such that  $\mathcal{I}(\phi) = \mathcal{V}$ .*

**Definition 4.2.9** (Presuppositional Propositions in ST5). *A proposition  $\phi \in \mathcal{L}$  is presuppositional with respect to a set of interpretation functions  $\mathcal{J}$  if there is a  $\mathcal{I} \in \mathcal{J}$  such that  $\mathcal{I}(\phi) \in \{\mathcal{P}^0, \mathcal{P}^1\}$ .*

**Definition 4.2.10** (Hybrid Propositions in ST5). *A proposition  $\phi \in \mathcal{L}$  is hybrid with respect to a set of interpretation functions  $\mathcal{J}$  if  $\phi$  is both presuppositional and vague with respect to  $\mathcal{J}$ .*

With restrictions to the models that we consider (and more specifically to the *interpretation functions* that we consider), these distinctions are assumed to fit the distinction in natural language between sentences that are vague, sentences that are presuppositional and sentences that are hybrid.

## 4.3 Hybrid Sentences

### 4.3.1 Conjunctions, Disjunctions and Implications in ST5

#### 4.3.1.1 An Example

Because ST5 deals with totally ordered values and defines its connectives in terms of *min* and *max*, it naturally makes predictions for conjunctions, disjunctions and implications combining vague and presuppositional propositions. Consider (50-b) repeated here that conjoins a vague sentence and a presuppositional sentence:

(56) The amplifiers are loud and they have stopped buzzing

Given that the amplifiers have *never* buzzed, if their volume is somewhere between clearly loud and clearly not loud, the first conjunct gets the value  $\mathcal{V}$  and the second conjunct gets the value  $\mathcal{P}^0$ . Therefore in these circumstances, the whole proposition gets the value  $\min(\mathcal{V}, \mathcal{P}^0) = \mathcal{P}^0$ : it is judged *both false and not false* (for the amplifiers were *not* buzzing before), and it's not judged *true*.

Here is a table summarizing the predictions of ST5 for hybrid conjunctions and disjunctions when the amplifiers (abbreviated as *A*) are borderline-loud and have never buzzed:

Proposition	Value	Judgment
A are loud	$\mathcal{V}$	Both true and false
A are not loud	$\mathcal{V}$	Both true and false
A have stopped buzzing	$\mathcal{P}^0$	Both false and not false
A have not stopped buzzing	$\mathcal{P}^1$	Both true and not true
A are loud & have stopped buzzing	$\mathcal{P}^0$	Both false and not false
A are not loud & have stopped buzzing	$\mathcal{P}^0$	Both false and not false
A are loud & have not stopped buzzing	$\mathcal{V}$	Both true and false
A are not loud & have not stopped buzzing	$\mathcal{V}$	Both true and false
A are loud or have stopped buzzing	$\mathcal{V}$	Both true and false
A are not loud or have stopped buzzing	$\mathcal{V}$	Both true and false
A are loud or have not stopped buzzing	$\mathcal{P}^1$	Both true and not true
A are not loud or have not stopped buzzing	$\mathcal{P}^1$	Both true and not true

Figure 4.1: Predictions of ST5 for hybrid conjunctions and disjunctions

#### 4.3.1.2 Left-Right Asymmetries

In view of these predictions, a word is in order about the left-right asymmetry of presupposition. It's been claimed since at least Stalnaker (Stalnaker 1974) and Heim (Heim 1983) that sentences such as (57-a) carry a presupposition while the corresponding reversed sentence (57-b) does not:

- (57)    a.    The amplifiers have stopped buzzing and they were buzzing before.  
           b.    The amplifiers were buzzing before and they have stopped buzzing.

In ST5, conjunctions are totally symmetric and (57-a) and (57-b) will get the same value when the amplifiers never buzzed:  $\min(\mathcal{P}^0, 0) = \min(0, \mathcal{P}^0) = 0$ . Therefore we predict that both (57-a) and (57-b) will be judged merely *false* when we know that amplifiers have never buzzed. It's unclear what *truth-value judgments* speakers would actually give for (57-a) and (57-b). We should, though, distinguish between

the question whether (57-a) as well as (57-b) should be judged *false* or whether they should come with different *truth-value judgments*, and the rather clear intuition that (57-b) is *utterable* in a broader range of conditions than (57-a).<sup>13</sup>

(Schlenker 2008) pointed out that this asymmetry in conditions of use could be related to a more general property of conjunctions. Indeed, the contrast we observed between (57-a) (which “sounds weird”) and (57-b) is somehow similar to the one we observe between (58-a) (which “sounds weird” too) and (58-b):<sup>14</sup>

- (58)    a.    John lives in Paris and he resides in France.  
           b.    John resides in France and he lives in Paris.

Schlenker therefore proposes a general constraint that has the effect of ruling out conjunctions where the first conjunct entails the second one. On the point of view I am adopting, the way to translate (57-a) into ST5 is as a proposition that we could schematize as  $\phi_p \wedge p$ .  $p$  is to be understood as a proposition expressing the presuppositional part of  $\phi_p$ . Therefore in the following discussion, we will only exclude models where the interpretation function does not assign  $\mathcal{P}^0$  to  $\phi_p$  when it assigns 0 to  $p$ . Schlenker’s principle rules out the expression of conjunctions of this form for their left part entails their right part. In classical logic, Schlenker’s condition to rule out the expression of conjunctions  $\psi \rightarrow \delta$  can be stated as  $\mathcal{M} \models \psi \rightarrow \delta$ . In ST5 as in classical logic, we have  $\psi \rightarrow \delta \equiv \neg\psi \vee \delta$ . Thus, for the classical notion of satisfaction as well as for the tolerant and the strict notions of satisfaction, we have  $\mathcal{M} \models \psi \rightarrow \delta$  if and only if  $\mathcal{M} \models \neg\psi \vee \delta$  if and only if  $\mathcal{M} \models \neg\psi$  or  $\mathcal{M} \models \delta$ . Because we think that the rejection of (57-a), that we modeled as  $\phi_p \wedge p$ , is due to the violation of Schlenker’s principle, we need to determine if a notion of satisfaction, strict or tolerant, makes  $\mathcal{M} \models \neg\phi_p$  or  $\mathcal{M} \models p$  hold for any model (compatible with  $p$  being the presuppositional part of  $\phi_p$ ). Table 4.1 dresses the list of the possible values of  $p$  and  $\phi_p$  in these models and states how Schlenker’s principle accordingly does or

---

<sup>13</sup>(i-b) uttered in the context described in (i-a) is an example of a sentence which is true and yet not utterable if it is of importance whether all of the amplifiers are buzzing and if it is clear that the utterer is in possession of this information. This is usually explained by the fact that (i-b) conveys a scalar implicature, namely that not all of the amplifiers are buzzing, which enters in contradiction with the context in (i-a). To this extent, pragmatic factors can thus obscure our truth-value judgments.

- (i)      a.      *Context: all the amplifiers in the room are buzzing.*  
           b.      Some of these amplifiers are buzzing.

<sup>14</sup>To insist on the need of distinguishing between giving a *non-classical truth-value* judgment for a sentence and feeling this sentence is “weird”, note that you will judge *both* (58-a) and (58-b) completely false if you know John lives in London, but still regard (58-a) as weird.

does not rule out the expression of  $\phi_p \wedge p$ , depending on the notion of satisfaction we consider.

$\mathcal{I}(p)$	$\mathcal{I}(\phi_p)$	$\mathcal{M} \models^t p$ or $\mathcal{M} \models^t \neg\phi_p$	$\mathcal{M} \models^s p$ or $\mathcal{M} \models^s \neg\phi_p$
$0/\mathcal{P}^0$	$\mathcal{P}^0$	Yes ( $\mathcal{I}(\neg\phi_p) \geq \mathcal{V}$ )	No ( $\mathcal{I}(p) < 1$ and $\mathcal{I}(\neg\phi_p) < 1$ )
$\mathcal{V}/\mathcal{P}^1$	$\mathcal{P}^0/\mathcal{P}^1$	Yes ( $\mathcal{I}(p) \geq \mathcal{V}$ )	No ( $\mathcal{I}(p) < 1$ and $\mathcal{I}(\neg\phi_p) < 1$ )
1	$0/\mathcal{P}^0/\mathcal{V}/\mathcal{P}^1/1$	Yes ( $\mathcal{I}(p) \geq \mathcal{V}$ )	Yes ( $\mathcal{I}(p) = 1$ )
$\phi_p \wedge p$ ruled out?		Yes	No

Table 4.1: Using the strict or the tolerant notion of satisfaction to adapt Schlenker’s principle in ST5 yield different acceptance diagnoses for  $\phi_p \wedge p$ .

Talbe 4.1 suggests that the correct way to implement Schlenker’s principle is by using the notion of tolerant satisfaction. One should note moreover that if the only constraint on the use of (57-a) were for the presupposition of its left conjunct to be fulfilled, then (57-a) should sound totally fine in cases where (57-b) is known to be true, but this is not the case: if we know that the amplifiers used to buzz, (57-a) “sounds weird” in a way in which (57-b) does not. To this extent, the strength of the contrast between (57-a) and (57-b) should not be raised in favor of the view that (57-a) is presuppositional while (57-b) is not: as a matter of fact, we can’t use our judgments on (57-a) to clearly distinguish between cases where the presupposition of its left conjunct is fulfilled from cases where it is not.<sup>15</sup>

If one thinks that, nonetheless, these sentences should receive different *truth-value judgments*, a possibility is to revise the semantics of the conjunction operator so that it gives the value  $\mathcal{P}^0$  to a conjunction whenever it has a proposition of value  $\mathcal{P}^0$  on its left: with such a semantics, and contrary to the option above,  $\phi_p \wedge p$  (the formalization we proposed for (57-a)) is presuppositional with respect to any set of interpretation functions treating  $p$  as the presuppositional part  $\phi_p$ , since  $\phi_p \wedge p$  gets the value  $\mathcal{P}^0$  in at least one of the models under consideration. As (Fox 2008) and (George 2008) point out, one can extend this kind of considerations to all the connectives in the system by resorting to a unifying principle in the spirit of the one

<sup>15</sup>However, as (Fox 2010) suggests, sentences like (i) where the right conjunct is more informative than the presupposition of the left conjunct are not ruled out by Schlenker’s principle.

(i) John is unaware that he is sick and he has cancer.

It seems to me that if I heard (i) in situations where I know that John is actually healthy, I would judge (i) plainly false.

proposed by Schlenker. However it is not clear whether disjunctions and implications show the same asymmetry (see (59)), and so whether one should or not revise the semantics of the connectives in the system.

- (59)
- a. The amplifiers have stopped buzzing or they were not buzzing before.
  - b. The amplifiers were not buzzing before or they have stopped buzzing.
  - c. The amplifiers have stopped buzzing, if they were buzzing before.
  - d. If the amplifiers were buzzing before, they have stopped buzzing.

### 4.3.2 Deriving Presuppositions From Words

ST5 comes with vague and presuppositional propositions, which express simple vague and presuppositional sentences of natural language. The last section showed that ST5 can also deal with complex sentences (ie. concatenations of simple sentences with connectives) and more specifically that it can derive the presuppositions of these complex sentences from its parts. But so far, we have only considered simple sentences whose presuppositions were merely fulfilled or unfulfilled (see the stipulation in (54)). That is to say, we have only considered situations in which presuppositions could be expressed by propositions receiving a bivalent value (0 or 1). But as it turns out, some presuppositions are to be expressed by propositions that themselves involve vague and presuppositional expressions. Think of sentences such as (60-a) or (60-b) whose presuppositions can respectively be expressed by (60-a-i) and (60-b-i) (which is the repetition of (51)).

- (60)
- a. The amplifiers have stopped being loud
    - (i) The amplifiers were loud
  - b. John knows that the amplifiers have stopped buzzing
    - (i) The amplifiers have stopped buzzing

By hypothesis, in situations where the amplifiers were borderline loud and have never buzzed, (60-a-i) expresses a proposition that gets the value  $\mathcal{V}$  and (60-b-i) expresses a proposition that gets the value  $\mathcal{P}^0$ . To this extent, what values should receive the propositions expressing (60-a) and (60-b) in these situations? More generally, what effect does a presupposition with value  $\mathcal{V}$  or  $\mathcal{P}^0$  have on the value of the proposition as a whole? This section answers this question by considering how the presuppositions of simple sentences are derived from their lexical parts.

#### 4.3.2.1 Presuppositional Expressions in Traditional Truth-Conditional Semantics

It makes sense to say that (60-a-i) expresses the presupposition of (60-a) and that (60-b-i) expresses the presupposition of (60-b) because the verbs *stop* and *know* generate presuppositions on the basis of their complement in a systematic way. Indeed, sentences of the form *X has stopped V-ing* generate the presupposition that *X* used to *V*, and sentences of the form *X knows that S* generate the presupposition that *S*. In truth-conditional semantics, there is by now a traditional approach of presupposition in terms of partial functions. (Heim & Kratzer 1998)’s formalization offers a formal representation of presuppositions as domain conditions on the functions corresponding to the interpretation of presuppositional sentences.<sup>16</sup> On this view, to say that a sentence *S* has truth value *n* is to say that  $\llbracket S \rrbracket$  yields *n* for the actual world. For any *S*,  $\llbracket S \rrbracket$  is a function of the form  $\lambda w_s : P. Q$ , where *P* and *Q* are statements referring to *w*, whose domain is  $\{w : P\}$  and that, for any *w* in its domain, return 1 if *Q* and 0 otherwise. On this basis, I will say that we derive a presupposition failure when the actual world is not in the domain of  $\llbracket S \rrbracket$ : the resulting proposition fails to have a bivalent truth-value.

To illustrate this, consider the possible lexical entries for the presuppositional expressions *know* and *stop* in (61-a) and (61-b), where the respective presuppositions that they generate consist in the propositions between  $\lambda w_s :$  and the next dot.<sup>17</sup>

- (61) a.  $\llbracket \text{know} \rrbracket = \lambda \phi_{st} . \lambda x_e . \lambda w_s : \phi(w) = 1$ . For all *w'* compatible with *x*’s beliefs in *w*,  $\phi(w') = 1$ .  
b.  $\llbracket \text{stop} \rrbracket = \lambda P_{<e,st>} . \lambda x_e . \lambda w_s : \text{there is a } w' \text{ anterior to } w \text{ such that } P(x)(w') = 1. P(x)(w) = 0$ .

The presupposition that obtains when (61-a) combines with its arguments is represented as “ $\phi(w) = 1$ ” which states the truth of the complement proposition, and the presupposition that obtains when (61-b) combines with its arguments is represented as “there is a *w'* anterior to *w* such that  $P(x)(w') = 1$ ” which states that the complement predicate was true of the subject at some point in the past. In situations where

<sup>16</sup>Even though they adopt a notation formed with a statement expressing the domain condition and a statement describing the value, one should resist viewing (Heim & Kratzer 1998)’s system as a bidimensional treatment of presupposition: from a very formal point of view, their system does not make it possible to access the statement describing the value when the domain condition is unfulfilled. They share this property with traditional trivalent approaches of presupposition. See (George 2008) on this point.

<sup>17</sup>Here, for simplicity, I consider variables of type *s* to represent world-time pairs.

this is not the case, we derive a presupposition failure: the resulting propositions fails to have a bivalent truth-value. Therefore, when (61-b) enters in the construction of (60-b-i), we derive a proposition which fails to have a truth-value in situations where the amplifiers have never buzzed. In deriving (60-b), (60-b-i) then combines with *know*. In the end, the function  $\llbracket (60-b) \rrbracket$  that we compute is the one in (62).

$$(62) \quad \lambda w_s : \llbracket (60-b-i) \rrbracket(w) = 1. \text{ For all } w' \text{ compatible with } x\text{'s beliefs in } w, \\ \llbracket (60-b) \rrbracket(w') = 1.$$

There are two possible positions regarding how the computation of (62) proceeds in situations where the amplifiers have never buzzed. The first position is to consider that the computation stops as soon as we evaluate a proposition that fails to receive a bivalent truth-value. Given that to determine the truth-value of (62), we first need to determine whether its domain condition is fulfilled, we first need to check the value of  $\llbracket (60-b-i) \rrbracket$ . And given that  $\llbracket (60-b-i) \rrbracket$  fails to receive a bivalent truth-value in the situations under consideration, the computation would stop here, without any further consideration. The second position is to consider that when a proposition fails to have a bivalent truth-value, it gets a *third* truth-value. We would thus be able to determine the value that  $\llbracket (60-b-i) \rrbracket$  would return in the situations under consideration, which is precisely the third truth-value. And given that this value is different from 1, the domain condition of (62) would fail to be satisfied. In the end, both options predict (60-b) to yield a presupposition failure in situations where (60-b-i) would yield a presupposition failure: this property is welcome, given that *know* is said to *project* the presuppositions of its complement proposition. However, whereas it is a general consequence of the first option that we eventually observe a presupposition failure whenever a proposition that fails to get a bivalent truth-value enters in the computation of the final proposition, the predictions of the second option crucially rely on the formulation of the domain condition. Importantly, there is no constraint in this framework which would exclude a lexical entry identical to (61-a) with the exception that the domain condition would be formulated as  $\phi(w) \neq 0$  (the two formulations would be equivalent in a bivalent framework). *Ceteris paribus*, such a lexical entry would *not* project the presuppositions of its complement proposition with the second option, contrary to the apparent behavior of *know*. This approach runs into further similar problems when we try to implement vague propositions and consider sentences like (60-a). I will not discuss these problems here, but their investigation follows the very same pattern that we just went through, with the additional consideration of the specificity of projection of vague propositions.

#### 4.3.2.2 Presuppositional Expressions in ST5

Sentences (60-a) and (60-b) show that even sentences with no connective can have complex presuppositions, that is to say presuppositions resulting in a systematic way from the parts of the sentence. I claim that in addition to derive presuppositions of complex sentences from the presuppositions of the sentences that they concatenate, ST5 also provides us with the tools to derive the presuppositions of simple sentences from their parts. I will propose my own adaptation of (Heim & Kratzer 1998)'s notation to deal with subpropositional objects in ST5. This will allow us to deal with more subtle presuppositions by taking the five different values of the system into account, and to provide a systematic notion of presupposition fulfillment based on the two notions of satisfaction of the system.

To this end, I will also need to turn ST5 semantics into an intensional system. This can be achieved by adding a set of possible world-time pairs to the model, and by relativizing the definitions of vague predicates, of the connectives and of tolerant and strict satisfactions to a world-time pair.

**Definition 4.3.1** (Intensional Semantics in ST5). *A model  $\mathcal{M}$  consists of a non-empty domain of individuals  $\mathcal{D}$ , an interpretation function  $\mathcal{I}$  and a set of world-time pairs  $\mathcal{W}$  such that:*

- *For any predicate  $P \in \mathcal{L}$ , any individual name  $a \in \mathcal{L}$  and any world-time pair  $w \in \mathcal{W}$ ,  $\mathcal{I}(Pa)(w) = \mathcal{V}$  iff  $\mathcal{I}(a)$  is a borderline case for  $\mathcal{I}(P)$  at  $w$ ,  $\mathcal{I}(Pa)(w) \in \{0, 1\}$  otherwise.*
- *For any wff  $\phi$  and any  $w \in \mathcal{W}$ ,  $\mathcal{I}(\neg\phi)(w) = 1 - \mathcal{I}(\phi)(w)$ .*
- *For two wff  $\phi$  and  $\psi$  and any  $w \in \mathcal{W}$ ,  $\mathcal{I}(\phi \wedge \psi)(w) = \min(\mathcal{I}(\phi)(w), \mathcal{I}(\psi)(w))$ ,  $\mathcal{I}(\phi \vee \psi)(w) = \max(\mathcal{I}(\phi)(w), \mathcal{I}(\psi)(w))$  and  $\mathcal{I}(\phi \rightarrow \psi)(w) = \mathcal{I}(\neg\phi \vee \psi)(w)$*

*Strict and tolerant satisfactions are consequently relativized to a world-time pair:*

- **Strict satisfaction:**  $\mathcal{M} \models^{s,w} \phi$  iff  $\mathcal{I}(\phi)(w) = 1$
- **Tolerant satisfaction:**  $\mathcal{M} \models^{t,w} \phi$  iff  $\mathcal{I}(\phi)(w) \geq \frac{1}{2}$

As a consequence, the interpretation of any well-formed formula is a function from a possible world to a logical truth-value. We will therefore use our version of the lambda notation to represent the interpretation of well-formed formulas. For instance, to talk about (63-a) using the resources of ST5, we would talk about a proposition  $\phi$  and a model whose interpretation yields for  $\phi$  a function we would write as in (63-b).



- (63) a. The amplifiers are loud.  
 b.  $\mathcal{I}(\phi) = \lambda w_s. 1$  iff  $A$  are clearly loud at  $w$ ,  $\mathcal{V}$  iff  $A$  are borderline-loud at  $w$ ,  $0$  otherwise.

Functions of type  $\langle s, t \rangle$  take world-time pairs as inputs and return logical values in  $\{0, \mathcal{P}^0, \mathcal{V}, \mathcal{P}^1, 1\}$  as outputs.

Before introducing lexical entries for presuppositional expressions in ST5, let us look back at (60-a) and (60-b) and consider the range of possible situations in regard of the fulfillment of their presuppositions. In situations where the amplifiers were clearly loud before, (60-a) would not really be problematic, for (60-a-i) would simply be true (truth-value 1). As a consequence, (60-a) would be simply true (truth-value 1) if the amplifiers are still clearly loud, simply false (truth-value 0) if they are now clearly not loud, and both true and false (truth-value  $\mathcal{V}$ ) if they are borderline-loud. Note that even in this situation of bivalence regarding the presupposition, the vague dimension of the assertive part has to be taken in consideration. In situations where the amplifiers buzzed before but are not currently buzzing, (60-b-i) would be simply true (truth-value 1) and (60-b) would be simply true (truth-value 1) if John believes so and false (truth-value 0) if he does not.<sup>18</sup> In parallel, in situations where the amplifiers were clearly not loud before, (60-a-i) would be simply false (truth-value 0), thus yielding a presupposition failure (truth-value  $\mathcal{P}^0$ ) when evaluating (60-a) as a whole. Similarly, in situations where the amplifiers buzzed before and are currently buzzing, (60-b-i) would be simply false (truth-value 0), thus yielding a presupposition failure (truth-value  $\mathcal{P}^0$ ) when evaluating (60-b) as a whole. Things get more complex when we consider situations where the amplifiers were borderline-loud before. In such a situation, (60-a-i) would be both true and false (truth-value  $\mathcal{V}$ ). It seems to me that (60-a) would somehow yield a presupposition failure for the presupposition cannot properly be said to be fulfilled. On the other hand, it cannot properly be said to be unfulfilled either. Subsequently, if the amplifiers further turn out to be clearly not loud, I have the impression that (60-a) would be on the *true* side (truth-value  $\mathcal{P}^1$ ); and if the amplifiers are as loud as or even louder than before, (60-a) would be on the *false* side (truth-value  $\mathcal{P}^0$ ). Things appear to be simpler when we consider (60-b) in situations where the amplifiers have never buzzed. In such a situation, the

<sup>18</sup>I take “ $X$  believes  $\phi$ ” to be the assertive part of “ $X$  knows  $\phi$ ” and to return a bivalent truth-value regardless of the truth status of  $\phi$ . It might well be the case that things are more complex, and that one should consider justified belief for the assertive part. But whatever we take to be the assertive part, the crucial point here is how each part contributes to the value of the whole proposition.

simple sentence (60-b-i) would be associated with a presupposition failure (truth-value  $\mathcal{P}^0$ ). As noted earlier, *know* seems to project the presuppositions associated with its complement proposition: (60-b) simply yields a presupposition failure too in this situation (truth-value  $\mathcal{P}^0$ ).

On the basis of these particular examples, I propose that there is a generalization about the way in which “presuppositional parts” of sentences contribute to truth values – and one that can be expressed naturally in terms of the intensional version of ST5. Specifically, the intensional version of ST5 should be used to model natural language in the following manner. Suppose that we have a sentence  $S$  whose assertive part can be paraphrased by a sentence  $A$  and whose presuppositional part can be paraphrased by a sentence  $P$ . In that case, we should imagine that we are considering a proposition  $\phi$  evaluated with respect to a model that establishes the following relations between  $\phi$  and two other propositions  $\psi$  and  $p$  (corresponding respectively to  $S$  and  $A$ ):

- $\mathcal{I}(\phi)(w) = \mathcal{I}(\psi)(w)$  if  $\mathcal{I}(p)(w) = 1$
- $\mathcal{I}(\phi)(w) = \mathcal{P}^1$  if  $\mathcal{P}^0 < \mathcal{I}(p)(w) < 1$  and  $\mathcal{I}(\psi)(w) = 1$
- $\mathcal{I}(\phi)(w) = \mathcal{P}^0$  otherwise

The generalization is stated in terms of logical truth-values in order to make the comparison with the considerations above easier, but one should note that it can easily be restated in terms of tolerant and strict satisfaction of the presuppositional and the assertive parts. In order to derive truth-values fitting the generalization from the lexical entries composing a proposition, I propose that the interpretation of lexical entries of presuppositional words involve a special function,  $\star$ , which takes two logical values as its arguments and returns a logical value as its output. Its semantics is defined in Definition 4.3.2 and the resulting truth-table is represented in Table 4.2.

**Definition 4.3.2** ( $\star$  Function).  *$\star$  is a function from pairs of truth-values to truth-values such that, for any well-formed formulas  $\phi$  and  $\psi$ , any  $w \in \mathcal{W}$  and any model  $\mathcal{M}$  with an interpretation function  $\mathcal{I}$ ,*

$$\begin{aligned}
& \mathcal{I}(\phi)(w) \star \mathcal{I}(\psi)(w) \\
&= \mathcal{I}(\psi)(w) \text{ iff } \mathcal{M} \models^{s,w} \phi \\
&= \mathcal{P}^1 \text{ otherwise if } \mathcal{M} \models^{t,w} \phi \text{ and } \mathcal{M} \models^{s,w} \psi \\
&= \mathcal{P}^0 \text{ otherwise}
\end{aligned}$$

$\star$	0	$\mathcal{P}^0$	$\mathcal{V}$	$\mathcal{P}^1$	1
0	$\mathcal{P}^0$	$\mathcal{P}^0$	$\mathcal{P}^0$	$\mathcal{P}^0$	$\mathcal{P}^0$
$\mathcal{P}^0$	$\mathcal{P}^0$	$\mathcal{P}^0$	$\mathcal{P}^0$	$\mathcal{P}^0$	$\mathcal{P}^0$
$\mathcal{V}$	$\mathcal{P}^0$	$\mathcal{P}^0$	$\mathcal{P}^0$	$\mathcal{P}^0$	$\mathcal{P}^1$
$\mathcal{P}^1$	$\mathcal{P}^0$	$\mathcal{P}^0$	$\mathcal{P}^0$	$\mathcal{P}^0$	$\mathcal{P}^1$
1	0	$\mathcal{P}^0$	$\mathcal{V}$	$\mathcal{P}^1$	1

Table 4.2: The truth-table of the  $\star$  function. Rows represent the first argument of  $\star$ , and columns represent the second argument of  $\star$ .

Note that the  $\star$  function is essentially a semantic device for it directly operates on logical truth-values. I see no reason for enhancing ST5 with a new connective that effects what the star operator does. Relatedly, I think that there is no linguistic expression that would simply express the meaning of  $\star$ .

As a last step before proposing new lexical entries, I shall bring some modifications to (Heim & Kratzer 1998)’s notation. First, I shall only consider total functions: I shall therefore not state any domain condition. Second, I shall never directly *compare* truth-values in the statements describing the range of the functions: as a consequence, I will not state conditions that make use of  $=$ , like *if*  $\phi(w) = 1$  (however some statements explicitly mention the truth-value that the function *returns*). Finally, I introduce the two following functions to deal with intensional representations.

**Definition 4.3.3** (Big Disjunction over Intensions). *For any function  $f$  of type  $\langle s, t \rangle$  and set of world-time pairs  $S$ ,  $\max_S f$  is the highest value in  $\{f(w) : w \in S\}$ .*

**Definition 4.3.4** (Big Conjunction over Intensions). *For any function  $f$  of type  $\langle s, t \rangle$  and set of world-time pairs  $S$ ,  $\min_S f$  is the lowest value in  $\{f(w) : w \in S\}$ .*

Adopting a lambda notation building on that of (Heim & Kratzer 1998)’s (with the modifications described above) allows us to further adapt the lexical entries above and to derive the interpretations of (60-a) and of (60-b) in the same spirit.

(64) Lexical entries

- a.  $\llbracket \text{the amplifiers} \rrbracket = a, \llbracket \text{John} \rrbracket = j$
- b.  $\llbracket \text{not} \rrbracket = \lambda\phi_{st}.\lambda w_s.1 - \phi(w)$ .
- c.  $\llbracket \text{buzz} \rrbracket = \lambda x_e.\lambda w_s. 1$  iff  $x$  is buzzing in  $w$ , 0 otherwise.
- d.  $\llbracket \text{loud} \rrbracket = \lambda x_e.\lambda w_s. 1$  iff  $x$  is clearly loud in  $w$ ,  $\mathcal{V}$  iff  $x$  is borderline-loud, 0 otherwise.
- e.  $\llbracket \text{know} \rrbracket = \lambda\phi_{st}.\lambda x_e.\lambda w_s.\phi(w) \star \min_{\text{Dox}(x,w)} \phi$ .

$$f. \quad \llbracket \text{stop} \rrbracket = \lambda P_{\langle e, st \rangle} . \lambda x_e . \lambda w_s . \max_{\text{Before}(w)} P(x) \star \llbracket \text{not} \rrbracket (P(x))(w).$$

$$\begin{aligned}
(65) \quad & \llbracket \text{The amplifiers have stopped being loud} \rrbracket \\
& a. = \llbracket \text{stop} \rrbracket (\llbracket \text{loud} \rrbracket) (\llbracket \text{the amplifiers} \rrbracket) \\
& b. = \lambda P_{\langle e, st \rangle} . \lambda x_e . \lambda w_s . \max_{\text{Before}(w)} P(x) \star \llbracket \text{not} \rrbracket (P(x))(w) . (\llbracket \text{loud} \rrbracket) (\llbracket \text{the amplifiers} \rrbracket) \\
& c. = \lambda w_s . \max_{\text{Before}(w)} \llbracket \text{loud} \rrbracket (a) \star \llbracket \text{not} \rrbracket (\llbracket \text{loud} \rrbracket (a))(w) . \\
& d. \text{ Therefore, from Def. 4.3.2, for any } w \in \mathcal{W}, \\
& \quad \llbracket \text{the amplifiers have stopped being loud} \rrbracket (w) \\
& e. = \llbracket \text{not} \rrbracket (\llbracket \text{loud} \rrbracket (a))(w) \text{ iff there is a } w' \text{ prior to } w \text{ such that} \\
& \quad \llbracket \text{loud} \rrbracket (a)(w') = 1, \mathcal{P}^1 \text{ iff the former is not the case but there is a } w' \text{ prior} \\
& \quad \text{to } w \text{ such that } \mathcal{P}^0 < \llbracket \text{loud} \rrbracket (a)(w') < 1 \text{ and } \llbracket \text{not} \rrbracket (\llbracket \text{loud} \rrbracket (a))(w) = 1, \\
& \quad \mathcal{P}^0 \text{ otherwise.} \\
& f. = 1 \text{ iff the amplifiers were clearly loud before and are now clearly not} \\
& \quad \text{loud, } \mathcal{V} \text{ iff the amplifiers were clearly loud before and are now borderline-} \\
& \quad \text{loud, } 0 \text{ iff the amplifiers were clearly loud before and are still clearly loud,} \\
& \quad \mathcal{P}^1 \text{ iff the amplifiers were borderline-loud before and are clearly not loud} \\
& \quad \text{now, } \mathcal{P}^0 \text{ otherwise.}^{19}
\end{aligned}$$

$$\begin{aligned}
(66) \quad & \llbracket \text{John knows that the amplifiers have stopped buzzing} \rrbracket \\
& a. = \llbracket \text{know} \rrbracket (\llbracket \text{stop} \rrbracket (\llbracket \text{buz} \rrbracket) (\llbracket \text{the amplifiers} \rrbracket)) (\llbracket \text{John} \rrbracket) \\
& b. = \lambda \phi_{st} . \lambda x_e . \lambda w_s . \phi(w) \star \min_{\text{Dox}(x,w)} \phi . \\
& \quad (\llbracket \text{the amplifiers have stopped buzzing} \rrbracket) (\llbracket \text{John} \rrbracket) \\
& c. = \lambda w_s . \llbracket \text{the amplifiers have stopped buzzing} \rrbracket (w) \star \\
& \quad \min_{\text{Dox}(j,w)} \llbracket \text{the amplifiers have stopped buzzing} \rrbracket . \\
& d. \text{ Therefore, from Def. 4.3.2, for any } w \in \mathcal{W}, \\
& \quad \llbracket \text{John knows that the amplifiers have stopped buzzing} \rrbracket (w) \\
& e. = \min_{\text{Dox}(j,w)} \llbracket \text{the amplifiers have stopped buzzing} \rrbracket \text{ iff} \\
& \quad \llbracket \text{the amplifiers have stopped buzzing} \rrbracket (w) = 1, \\
& \quad \mathcal{P}^1 \text{ iff } \mathcal{P}^0 < \llbracket \text{the amplifiers have stopped buzzing} \rrbracket (w) < 1 \text{ and} \\
& \quad \min_{\text{Dox}(j,w)} \llbracket \text{the amplifiers have stopped buzzing} \rrbracket = 1, \mathcal{P}^0 \text{ otherwise.}
\end{aligned}$$

<sup>19</sup>The reader may want to consult the following table to help with the last step.

Loud before	Loud now	(65)
0	0/ $\mathcal{V}$ /1	$\mathcal{P}^0$
$\mathcal{V}$	0/ $\mathcal{V}$	$\mathcal{P}^0$
	1	$\mathcal{P}^1$
	0	0
1	$\mathcal{V}$	$\mathcal{V}$
	1	1

- f. = 1 iff the amplifiers were buzzing before and are not buzzing now and John believes so, 0 iff the amplifiers were buzzing before and are not buzzing now and John’s beliefs do not exclude that they still are,  $\mathcal{P}^0$  otherwise.<sup>20</sup>

The intuitions about our truth-value judgments for hybrid sentences like (60-b) and (60-a) in critical situations seem a bit difficult to access, maybe because of a demanding processing. In the end, maybe only experimental data can discriminate between theories that make different predictions regarding truth-value judgments for these kinds of sentences. Nonetheless, any theory has to make *some* predictions for these sentences. As shown earlier, (Heim & Kratzer 1998)’s formalization permits several approaches of sentences like (60-b), but some theories of presupposition are more precise on this issue. For example, (Karttunen 1973) proposed to categorize factives and aspectual verbs (such as *know* and *stop*) as what he famously called *holes*:

“If the main verb of the sentence is a hole, then the sentence has all the presuppositions of the complement sentences embedded in it.”

This makes direct predictions regarding (60-b), but it provides no way of distinguishing between a situation where the amplifiers are still buzzing (which could be referred to as a “matrix presupposition” failure) and a situation where the amplifiers have never buzzed (which could be referred to as an “embedded presupposition” failure): in the first situation, the complement of the factive is false so it yields a presupposition failure; in the second situation the *inherited* presupposition is *unfulfilled* so it *also* yields a presupposition failure. On the contrary, ST5 provides us with the tools to deal with this variety of situations because the presuppositional part of the whole proposition would have the value 0 in the first case and the value  $\mathcal{P}^0$  in the second case. It is not clear whether speakers would give different truth-value judgments in these two situations for (60-b), and I decided here to treat them equally (as can be seen in the truth-table in Table 4.2), as does a theory *à la* Karttunen.

<sup>20</sup>The reader may want to consult the following table to help with the last step.

Stopped buzzing	John believes	(66)
0 $\mathcal{P}^0$	0/ $\mathcal{P}^0$ /1	$\mathcal{P}^0$
1	0 $\mathcal{P}^0$ 1	0 $\mathcal{P}^0$ 1

The reason why some existing theories of presupposition provide no clear predictions concerning sentences like (60-b) and (60-a) is because they only consider *bivalent* presuppositions. As long as a theory of presupposition treats the presuppositional content as bivalent, it will face difficulties when trying to account for sentences where the presuppositional content is vague. This is precisely the weakness that we avoid with a  $\star$  function that can be seen as based on the notions of strict and tolerant satisfaction: it allows us to escape the traditional duality of either “fulfilled” or “unfulfilled” presuppositions. The first clause of Def. 4.3.2 states that when the presuppositional part of a proposition is strictly satisfied, the whole proposition gets the value of its assertive part: in this situation one would traditionally say that the presupposition is “fulfilled”. The second clause considers the case where the presupposition is only tolerantly satisfied. To some extent, one could see this as a condition where the presupposition is “partly fulfilled”. The whole proposition will be “partly true” if the assertive part is true itself: that’s what  $\mathcal{P}^1$  stands for. Finally, the third clause states that even if the presupposition is tolerantly satisfied (“partly fulfilled”), the whole proposition should not be considered “partly true” if the assertive part is not strictly satisfied; nor if the presupposition is not satisfied at all. But still, such a proposition should not be merely false, because the presupposition is not “fulfilled”: that’s what  $\mathcal{P}^0$  stands for.

A final word is in order regarding the way we have proposed to derive presuppositions from lexical entries. Def. 4.3.2 has the effect of assigning propositions a truth-value depending on the strict and tolerant satisfactions of its presuppositional and assertive parts. Importantly, no lexical entry explicitly checks a logical truth-value on mere stipulative grounds: the only function from logical truth-values to logical truth-values is  $\star$ , and its definition echoes the theoretical notions of satisfaction of ST5, as we just reminded.

One might argue however, on the basis of the apparent projection-behavior of expressions such as *think* as exemplified in (67), that we need to retrieve and check the logical truth-value of the presuppositional part of the complement proposition (*think* might be a filter in Karttunen’s terminology). Indeed, (67) seems to be true when John thinks that the amplifiers were buzzing before and are not anymore, false when John thinks the amplifiers were buzzing before and are currently buzzing, and it seems to yield a presupposition failure when John thinks the amplifiers have never buzzed, regardless of whether the amplifiers were actually buzzing before.<sup>21</sup> This

---

<sup>21</sup>It is not totally clear what truth-value judgment we would give for (i) in situations where John thinks that the amplifiers are buzzing right now but is ignorant as to whether the amplifiers were

suggests that (67) should be modeled as a linguistic proposition with (67-a) as its presuppositional part, and a way to express this presuppositional part would be to check that the linguistic proposition corresponding to “the amplifiers have stopped buzzing” never gets  $\mathcal{P}^0$  nor  $\mathcal{P}^1$  in any world compatible with John’s beliefs. This is proposed in (67-b).

- (67) John thinks that the amplifiers have stopped buzzing.
- a. John does not think that the amplifiers have never buzzed.
  - b.  $\lambda w_s. [\lambda u_s. 1 \text{ iff } \max_{Dox(j,u)} \llbracket (67\text{-a}) \rrbracket \notin \{\mathcal{P}^0, \mathcal{P}^1\}, 0 \text{ otherwise.}](w) \star \min_{Dox(j,w)} \llbracket (67\text{-a}) \rrbracket.$

As said just before, (67-b) cannot obtain in our present system because we do not allow lexical entries to explicitly check for logical values. But this is actually not a problem, because the presupposition-filtering aspect of *think* is in fact directly derived from its assertive part, as can be seen in (68).

- (68)  $\llbracket \text{think} \rrbracket = \lambda \phi_{st}. \lambda x_e. \lambda w_s. \min_{Dox(x,w)} \phi.$
- a.  $\llbracket \text{John thinks that the amplifiers have stopped buzzing} \rrbracket$
  - b.  $= \llbracket \text{think} \rrbracket (\llbracket \text{the amplifiers have stopped buzzing} \rrbracket) (\llbracket \text{John} \rrbracket)$
  - c.  $= \lambda \phi_{st}. \lambda x_e. \lambda w_s. \min_{Dox(x,w)} \phi. (\llbracket \text{the amplifiers have stopped buzzing} \rrbracket)(j)$
  - d.  $= \lambda w_s. \min_{Dox(j,w)} \llbracket \text{the amplifiers have stopped buzzing} \rrbracket.$
  - d. From Def. 4.3.4, for any  $w \in \mathcal{W}$   
 $\llbracket \text{John thinks that the amplifiers have stopped buzzing} \rrbracket(w) =$   
 $\mathcal{P}^0$  if  $\llbracket \text{the amplifiers have stopped buzzing} \rrbracket(w') = \mathcal{P}^0$  for all  $w'$  compatible with John’s beliefs in  $w$ .

In other words, (68) shows that a presupposition failure obtains whenever John thinks that the amplifiers have never buzzed. This result suggests that (Karttunen 1973)’s distinction between *filters* and *holes* may result from independently motivated lexical aspects of words. For instance, *know* is a hole because it presupposes its complement proposition: as a result, from Def. 4.3.2, it inherits the presuppositions of its complement. And *think* is a filter because of its attitude orientation: the logical

---

buzzing before. It seems to me though that I would consider “John thinks that the amplifiers have stopped buzzing” false, given that his knowledge is compatible with the amplifiers being buzzing before and being still buzzing.

value of its assertive part depends on the logical value its complement proposition gets when evaluated against its subject’s beliefs.

In conclusion, ST5 predicts more nuanced judgments for presuppositional sentences than its competing because it takes the relative “gradedness” of the presuppositions into account. Even though some theories do deal with hybrid sentences like (60-b), none of them deal with hybrid sentences like (60-a) to my knowledge. Because Definition 4.3.2 covers all the satisfaction possibilities, it is easy to see that the system is now completely predictive with respect to the kind of proposition (ie. bivalent, vague, presuppositional or hybrid<sup>22</sup>) that appears as a presupposition of the whole sentence.

## 4.4 Conclusions

ST provides us with a notion of assertoric ambiguity that, along with some bridge principles, lets us explain our conflicting truth-value judgments in case of vagueness. Adding two symmetrical values around  $\frac{1}{2}$  has made it possible to capture the difference between *not true* and *false* judgments and between *not false* and *true* judgments by virtue of bridge principles based on ST notions of satisfaction. Moreover, these values lend themselves naturally to an account for the asymmetry of truth-value judgments concerning the positive and negative counterparts of presuppositional sentences. Furthermore, we now have a system that incorporates *both* vagueness and presupposition while *also* accounting for the differences in the judgments they trigger. At the same time, there is clearly more to be said about how the presuppositions of complex sentences depend on the presuppositions of the simple sentences they embed; here we had to add some stipulations. More data would be welcome in order to test the predictions of ST5. The next two chapters will present an experimental design for eliciting truth-value judgments for vagueness and presupposition.

---

<sup>22</sup>As an example of how ST5 deals with hybrid *presuppositions*, consider (i-a), its presupposition being (i-b):

- (i)    a.    John knows that the amplifiers have stopped being loud.
- b.    The amplifiers have stopped being loud.

We saw earlier that in cases where the amplifiers were borderline-loud before decreasing in volume, the hybrid proposition expressed by (i-b) gets the value  $\mathcal{P}^0$ , which prevents it from being even tolerantly satisfied; therefore (i-a) will also get the value  $\mathcal{P}^0$  by Definition 4.3.2.



# Chapter 5

## Three Experiments<sup>1</sup>

This chapter reports the results of two experiments concerning truth-value judgments for sentences with vague predicates and with presuppositional expressions. Our initial goal in conducting these experiments was to test the predictions of ST5. The development of ST5 was largely motivated by the idea that there is an important distinction between vagueness and presupposition: the former licenses *both true and false* (glutty) judgments and *neither true nor false* (gappy) judgments about borderline cases while the latter license gappy judgments but exclude glutty judgments in case of presupposition failure. And, beyond this, the system makes further, finer predictions about truth-value judgments. However, the interest of the experimental results reported here goes beyond their potential relevance to evaluating ST5. Overall, they establish that the pattern of truth-value judgments for presuppositional sentences is different from the pattern of truth-value judgments for vague sentences.

Section 5.1 presents a first experiment where participants were asked to rate the truth of vague and presuppositional sentences on a 5 point scale. The very format of this scale aimed at directly testing the formal aspects of ST5, but the results turned out to be quite messy. They clearly showed, though, that vague sentences and presuppositional sentences give rise to different patterns of judgments. Section 5.2 presents a second experiment where the scale was replaced with a threefold choice between “Completely true”, “Completely false”, and “Neither”. In addition, in this second version, pictures associated with presuppositional sentences described a sequence of states whereas they presented isolated situations in the first experiment.

---

<sup>1</sup>In this chapter and Chapter 7, I present descriptive analyses, but most aspects of the discussion in these chapters rely on robust observations, so whether more elaborate analyses bring new significant differences or dismiss an alleged effect should not have any dramatic impact on the content of my discussion. In any event, the data that I collected are available for statistical treatment – readers wishing to examine the data should contact me at [jeremy.e.zehr@gmail.com](mailto:jeremy.e.zehr@gmail.com).

I wanted to come up with a protocol that would reveal a clear distinction in (non-classical) truth-value judgments for vague sentences describing borderline cases, on the one hand, and for presuppositional sentences whose presuppositions are unfulfilled, on the other. The results of the first experiment were replicated for vagueness (modulo the different format of the answers). The answers that participants gave for presuppositional sentences can be divided in two groups depending on the pictures they were presented with. In any case, they show a distribution which is clearly and systematically different from their answers for vague sentences.

## 5.1 First Experiment: Testing ST5

The first experiment aimed to test some major properties of the system ST5. Importantly, ST5 makes predictions regarding the truth-value judgments that speakers report when describing borderline cases and in situations of presupposition failure. Among the experiments on vagueness mentioned in Chapt. 3, only (Serchuk et al. 2011) directly looked at gappy and glutty judgments expressed with *truth-predicates*, such as *both true and false* and *partially true and partially false*. They did find some *neither true nor false* judgments but they found fewer *partially true and partially false* judgments.<sup>2</sup> One aspect of their design was that participants were asked to *choose* between the following truth-value judgments: *true*, *false*, *neither*, *both*, *partially true and partially false*, and *don't know*. It is possible that showing all these possibilities at once and forcing participants to choose only one of them could have masked the variety of available judgments in critical situations. In particular, ST5 predicts vague descriptions of borderline cases to be possibly judged *neither true nor false* but also *both true and false*.

I therefore conducted a pilot experiment not reported here, drawing on their design but where I tested each of the various truth-value judgments separately. Unfortunately, participants reported that the task was too heavy and too confusing, especially when it came to judging the falsity of negative sentences. Consequently, the few results I obtained were clearly impossible to analyze. I finally decided to set up a simpler task for participants which would nonetheless test the system ST5.

---

<sup>2</sup>As a reminder, they found very few *both true and false* judgments. (Alxatib & Pelletier 2011) found a significant acceptance of “Both ADJ and not ADJ” descriptions for borderline cases, but it is not clear whether it was significantly lower than the acceptance of “Neither ADJ nor nor ADJ”.

### 5.1.1 Design

The first experiment was designed as a truth-value judgment task. Subjects were shown pictures along with descriptions involving a vague or a presuppositional expression and asked to give their judgment on a 5-point scale going from “Completely false” to “Completely true”. No value was explicitly associated with the 5 buttons on the scale, even though the leftmost button was near “Completely false” and the rightmost button was near “Completely true” (see for instance Figure 5.1). In the critical conditions, the pictures associated with the vague descriptions represented borderline cases and the pictures associated with the presuppositional descriptions represented situations of presupposition failure. In the control conditions, the pictures represented clear instances or clear counter-instances of the vague or the presuppositional expression. I expected subjects’ answers in the *control* conditions to lie at the extremes of the 5-point scale, i.e. subjects were expected to give clear *true* and clear *false* judgments in the control conditions. By contrast, subjects’ answers in the *critical* conditions were expected to range somewhere between these two points, based on the non-bivalent status that authors have attributed to vague and presuppositional propositions in contexts of this sort.

In addition, each sentence used in the descriptions was tested along with its negative counterpart. Negation constitutes a hallmark in the domain of presupposition and ST5 makes strong predictions about it. More precisely, based on the negation operator defined in ST5, subjects’ answers for the *vague* descriptions in the critical contexts were expected to be the same across negation. By contrast, subjects’ answers for the *presuppositional* descriptions in the critical contexts were expected to vary across negation.

Whereas Serchuk&al. directly asked people to imagine borderline cases on “the spectrum of rich/heavy women”, the present experiment presented subjects with pictures. Visually representing borderline cases is a challenging task, because different persons situate them in different areas on the scale associated with the vague term under evaluation. (Ripley 2011) and (Egré et al. 2013) showed subjects a series of entities of various enough measurements to be sure that subjects would treat some of these entities as borderline cases. However, this strategy was not available for the present purposes. (Klein 1980) proposes a process of partial categorization with vague adjectives which sorts their arguments into a positive, a negative, and a “gappy” extension. Importantly, this process is *sensitive to the comparison class*, which means that entities that fall into the gappy extension when compared with one set of entities will not necessarily fall into the gappy extension when compared with a different set

of entities. Assuming that entities who lie in the middle with respect to the measure expressed by the vague adjective are to fall into the gappy extension, I presented subjects with pictures representing 3 entities such that one of them lied halfway between the two others in regard of the measure associated with the vague adjective used in the description. As expected, subjects naturally treated the entity with a central measure as a borderline case.<sup>3</sup> To avoid systematic symmetry between the relative order of the described entities in the picture and the relative order of the expectedly corresponding button on the 5-point scale, the order of the entities in the picture was pre-randomized. An arrow was placed under the object against which participants had to evaluate the description. Therefore, the critical conditions for the vague descriptions consisted in showing a picture where the arrow designated the entity whose measure was central, whereas the arrow indicated one of the two other entities in the control conditions. All the vague descriptions were of the form “The object indicated by an arrow is (not) ADJ”. An example of an affirmative vague description to be evaluated in a critical context is provided in Figure 5.1.

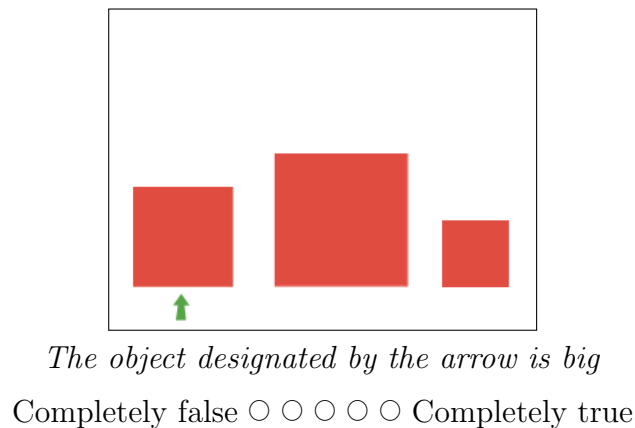
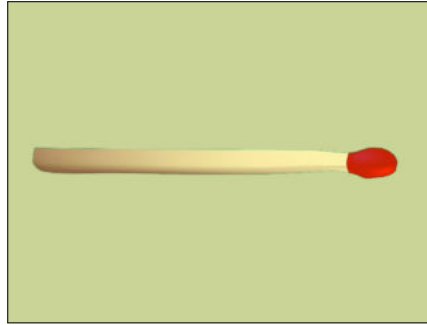


Figure 5.1: Example of an *affirmative vague* description in a *critical* context. The pointed square is supposed to lie in neither of *big*’s positive or negative extension.

It was obviously not possible to use the same strategy in building the pictures for the presuppositional descriptions. Because I wanted to control for the effect of the precise presuppositional expression used in the description, I chose to restrict my attention to the aspectual verb *stop* which is a usual example of a presuppositional expression. For this reason, every presuppositional description was of the form “X has stopped V-ing”. This form imposed some constraints on the pictures used as

<sup>3</sup>In a similar way, (Alxatib & Pelletier 2011) presented their participants with a picture representing 5 men of different heights to evaluate the vague adjective *tall*. It seems that most participants treated the man whose height was in the middle as a borderline case.

contexts of evaluation for the description. But the most important constraint came from the consideration of the process of global accommodation which is known to occur in cases of presupposition failure. Indeed, even in the absence of evidence for the fulfillment of a presupposition, speakers tend to consider it as fulfilled to make the discourse coherent. The pictures therefore had to describe situations clear enough to prevent participants from resorting to this strategy of global accommodation in the critical contexts. The pictures providing the control contexts represented situations where the process described by the complement verb was clearly over or was clearly still going on; and the pictures providing the critical contexts represented situations where the process described by the complement verb had clearly not started yet. Figure 5.2 exemplifies a presuppositional description to be evaluated in a critical context.



*The match has stopped burning*

Completely false ○ ○ ○ ○ ○ Completely true

Figure 5.2: Example of an *affirmative presuppositional* description in a *critical* context. The *presuppositional* expression is *stop burning*, and the unstruck match targets a situation where the event of burning has not started yet.

Concerning the vagueness conditions, in addition to pictures representing squares of various sizes, participants also saw pictures representing balls of various prices (indicated by ‘\$’, ‘\$\$’ and ‘\$\$\$’) and pictures representing speakers of various volumes (indicated by one, two or three waves). Concerning the presupposition conditions, in addition to pictures representing a *match* at different stages in a burning process (indicated by an unstruck, a burning and a burnt match), participants also saw pictures representing a *monkey* at different stages in an eating process (indicated by an unpeeled, a being-eaten and an empty banana) and pictures representing a *skydiver* at different stages in a falling process (indicated by the character being on a plane, in the air on on landed the ground). Each condition was thus tested three times.

This yielded a  $2 \times 2 \times 3$  interaction design, where the factors were **Description Type** (vague vs. presuppositional), **Polarity** (affirmative vs. negative) and **Context**

(critical vs. instance-oriented vs. counter-instance-oriented). Importantly, note that there is *a priori* an interaction between control contexts and **Polarity**:

- (69)
- a. *instance-oriented* contexts correspond to the pictures that make the affirmative descriptions *true*;
  - b. *counter-instance-oriented* contexts correspond to the pictures that make the affirmative descriptions *false*;
  - c. *instance-oriented* contexts correspond to the pictures that make the *negative* descriptions *false*;
  - d. *counter-instance-oriented* contexts correspond to the pictures that make the *negative* descriptions *true*

Figure 5.3 and Figure 5.4 respectively exemplify an affirmative vague description in an instance-oriented context and a negative presuppositional description in a counter-instance-oriented context.

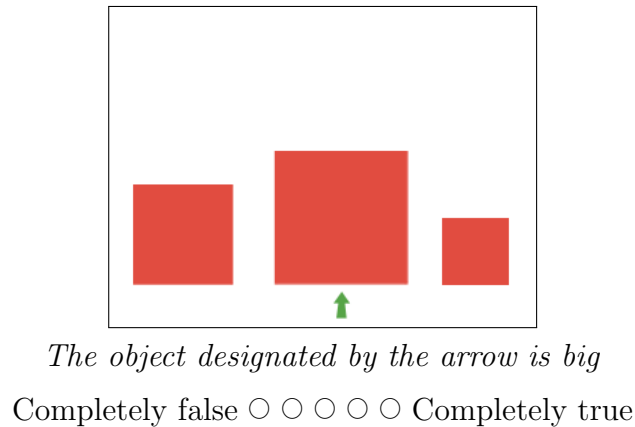


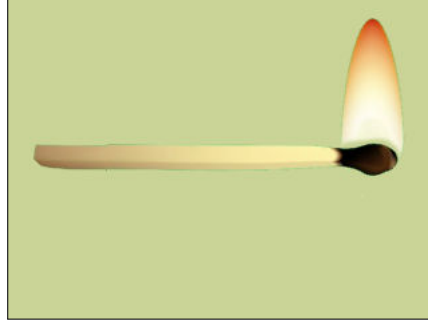
Figure 5.3: Example of an *affirmative vague* description of an *instance*. The pointed square lies in the positive extension of *big*.

#### 5.1.1.1 Predictions

One can imagine at least two types of predictions based on ST5 regarding this design: one can entertain a strong hypothesis (SH) or a weak hypothesis (WH).

On the strong hypothesis (SH), speakers distinguish the different truth-values in their answers to the extent possible, and thus match the five buttons to the five truth-values. The strong hypothesis makes the following predictions:

- i. **an interaction between Context and Polarity in the control contexts:** regardless of **Description Type**, subjects would click the rightmost button



*The match has not stopped burning*

Completely false ○ ○ ○ ○ ○ Completely true

Figure 5.4: Example of a *negative presuppositional* description of a *counter-instance*. The burning match lies in the negative extension of *stop burning*.

- (“Completely true”) when judging the *affirmative* descriptions in the *instance-contexts* and when judging the *negative* descriptions in the *counter-instance* contexts; they would click the leftmost button (“Completely false”) when judging the *negative* descriptions in the *instance-contexts* and when judging the *affirmative* descriptions in the *counter-instance* contexts.
- ii. **no effect of Polarity for the vague descriptions in the critical contexts:** subjects would click the button in the middle (interpreting it as “both true and false”) for both the affirmative and the negative descriptions.
  - iii. **an effect of Polarity for the presuppositional descriptions in the critical contexts:** subjects would click the middle-left button (interpreting it as “false but not completely false”) for the affirmative descriptions and the middle-right button (interpreting it as “true but not completely true”) for the negative descriptions.

Figure 5.5 presents fictive results compatible with SH.

Under a weaker hypothesis (WH), speakers do not maximally distinguish the five truth-values in their answers, but nonetheless indicate the distinctions that they can, by exploiting the fact that ST5 truth-values enter in an order relation. Subjects would not display a systematic correspondance between the button they click and the truth-value that ST5 assigns to the description in the specified context. However, the different buttons the subjects would click across the different conditions would globally enter the same order relation as the different logical truth-values that ST5

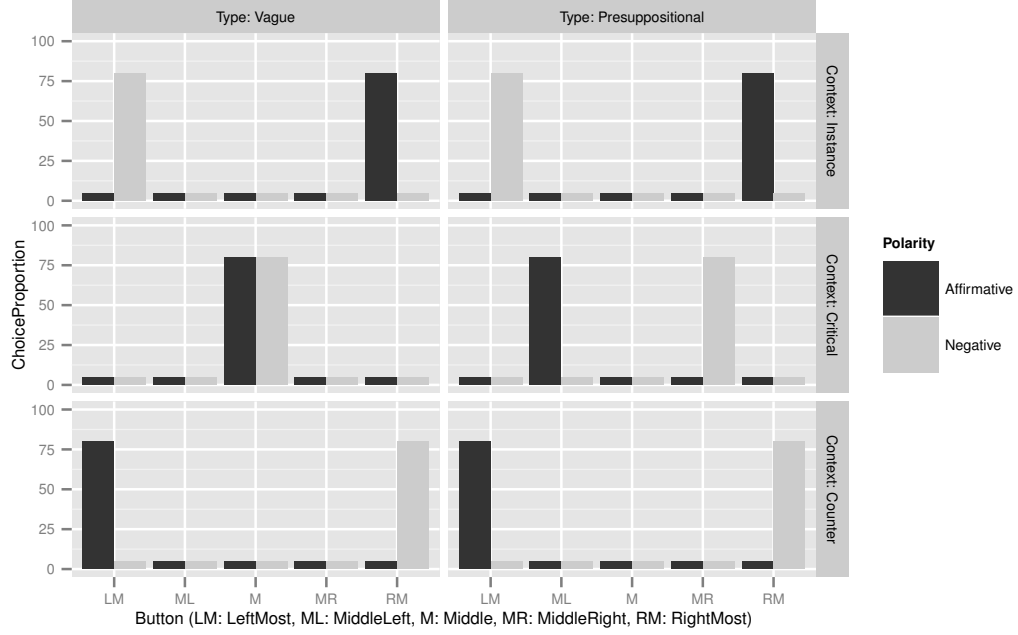


Figure 5.5: The predicted percentages of clicks on each button according to SH for Experiment 1. The left-most button was adjacent to the text *completely false*, and the right-most button was adjacent to the text *completely true*.

associates with each condition. More specifically, WH predicts:<sup>4</sup>

- i. **a *global* effect of Context:** in the critical contexts (logical truth-value in  $\{\mathcal{P}^0, \mathcal{V}, \mathcal{P}^1\}$ ), subjects would tend to click more central buttons than in the control contexts (logical truth-value in  $\{0, 1\}$ ).
- ii. **no effect of Polarity with the *vague* descriptions in the *critical* contexts:** subjects would tend to click central buttons regardless of the polarity (logical truth-value  $\mathcal{V}$ ).
- iii. **an effect of Polarity with the *presuppositional* descriptions in the *critical* contexts:** subjects would tend to click buttons on the left (i.e. close to “Completely false”) for the *affirmative* descriptions (logical truth-value  $\mathcal{P}^0$ ) and buttons on the right (i.e. close to “Completely true”) for the *negative* descriptions (logical truth-value  $\mathcal{P}^1$ ).
- iv. as a consequence, **an interaction between Description Type and Polarity in the critical contexts.**

<sup>4</sup>I indicate in parentheses the logical truth-value that ST5 assigns to the descriptions in the specified contexts.



## 5.1.2 Methods

### 5.1.2.1 Materials

As mentioned in the previous section, there were 3 sets of pictures for the vague descriptions and again 3 sets of pictures for the presuppositional descriptions, yielding 3 measures for each of the 12 conditions described above. In particular, each subject was presented with 3 affirmative and 3 negative descriptions for the vague sentences and again with 3 affirmative and 3 negative descriptions for the presuppositional sentences in the critical and in the control contexts. Figure A.1 and Figure A.2 in Appendix A.1.2 respectively list the three sets of pictures that were used to test the presuppositional descriptions and the three sets of pictures that were used to test the vague descriptions. In addition to these 36 items were included 6 fillers, ensuring participants' good understanding of the task. Half of the fillers used the 3 pictures built for the vague descriptions, with the arrow always indicating the object in the middle of the picture, and the other half used the 3 pictures built for the presuppositional descriptions in the critical contexts, i.e. pictures representing the stage preceding the described process. The descriptions in the fillers involved no vague adjective and no presuppositional verb. Four of them used affirmative sentences and the other two used negative sentences. The order of presentation of these 42 items was randomized for each subject.

### 5.1.2.2 Procedure and Participants

The experiment was on-line and participants were recruited via the Amazon Mechanical Turk platform, where they were informed of the linguistic nature of the experiment. They were then redirected to a personal server where the material was hosted. Appendix A.1.1 reports the instructions given to the participants. The proper experiment started after they judged a practice item.<sup>5</sup> A "Next" button appeared when the participants clicked on one of the 5 buttons on the scale, letting them display the next item. At the end of the experimental session, participants had to enter their Amazon Mechanical Turk ID to validate their participation. 49 Amazon workers (no requirement specified) participated and were remunerated \$1.5 for an average time of 6 minutes and 53 seconds.

---

<sup>5</sup>The practice item can be found in Appendix A.1.3.

### 5.1.3 Results

Figures 5.6 and 5.7 present the repartitions of participants’ button clicks across **Context** and **Polarity** for each repetition (i.e. each set of pictures), respectively for the vague descriptions and for the presuppositional descriptions. 6 participants who scored less than 75% accuracy on filler items (unambiguous) were excluded.

These graphs show a clear overall effect of **Context**. One can also see an interaction between **Polarity** and **Context**: there was an effect of **Polarity** in the control contexts and arguably no effect of **Polarity** in the critical contexts; along with an interaction between **Description Type** and **Context**: there was no effect of **Description Type** in the control contexts but there is a clear difference between the vague and the presuppositional descriptions in the critical contexts. Participants’ answers were quite constant across repetition. However, the pictures that presented mid-sized squares described by a vague description yielded less contrasted answers; and the pictures that represented a monkey with an unpeeled banana described by a presuppositional description flattened the distribution of answers while they possibly revealed an effect of **Polarity**.

### 5.1.4 Discussion

First of all, the clear and consistent results for the control contexts show that participants understood the task and behaved as expected. However, looking at the results for the presuppositional descriptions, some participants sometimes behaved as if there was no negation in the *negative* presuppositional descriptions. Indeed, for these descriptions, there was a non-negligible rate of clicks on the rightmost button (“Completely true”) in the *instance* contexts and on the leftmost button (“Completely false”) in the *counter-instance* contexts, i.e. some answers patterned with the distribution for the *affirmative* presuppositional descriptions. Double negations are known to be hard to process, and it might be that the verb *stop* was processed as a negation of the occurrence of a process. If this analysis is correct, the apparent ignorance of negation might have resulted from the process of negating a negation-flavored predicate, thus explaining the clicks patterning with the affirmative presuppositional descriptions.

#### 5.1.4.1 Vagueness

The results for the vague descriptions are mostly compatible with SH. In the control contexts, participants behaved as expected: **Context** and **Polarity** interacted

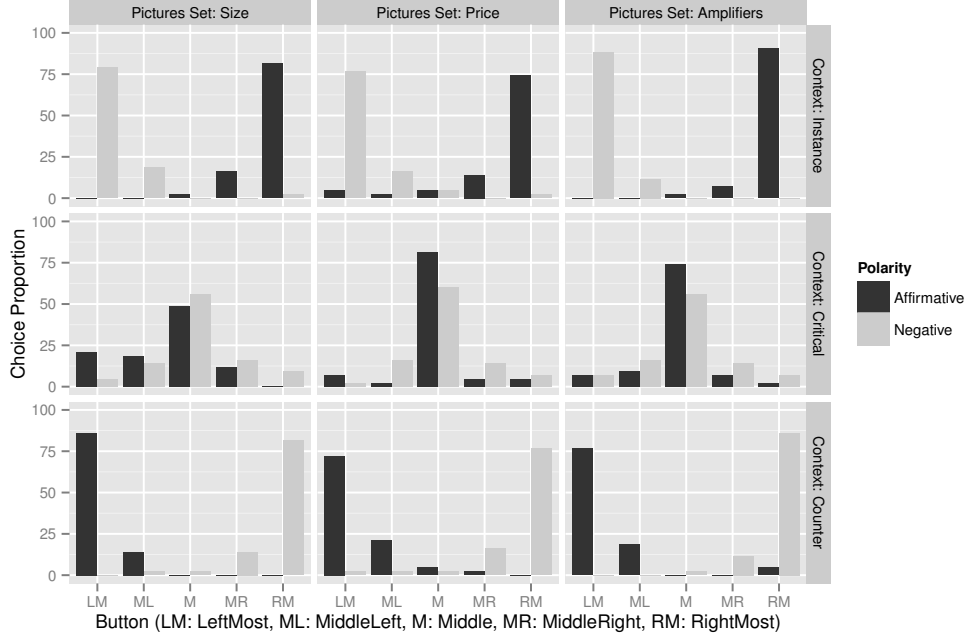


Figure 5.6: The observed percentages of clicks on each button for the *vague* descriptions in Experiment 1. The left-most button was adjacent to the text *completely false*, and the right-most button was adjacent to the text *completely true*. The three columns correspond to the three tested adjectives (*big*, *expensive*, *loud*). The middle row corresponds to the critical contexts where the pictures depicted borderline cases for the corresponding adjective.

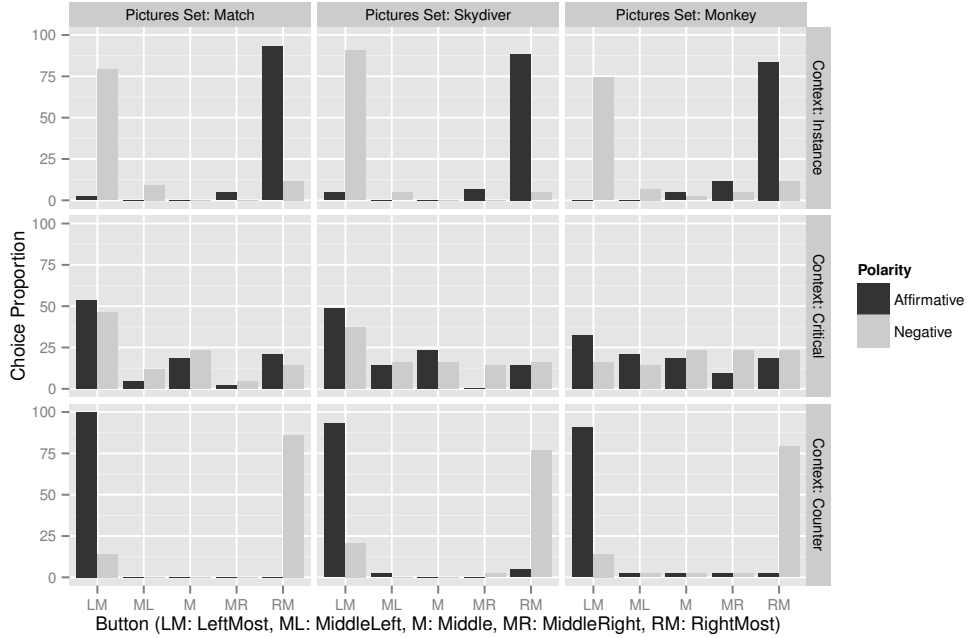


Figure 5.7: The observed percentages of clicks on each button for the *presuppositional* descriptions in Experiment 1. The left-most button was adjacent to the text *completely false*, and the right-most button was adjacent to the text *completely true*. The three columns correspond to the three tested expressions (*stop burning*, *stop falling*, *stop eating*). The middle row correspond to the critical contexts where the pictures depicted situations where the corresponding event had not even started.

exactly as predicted by SH(i). In the critical contexts, participants showed a clear preference for the middle-button for either polarity: this is what SH(ii) predicted, and the effect of **Polarity**, if any, appears to be trivial. The results for the square-pictures in the critical contexts were flatter though and there was a small tendency toward left (false-oriented) buttons for the affirmative descriptions. With the other pictures, the distribution of clicks in the critical contexts was slightly flatter for the *negative* vague descriptions than for the *affirmative* vague descriptions: there were fewer clicks on the middle-button, and more clicks on the middle-right and on the middle-left buttons. This may reflect more hesitation in participants' judging the negative vague descriptions, but we would still have to explain where this hesitation came from. A possible explanation would resort to a threefold ambiguity of negation. Under this view, *A is not ADJ* would be ambiguous between a paraphrase as *A is clearly not ADJ*, a paraphrase as *A is not clearly ADJ* and a paraphrase where there is no clearness consideration involved. When describing borderline cases for *ADJ*, the first reading would be on the false side, the second reading would be on the true side and the third reading would be as true as false. To explain the majority of clicks on the middle-button, this explanation would need to add that the third reading is easier to access. Importantly, positing this *ambiguity of negation* is different from positing a possible resort to a covert *clearly* operator that could scope over or under negation: a covert operator analysis would derive an ambiguity for the affirmative vague descriptions too (depending on the presence/absence of the operator), and would therefore predict a tendency toward “*falsish*” answers for this polarity. Interestingly enough, the predictions of the first option (positing an ambiguity of negation of vague predicates) fit the results well for the *Price* and *Amplifiers* items, and the predictions of the second option (positing the existence of a covert *clearly* operator) fit the results well for the *Size* items. These questions definitely deserve more attention, and it would be interesting to investigate more deeply the asymmetry in the judgments for affirmative and negative vague descriptions of borderline cases.

#### 5.1.4.2 Presupposition

The results for the presuppositional descriptions in the control contexts conform to SH(i). However, they clearly invalidate SH(iii). But even though SH(iii) is now ruled out, one might wonder whether any of WH(i-iv) is confirmed by these results. WH(i) predicts a global effect of **Context**, and this is clearly observed. However, WH(i) more precisely predicts that we should observe a majority of clicks on non-extreme

buttons in the critical contexts, and this is not what we observe *for the presuppositional descriptions*: there is a relative majority of clicks on the leftmost (“Completely false”) button (the majority is even absolute in the *Match* items). WH(ii), predicting no effect of polarity for vagueness in the critical conditions, necessarily obtains because it is entailed by SH(ii) which obtains (participants clicked the middle button for both affirmative and negative vague descriptions of borderline cases). WH(iii) predicts an effect of **Polarity** in the critical contexts, namely that we should observe more clicks on the left for the affirmative descriptions and more clicks on the right for the negative descriptions. This effect is clearly not what we observe. However, the results conform to WH(iv): there is an interaction between **Description Type** and **Polarity** in the critical contexts. Indeed, the buttons that participants clicked for the *affirmative presuppositional* descriptions lie more on the *left* than the buttons they clicked for the *vague* descriptions of *either polarity*. However, the buttons they clicked for the *negative presuppositional* descriptions do not lie more on the *right* than the buttons they clicked for the *vague* descriptions of *either polarity*, to the extent that participants produced similar pattern of answers for either polarity of the presuppositional descriptions in the critical contexts.

Two aspects of the results for the presuppositional descriptions in the critical contexts call for an explanation: first, why do we have an overall flat distribution slightly oriented to the left; and second, why don’t we observe an effect of **Polarity**? At this point, it should be noted that the results for the critical contexts associated with the *Monkey* items, illustrated in Fig. 5.8, are slightly different from the ones obtained for the two other sets of items: they are flatter and do suggest an effect of **Polarity**. I will put them aside for now, try to explain the results for the two other sets of pictures and come back on the *Monkey* set afterwards, in the light of the proposed explanations.



Figure 5.8: The picture used to test the description “The monkey has (not) stopped eating” in the critical contexts.

**Theoretical Explanations** There is now a widespread position claiming that when confronted to an affirmative sentence whose presupposition is unfulfilled, speakers may give a “false” truth-value judgment.<sup>6</sup> Bivalent systems that assign the value 0 to propositions with unfulfilled presuppositions would directly account for this tendency with the affirmative descriptions, but they would not account for the tendency found here with the negative descriptions, assuming that truth-functional negation applies there to the proposition with the unfulfilled presupposition. Trivalent systems equipped with an operator that maps propositions of the third value to propositions of value 0 do predict that both the affirmative and the negative presuppositional descriptions could yield “Completely false” truth-value judgments, but they also predict an effect of **Polarity** if the operator can scope under negation, to the extent that this would result in negative propositions of value 1 and therefore to “Completely true” truth-value judgments for the negative presuppositional descriptions.

Perhaps the most neutral position that would not be really threatened by these results is simply to say that a relative majority of participants chose the leftmost button as a way of signaling that they rejected the descriptions on the ground that their presuppositions were unfulfilled, regardless of their status toward the logical truth-values 0 and 1. Advocates of a trivalent approach of presupposition could invoke the operator to explain the rightmost clicks with the negative presuppositional descriptions. Note however that the neutral position is also compatible with ST5, which directly accounts for these clicks. The main threat now lies in the existence of rightmost clicks with the *affirmative* presuppositional descriptions.

A possible explanation for this lies in the already mentioned process of global accommodation. The idea of global accommodation is that, when speakers are presented with a sentence of which they ignore whether the presupposition is fulfilled or unfulfilled, they tend to assume that the presupposition *is fulfilled* for the sake of communication and/or computation. By resorting to global accommodation, participants would then have given their judgment regarding the assertive content of the description. If we consider the assertive part of sentences of the form “X has stopped V-ing” to be something like “X is not currently V-ing”, then the assertive parts of the presuppositional descriptions are true in the critical contexts, and this would motivate a click on the rightmost button. This explanation can account for the observation

---

<sup>6</sup>(Russell 1905) radically endorsed this position, but (Lasnik 1993) and (von Stechow 2004) more recently also allowed for such truth-value judgments in situations of presupposition failure. As discussed below, semantic approaches of presupposition that posit an accommodation operator as discussed in Chapt. 2 (again, see (Beaver 2001) for a discussion of the accommodation operator in semantic theories of presupposition) can also account for this kind of truth-value judgments.

of clicks on the rightmost button for the *Skydiver* and the *Monkey* pictures in the critical contexts. When they saw the picture of a skydiver on a plane along with the description “The skydiver has stopped falling”, participants could have gone through the following reasoning: the skydiver must have fallen before, he came back on the plane and he is not currently falling, so the description is true. Similarly, when they saw the picture of a monkey with an unpeeled banana along with the description “The monkey has stopped eating”, participants could have gone through the following reasoning: the monkey must have eaten something else before, it found this banana and it is not currently eating, so the description is true. But this explanation runs into problems with the picture of an unstruck match: participants could not consider a situation where the match had burnt before, because in such a situation the match precisely should *not* look unstruck.

At this point, a possibility would be to posit that *stop* is in fact ambiguous between a presuppositional and a non-presuppositional reading. We could then explain the results for the presuppositional descriptions in the critical contexts in the following way. First, regardless of the polarity of a presuppositional sentence, when the presupposition is known to be unfulfilled, there would be a strong tendency to reject it as false and a weaker tendency to consider the sentence neither true nor false. This would account for the clicks on the leftmost and the middle buttons in either polarity of the descriptions. Second, there would be a non-presuppositional reading of *stop* which would simply mean *is not currently*. This would account for the clicks on the rightmost button for the affirmative descriptions, but it would also predict more clicks on the leftmost button for the negative descriptions than for the affirmative descriptions.<sup>7</sup> Recall, however, that on the one hand advocates of trivalence would *also* assume that there is an operator that maps propositions of non-bivalent values to 0. This would be necessary to account for the clicks on the rightmost button when the description is negative: resorting to this strategy would yield true negative sentences when the operator scopes under negation. On the other hand, ST5 directly predicts the observation of clicks on buttons on the right for the negative descriptions (associated with the value  $\mathcal{P}^1$ ).

This multiplication of non-homogeneous explanations might make one suspicious of the experiment as a whole. A reasonable course at this point might be to look for flaws in the design of the experiment that could have served to conceal the true generalizations about the treatment of vague and presuppositional sentences. Indeed, it could well be that a general explanation like the neutral position is right, but that

---

<sup>7</sup>Though remember the comment above on the processing of double negations.

flaws in the design added some noise to the pattern of answers for the presuppositional descriptions.

**Questioning the Design** One possibility might be that participants developed a strategy during the experiment: maybe they did not know which button to click at the beginning of the experiment when confronted with the presuppositional descriptions in the critical contexts and therefore randomly clicked, but then they would have understood the task better and consistently chosen the leftmost button in these contexts. However when we consider only the first quartile of answers reported in Fig. 5.9, we still observe the same tendency toward the leftmost button and also a non-trivial proportion of clicks on the rightmost button, for either polarity.

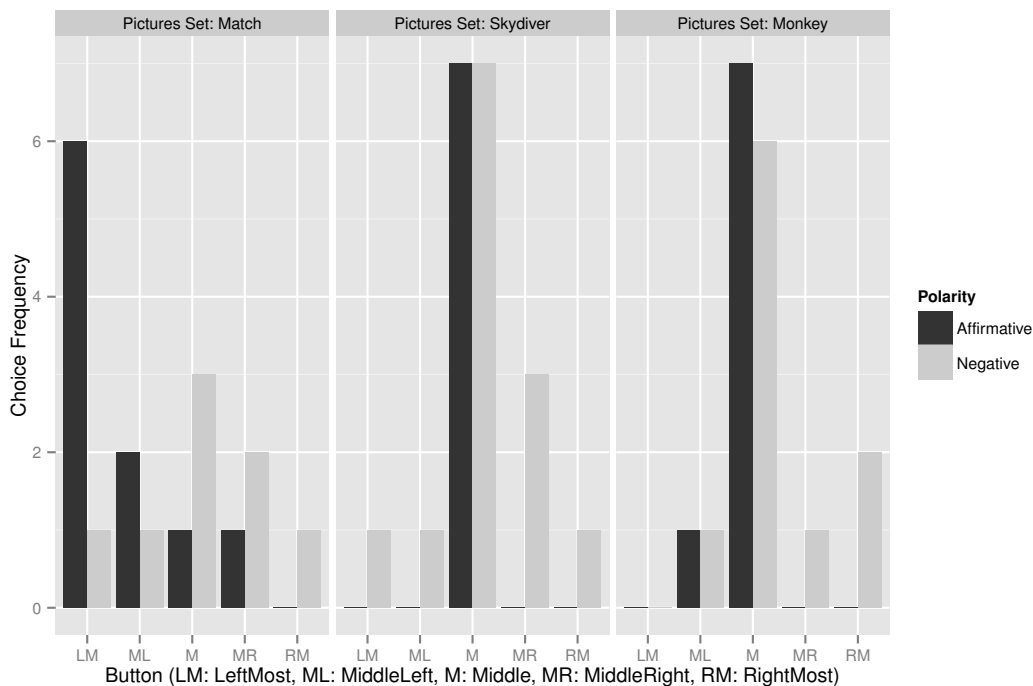


Figure 5.9: The repartition of absolute numbers of clicks for the presuppositional descriptions in the *critical* contexts in the first quartile. The left-most button was adjacent to the text *completely false*, and the right-most button was adjacent to the text *completely true*. The three columns correspond to the three tested expressions (*stop burning*, *stop falling*, *stop eating*).

Another possibility would be that the observed distributions result from different participant profiles. When we look at participants answers individually though, we see that very few participants systematically gave the same judgments for the three repetitions of the same condition, regardless of the polarity. Rather, participants who clicked the rightmost button (“Completely true”) on one trial generally clicked



another button on the other two trials, and more often than not the button they clicked was the leftmost one (“Completely false”).

A last possible analysis would be that each participant could adopt two strategies when confronted to a presupposition failure: either they clicked the leftmost button to convey that using the description in the critical context is incorrect, or they did not even map the description to a truth-value and therefore answered randomly on the scale.

These last considerations call for a new design that would provide us with clearer truth-value judgments distributions in cases of presupposition failure. In particular, the new design should block the strategy of clicking randomly to signal a presupposition failure and it should prevent any process of global accommodation. This experiment made it clear, though, that participants did not treat vagueness nor presupposition classically. In addition, the distributions of clicks for the vague descriptions and for the presuppositional descriptions in the critical contexts were clearly different. These results show that participants can exhibit non-bivalent behavior both for vagueness and presupposition, while differing in what non-bivalent judgments they might give for each type of sentences.

## 5.2 Second Experiment: A 3-Response Paradigm

The second experiment was designed in an effort to provide a protocol which would lead clear and systematic different truth-value judgments for vague and presuppositional sentences. For this reason, it was not designed with the primary aim of testing the predictions of ST5. In the meantime, I got acquainted with a series of experiments by (Križ & Chemla 2014) investigating truth-value gaps observed in situations of non-homogeneity. They obtained very clear results, establishing the validity of their designs. This second experiment was therefore designed with their “one shot ternary judgments” (*sic*) method.

### 5.2.1 Design

The 5-point scale of the first experiment was replaced here with a threefold choice between “Completely false”, “Completely true” and “Neither”. (Križ & Chemla 2014)’s results show that when presented with these options, subjects readily click on “Neither” to signal a truth-value gap. Therefore, I expected that in situations of presupposition failure participants would prefer to click on “Neither” rather than to click randomly to signal a lack of truth-value. Besides, in contrast with the previous

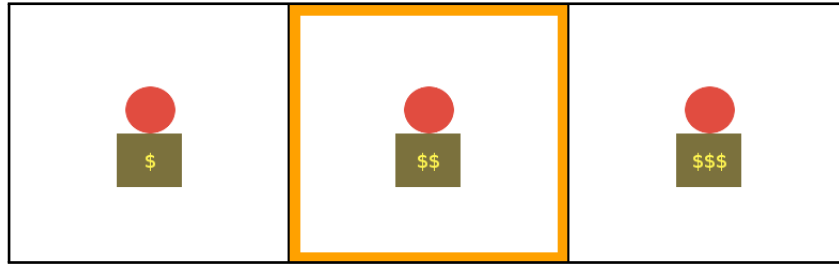
experiment where the radio buttons were not associated with labels, I suspected that labeling each button would help in identifying its meaning and in understanding the task. In an effort to prevent participants from resorting to global accommodation, the pictures associated with the presuppositional descriptions were modified to represent a *sequence* of three frames describing an event: the first frame showed the scene before the event taking place, the second frame showed the scene while the event was taking place and the last frame showed the scene after the event took place. Being presented with sequences of whole processes, subjects were expected to naturally treat the first frame as describing a situation where the event had not started yet. Accordingly, the pictures associated with the vague descriptions were modified to represent three frames, each containing one of the three objects present in the pictures of the previous experiment. In each condition, one of the frames was boxed and the description was to be evaluated against this frame. Thus, the context was defined by the boxed frame. For the affirmative presuppositional descriptions, “Completely true” answers were therefore expected when the last frame (the event being over) was boxed, “Completely false” answers were expected when the second frame (the event taking place) was boxed,<sup>8</sup> and “Neither” answers were expected when the first frame (before the event) was boxed. The frames for presupposition were always presented in this natural order for the reasons detailed above. The frames for vagueness were however still randomly ordered. The six sets of pictures (three sets by type of description) were used again in the hope that the modification described above would cancel the undesired specificity of the *Monkey* set, what turned out to be the case. Two sets of pictures reported in Appendix A.2.2 were added for each type of descriptions. This eventually led to 5 repetitions by condition. Each description was again presented both in the affirmative and in the negative polarity. As in the first experiment, this was a  $2 \times 2 \times 3$  interaction design, where the factors were **Description Type** (vague vs. presuppositional), **Polarity** (affirmative vs. negative) and **Context** (critical vs. instance-oriented vs. counter-instance-oriented). Figures 5.10 and 5.11 exemplify the task in the second experiment.

### 5.2.1.1 Predictions

Under the hypothesis that vagueness and presupposition yield non-bivalent truth-value judgments in borderline cases and in situations of presupposition failure, one

---

<sup>8</sup>It could be argued that showing the event currently taking place is not showing a situation where the event is obviously taking place *before*. However these frames described situations where the event had necessarily started before. In the end, it turned out that participants did answer as expected.



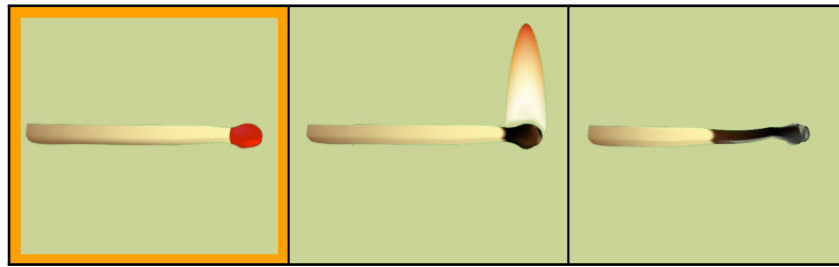
The object is not expensive

Completely true

Neither

Completely false

Figure 5.10: Example of a negative vague description in a critical context. The description was to be evaluated against the boxed frame, here targeting a borderline case for *expensive*. Subjects could click on one of the “Completely false”, “Neither” or “Completely true” buttons, which were always displayed in this order. The order of the frames was randomly pre-determined for the vague items.



The match has stopped burning

Completely true

Neither

Completely false

Figure 5.11: Example of an affirmative presuppositional description in a critical context. The description was to be evaluated against the boxed frame, here targeting a situation making the presupposition unfulfilled. Subjects could click on one of the “Completely false”, “Neither” or “Completely true” buttons, which were always displayed in this order. The order of the frames was always the same for the presuppositional items, to make clear that they entered in a temporal sequence.

prediction was that (i) we should observe an important rate of “Neither” answers both for the vague and for the presuppositional descriptions for either polarity in the critical contexts, as opposed to the control contexts. Under the hypothesis that vagueness and presupposition do not yield the *same* non-bivalent truth-value judgments in these contexts, another prediction was that (ii) for either polarity in the critical contexts, the distribution of answers for the vague descriptions should be different from the distribution of answers for the presuppositional descriptions.

Given the results of the first experiment, it was expected that (iii) in the critical contexts and for either polarity, participants would mostly answer “Neither” to vague descriptions and (iv) would distribute their answers over the three possibilities for presuppositional descriptions. However, if the new design succeeded in blocking the *random* strategy, and under the assumption that participants would not assign any truth-value to the descriptions whose presuppositions were unfulfilled, it was expected that (iv’) in the critical contexts, participants would divide their answers between “Neither” to signal a truth-value gap and “Completely false” to signal the rejection of an unfelicitous description, regardless of its polarity.

## **5.2.2 Methods**

### **5.2.2.1 Materials**

As mentioned in the previous section, there were 5 sets of pictures for the vague descriptions and again 5 sets of pictures for the presuppositional descriptions, yielding 5 measures for each of the 12 conditions described above. In particular, each subject was presented with 5 affirmative and 5 negative descriptions for the vague sentences and again with 5 affirmative and 5 negative descriptions for the presuppositional sentences in the critical and in the control contexts. In addition to these 60 items were included 8 fillers, ensuring participants’ good understanding of the task. Half of the fillers used 4 of the 5 pictures built for the vague descriptions, with the boxed frame always being the middle one, and the other half used 4 of the 5 pictures built for the presuppositional descriptions, with the boxed frame always being the first one (i.e. before the event taking place). The descriptions in the fillers involved no vague adjective and no presuppositional verb. Half of them used affirmative sentences and the other half used negative sentences. The order of presentation of these 68 items was randomized for each subject.

### **5.2.2.2 Procedure and Participants**

The second experiment was on-line too and participants were again recruited via Amazon Mechanical Turk platform, where they were identically informed of the linguistic nature of the experiment. This time, they were redirected to the Ibex Farm servers where the material was hosted. The experiment was implemented with the Ibex software. Appendix A.2.1 reports the instructions given to the participants. The task divided in two sessions: first a training session of two non problematic items for which participants received feedback (i.e. they were told whether their judgment

was correct or incorrect) and then the experimental session itself.<sup>9</sup> At the end of the experimental session, participants had to enter their Amazon Mechanical Turk ID and to indicate whether they were native speakers of English to validate their participation. They could also indicate their sex (Male or Female), their age and leave a comment.

50 Amazon workers (no requirement specified) participated and received \$1.0 as a retribution for an average time of 9 minutes. 2 participants were excluded from the analyses because they already took part in the first experiment. Out of the 48 remaining participants, 24 identified as females, 22 identified as males and 2 did not provide the information; and 41 participants defined themselves as native speakers of English. Collected ages ranked from 20 to 60.

### 5.2.3 Results

Figures 5.12 and 5.13 present the repartitions of participants’ button clicks across Context and Polarity for each repetition (i.e. each set of pictures), respectively for the vague descriptions and for the presuppositional descriptions. 7 participants who scored less than 75% accuracy on 7 of the 8 filler items (unambiguous) were excluded.<sup>10</sup>

Here again, there was a clear overall effect of **Context**. In the control contexts, there was a clear effect of **Polarity** and subjects barely ever clicked the “Neither” button for any picture. In the critical contexts, there was apparently no effect of **Polarity** and participants mostly clicked the “Neither” button for every repetition of the vague descriptions. In contrast, and based on the distribution of clicks for the presuppositional descriptions, the two picture sets introduced in this experiment seem to group apart from the picture sets already present in the first experiment. **Polarity** seems to have had an effect for the two new sets of pictures and possibly a minor effect in the critical contexts for the three pictures from experiment 1. In these contexts, participants mainly distributed their clicks over “Completely false” and “Neither” for either polarity of the presuppositional descriptions with the three latter pictures (except for the negative descriptions of the *Monkey* items) whereas the clicks for the new picture sets patterned closer to those in the instance contexts.

---

<sup>9</sup>The trial materials can be found in Appendix A.2.3

<sup>10</sup>One of the filler items proved to be ambiguous, with participants evenly distributing their clicks over the three buttons.

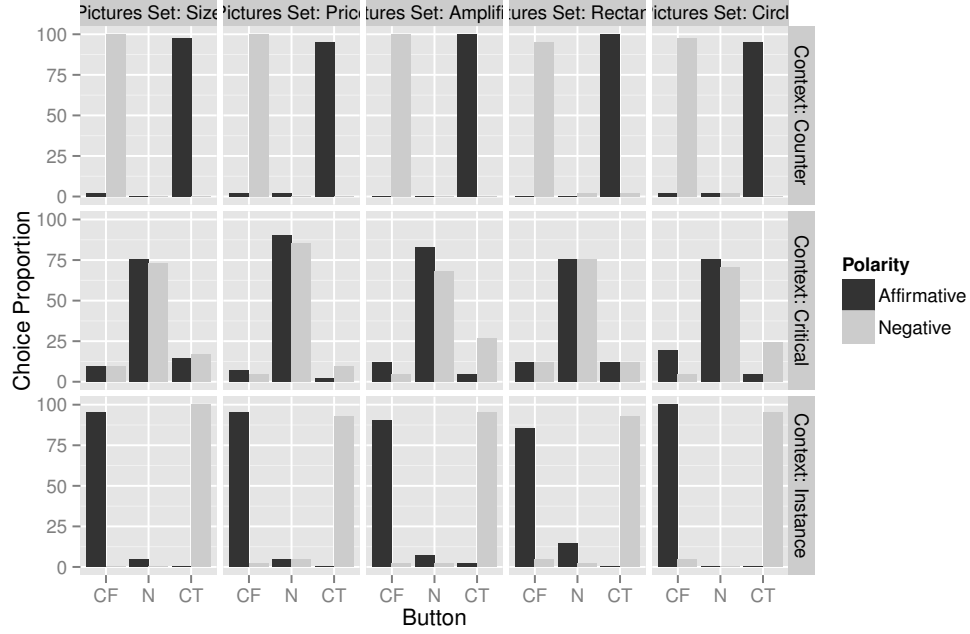


Figure 5.12: The observed percentages of clicks on each button for the *vague* descriptions in Experiment 2. *CF* stands for the button *Completely false*, *N* for the button *Neither*, and *CT* for the button *Completely true*. The five columns correspond to the five tested adjectives (*big*, *expensive*, *loud*, *wide* and *close*). The middle row corresponds to the critical contexts where the framed box of the pictures depicted borderline cases for the corresponding adjective.

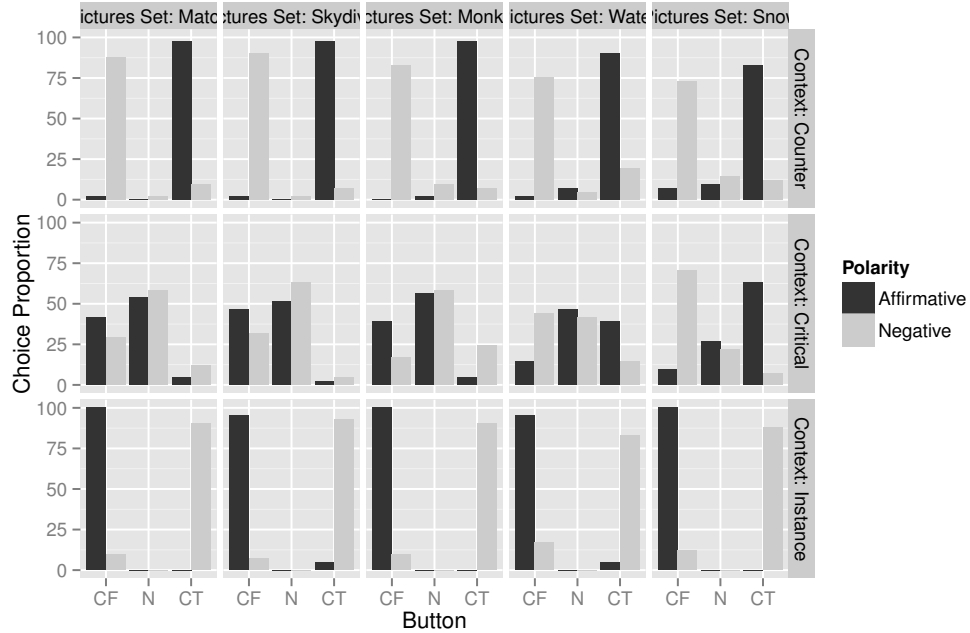


Figure 5.13: The observed percentages of clicks on each button for the *presuppositional* descriptions in Experiment 2. *CF* stands for the button *Completely false*, *N* for the button *Neither*, and *CT* for the button *Completely true*. The five columns correspond to the five tested expressions (*stop burning*, *stop falling*, *stop eating*, *stop flowing* and *stop snowing*). The middle row corresponds to the critical contexts where the framed box of the pictures (the first frame) depicted situations where the event had not even started.

## 5.2.4 Discussion

Prediction (i) was borne out to the extent that the distributions of clicks in the critical contexts were clearly different from the distributions of clicks in the control contexts. In addition, to the extent that the distributions of clicks in the control contexts for the presuppositional descriptions were clearly different from the distributions of clicks for the vague descriptions for either polarity, prediction (ii) was also borne out. Expectation (iii) was met: participants mainly clicked “Neither” for either polarity of the vague descriptions in the critical contexts. Expectation (iv) was not met: in the critical contexts, participants barely ever clicked the “Completely true” button for the affirmative presuppositional descriptions of the pictures from experiment 1. However expectation (iv’) was met for these pictures: this suggests that the new design successfully blocked the *random* strategy.

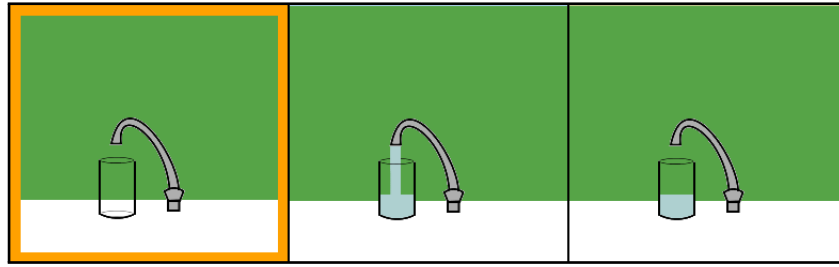
It appears that a non-negligible amount of participants clicked on the “Completely true” button when judging the negative description of the *Monkey* picture. The reason for this could be that the picture describing the critical context did not make it clear enough that the monkey was *not* eating the banana. Participants might have imagined that the picture was showing a monkey about to open and eat the banana, and they might have therefore considered that the event of eating had already started. In that case, the description would be true: “the monkey has not stopped eating: it has just started”.

Given that participants did not distribute their clicks similarly when judging the presuppositional descriptions in the critical contexts against the two new pictures, there must be something special about them. These pictures are represented in Fig. 5.14 and Fig. 5.15.



It has (not) stopped snowing

Figure 5.14: The presuppositional item corresponding to the critical conditions of the affirmative and negative presuppositional descriptions for the expression *stop snowing*. The unstruck match targets a situations where the event of burning has not started yet.



The water has (not) stopped flowing

Figure 5.15: The presuppositional item corresponding to the critical conditions of the affirmative and negative presuppositional descriptions for the expression *stop flowing*. The empty glass under the closed faucet targets a situation where the event of the water flowing has not started yet.

Again, my proposal is that global accommodation was available for these pictures even though sequencing ruled it out for the three pictures coming from the experiment 1. I suspect this is due to the very cyclic nature of the described events: a faucet is made for being repetitively used, so it is very natural, when confronted to the picture in Fig. 5.15, to assume that it had already been used before the first frame and that water has therefore stopped flowing since the last use.<sup>11</sup> In a somehow similar way, weather is a cyclic thing: an area has a usual climate pattern. When confronted to the picture in Fig. 5.14, one might reasonably assume, considering the last two frames, that in the region of the presented landscape it snows during winter. But the first frame shows no sign of snow: then it seems to have stopped snowing for quite a long time, the winter might be over and a new spring has come.

In contrast, we use a match just once: there's definitely no consideration of cyclic-ity going on when judging the sentence "The match has (not) stopped burning"; and skydiving is a rather exceptional event: it would not be very reasonable to assume that someone who is in a skydiving suit on a plane has already jumped, then went back on the plane and is now waiting for the landing.

In order to verify these conjectures, I therefore conducted a follow-up experiment study where I asked subjects to provide justifications for their answers to the presuppositional descriptions.

<sup>11</sup>In addition, one might understand "The water has stopped flowing" as suggesting that the faucet used to leak but does not anymore: this information would then be confirmed by the fact that the glass under the faucet is empty.



## 5.3 Follow-Up Experiment: Investigating Participants' Motivations for Presupposition

The aim of this follow-up experiment was to investigate the motivations for accepting the affirmative presuppositional descriptions of the *Snow* and *Water* pictures as true in the critical contexts. The hope was to show that subjects treated these pictures differently, namely that they went through a process of global accommodation when asked to judge the presuppositional descriptions for these pictures. The follow-up experiment was therefore designed to collect subjects' motivations for the truth-value judgments they gave.

### 5.3.1 Design

I used the same materials as in the previous experiment. All the vague and negative descriptions were removed, though, for the experiment not to get too long. A screen asking subjects to indicate their motivations by checking one or several boxes in a given list was inserted after each presuppositional description. On that screen, subjects were shown the picture and the description they just evaluated and they were reminded of the button they clicked. For each picture, there were 4 suggested reasons in the list of the following form:

X was V-ing before	X was not V-ing before
X is currently V-ing	X is not currently V-ing

In addition, a text-box appeared if the subjects checked *Other*, allowing them to provide their own explanations. Figure 5.16 exemplifies such a screen.

#### 5.3.1.1 Predictions

The results of the previous experiment for the affirmative presuppositional descriptions were expected to be replicated. Concerning the results for the motivations, subjects were expected to motivate a click on the “Neither” or the “Completely false” button in the critical contexts by checking “X was not V-ing before”, indicating a case of presupposition failure. They were expected to motivate a click on the “Completely true” button in these contexts by checking “X was V-ing before” and “X is not currently V-ing”, indicating that they globally accommodated the presupposition.



*It has stopped snowing*

You answered **Completely true**. Why?

Please check the reason(s) that best fit your motivations:

- |  |  |
|--|--|
| <input type="checkbox"/> The snow was falling before     | <input type="checkbox"/> The snow is currently falling     |
| <input type="checkbox"/> The snow was not falling before | <input type="checkbox"/> The snow is not currently falling |
| <input type="checkbox"/> Other                           |  |

Figure 5.16: The motivation screen that appeared when a subject clicked on the button *Completely true* after judging the *affirmative presuppositional* description *It had stopped snowing* in a *critical* context. Subjects had to check one or several of the listed motivations, and had the possibility to check *Other*, in which case a text field appeared where they could explain their motivation.

## 5.3.2 Methods

### 5.3.2.1 Materials

Target items consisted in sequences of two screens. The set of first screens was the set of the 5 affirmative presuppositional descriptions coming from experiment 2, presented in each of the 3 contexts. The set of second screens consisted in the corresponding motivation screens, as illustrated in Fig. 5.16. In addition, participants saw the 4 affirmative fillers coming from experiment 2. There was no motivation screen for the fillers. The order of presentation of these 19 items was randomized for each subject.

### 5.3.2.2 Procedure and Participants

Participants were again recruited via Amazon Mechanical Turk and redirected to the Ibex Farm servers. The following line was added in the instructions:

For some descriptions, you will be asked to inform us on the justifications for your judgment. In those cases, you will be presented with a list of several possible reasons. Please check those that you feel best correspond to your motivations for the judgment you gave. There is no right or wrong answer, we are interested in what you think.

The first item of the training session contained no motivation screen. The second item presented a first screen with a picture of a square along with a sentence describing

it as a circle, followed by a second screen asking the subjects to check one or several motivations in the following list for the truth-value judgments they gave:

The figure has four sides                      The figure is round  
The figure does not have four sides      The figure is not round

40 Amazon workers (no requirement specified) participated and received \$0.5 as a retribution for an average time of 4 minutes and 58 seconds. 1 participant was excluded from the analyses for having participated to experiment 2. Out of the 39 remaining participants, 16 identified as females, 21 identified as males and 2 did not provide the information; and 32 defined themselves as native speakers of English. Collected ages ranked from 19 to 58.

### 5.3.3 Results

Quite surprisingly, as can be seen in Fig. 5.17, the results of experiment 2 were not replicated: participants mostly clicked the “Completely false” button for 4 pictures out of 5 in the critical contexts. Nonetheless, only the *Snow* and *Water* pictures revealed a non-negligible rate of clicks on “Completely true” in the critical contexts.

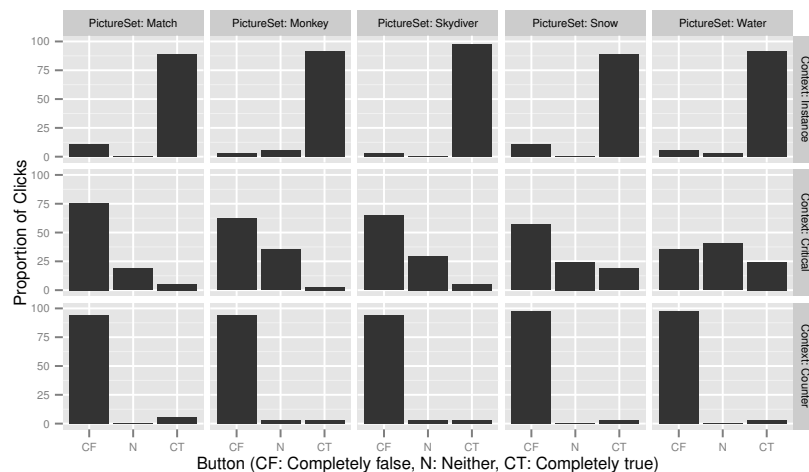


Figure 5.17: The observed percentages of clicks on each button in the follow-up experiment. *CF* stands for the button *Completely false*, *N* for the button *Neither*, *CT* for the button *Completely true*. The five columns correspond to the three tested expressions (*stop burning*, *stop eating*, *stop falling*, *stop flowing* and *stop snowing*). Note that the follow-up experiment contained only affirmative descriptions. The middle row corresponds to the critical contexts where the framed box of the pictures targeted a situation where the event had not even started.

Figure 5.18 shows the percentages of times subjects checked each reason depending on what judgment they gave. Cell percentages do not sum up to 100 because subjects

had the possibility to check several reasons at the same time. As an indication, light gray bars report the corresponding absolute number of checks, from 0 to 177. The *Reason* labels are to be mapped as follows:

- *WasBef*  $\rightarrow$  “X was V-ing before”
- *WasNotBef*  $\rightarrow$  “X was not V-ing before”
- *IsNow*  $\rightarrow$  “X is currently V-ing”
- *IsNotNow*  $\rightarrow$  “X is not currently V-ing”

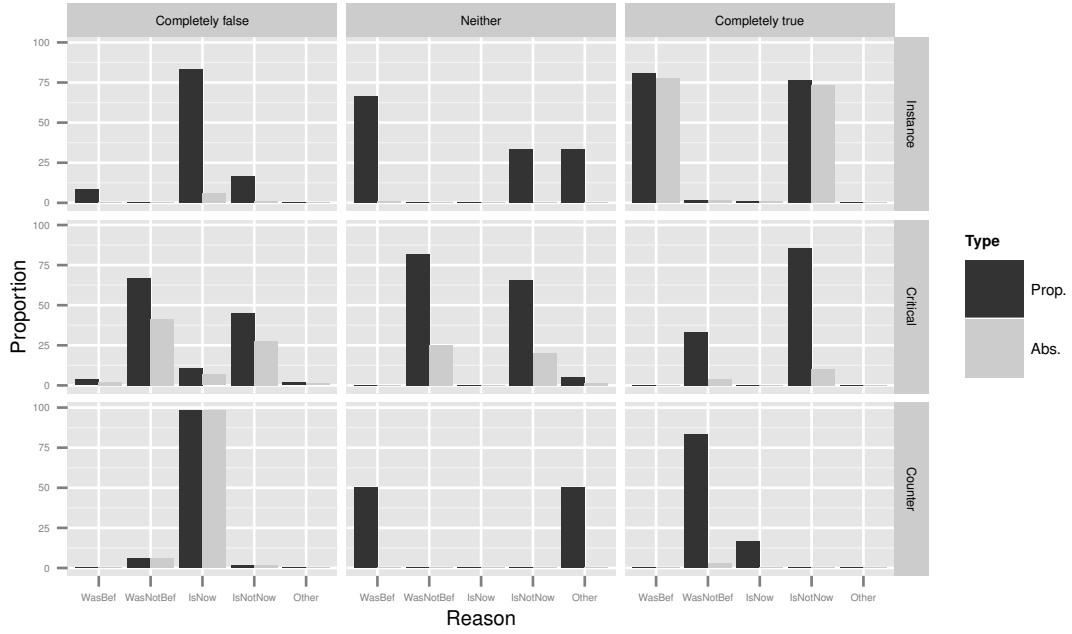


Figure 5.18: The observed distribution of clicks on each motivation in the follow-up experiment. Light gray bars report the absolute number of clicks, and dark gray bars report the percentages of clicks for the condition corresponding to the cell. Note that because participants had the possibility to check several motivations, the percentages in each cell can sum up to more than 100%. The left column corresponds to the motivations that the participants checked after giving a *Completely false* judgment, the middle column corresponds to the motivations that the participants checked after giving a *Neither* judgment and the right column corresponds to the motivations that the participants checked after giving a *Completely true* judgment. The top and bottom rows correspond to the motivations that the participants checked after judging a control item and the middle row corresponds to the motivations that the participants checked after judging a *critical* item (where the first box was framed).

Importantly, no participant checked “X was V-ing before” after judging a presuppositional description “Completely true” in a critical context. The majority of the few clicks on “Completely true” in these conditions were associated with the reason “X is not currently V-ing”.

### 5.3.4 Discussion

It seems that the modification in the design strongly favored clicks on “Completely false” in the critical contexts. This might be due to the quasi-absence of non-presuppositional items (there were only 4 fillers). Being presented with only one kind of items, participants might have developed a strategy of systematically clicking “Completely false” when the first frame was boxed. But this is not very plausible, given that the *Water* pictures resulted in a different clicks distribution. More probably, this overall tendency might be due to the insertion of the motivation screens. Having the possibility of providing a justification for their “Completely false” judgments in the critical contexts might have encouraged participants to give this judgment even though they felt the descriptions in these contexts to have a status different from their status in the control false contexts. The reason why participants were more reluctant to click the “Completely false” button in the previous experiments might be that they did not want to commit into accepting the truth of the presupposition. In this version of this experiment, they could make it clear that they held the presupposition to be false even when they clicked on “Completely false”.

When we focus on the two pictures that were problematic in the previous experiments, we observe a slightly higher rate of clicks on “Completely true” than what we observe for the three other pictures. Fig. 5.18 shows that in these conditions, no participant justified their choice by checking “X was V-ing before”. This is informative, given that participants did check this reason in the instance-contexts. This suggests that the reason why some participants clicked the “Completely true” button in the critical contexts was not because they imagined a situation that made the critical context similar to an instance-context. In other words, participants might actually not have gone through a global accommodation process. These participants justified their choice by checking “X is not currently V-ing”. This is compatible with an ambiguity of *stop* between a presuppositional and a non-presuppositional reading, as described earlier. But this position does not explain why the *Snow* and the *Water* pictures in particular were treated differently. Unfortunately, no participant typed in any explanation for an “Other” motivation.

## 5.4 Conclusions

### 5.4.1 Experimental Considerations

The presuppositional descriptions, contrary to the vague descriptions, yielded results that are difficult to analyze. This difficulty might result from the availability of a process of accommodation. With the worry of preventing this process, I built a set of pictures which I thought described situations making global accommodation implausible. The analysis of the results of experiment 1 suggested that global accommodation was nonetheless readily adopted by the participants. I therefore brought some modifications in the presuppositional stimuli and in the design of experiment 2. I thought that presenting a sequence of pictures manifestly describing the different stages of an event, and pointing to the frame showing the moment preceding the beginning of the event would rule out any process of global accommodation. Besides, I borrowed (Križ & Chemla 2014)’s format of answers. These decisions seem to have been efficient for three out of five sets of pictures. Yet, the two problematic sets of pictures nonetheless yielded “Completely true” judgments in the critical contexts, suggesting that global accommodation was still available. The conclusion to be drawn from a practical point of view is that the process of global accommodation is very robust and consequently very special attention should be paid to the presuppositional stimuli presented to the participants. In particular, the presuppositional expression used in the description should make it possible to present uncontroversial contexts of presupposition failure.

In contrast, the stimuli for vagueness turned out to be very efficient both with the design of experiment 1 and with the design of experiment 2. When subjects are presented with three entities of different measures and are asked to judge a vague description of the entity whose measure situates in the middle, they readily reject both “Completely true” and “Completely false” judgments. Importantly, we can argue from the results of experiment 2 that (Križ & Chemla 2014)’s “one shot ternary judgments” which they claim to elicit truth-value gaps due to homogeneity also successfully elicit truth-value gaps due to vagueness and truth-value gaps due to presupposition.<sup>12</sup> It is worth noting that even though this threefold option appears limited compared to the five-point scale of experiment 1, it was still able to elicit the distinction between vagueness and presupposition in the critical contexts.

---

<sup>12</sup>Although the interpretation of the clicks on “Neither” in reaction to the vague descriptions can also be interpreted as reflecting a truth-value *glut*. See the discussion of vagueness in Chapters 2 and 3.

The results of the follow up study made it clear that its design, in its current state, cannot be used to efficiently elicit participants' motivations in the contexts of presupposition failure. Interestingly though, it shed some light on how the possibility for participants to justify their judgments could precisely influence the truth-value judgments that they would report.

Statistical analyses of the results still need to be conducted, but some patterns clearly emerge. The next section tries to draw general conclusion on their basis.

### 5.4.2 Theoretical Considerations

In both experiments, participants treated both vague and presuppositional descriptions differently in the critical contexts and in the control contexts. This result argues in favor of the non-bivalent aspect of the truth-value judgments specific to vagueness and presupposition. Experiment 2 made this aspect especially clear when we focus on the three pictures coming from experiment 1.

Overall, it seems that in the critical contexts, negation had no or very little effect regardless of **Description Type**. In particular, putting aside the problematic items, the affirmative and negative descriptions with an unfulfilled presupposition were associated with the same set of truth-value judgments, and “Completely true” judgments were clearly excluded from this set in experiment 2. These results echo the truth-value judgments that (Abrusán & Szendrői 2013) observed for sentences of the form of (70-a) and of its negative counterpart (70-b), which were both globally judged false to the extent that the definite descriptions did not refer.<sup>13</sup>

- (70)    a.    The king of France is bald  
           b.    The king of France is not bald

These are rather challenging results for several theories of presupposition. Trivalent theories that use an accommodation operator to account for the *false* judgments would have to prevent the operator from scoping under negation and yielding propositions of value 1 that would then be associated with unwelcome “True” judgments. Bivalent theories of presupposition predict negative descriptions with an unfulfilled presupposition to be semantically true, but one could imagine that some pragmatic principles, in the spirit of what (Abrusán & Szendrői 2013) imagine for Stalnakerian views of presupposition, prevent speakers from reporting a “True” judgment in these situations and make them prefer “False” and “Neither” judgments. Finally, these

---

<sup>13</sup>Even if, as noted in Chapter 3, subjects also gave a non negligible amount of “Can’t say” judgments for the negative counterparts, whereas they were quasi absent for the affirmative sentences.

results also challenge ST5. ST5 straightforwardly accounts for the “Neither” judgments, given that propositions of value  $\mathcal{P}^0$  and propositions of value  $\mathcal{P}^1$  are neither completely true nor completely false. The “False” judgments associated with negative presuppositional descriptions can also easily be accounted for once we revise the bridge principles so as to allow “False” judgments for any proposition of a value lower than 1, and so for propositions of value  $\mathcal{P}^1$  in particular.<sup>14</sup> However, ST5 does predict the observation of “True” judgments for negative sentences whose presuppositions are unfulfilled. We cannot revise the bridge principle for “True” to exclude propositions of values lower than 1, for we would then no longer account for glutty judgments associated with vagueness (i.e. propositions of value  $\mathcal{V}$ ). In the end, ST5 would need to be augmented with the same kind of pragmatic principles as those proposed for the bivalent approaches of presupposition.

One can imagine an alternative position in the spirit of (Strawson 1950), according to which participants do not even assign a truth-value to a sentence that suffers from presupposition failure. In that case, we would indeed expect no difference between the affirmative and negative counterparts discussed just above. But recall that we considered this position to explain the distribution of answers in the first experiment: we imagined an account on which participants either chose “False” to signal their disagreement or answered completely at random. Given that the results of the second experiment suggest that the “True” answers of the first experiment did not genuinely correspond to presupposition failures, we would have to revise this account. We would have to say that qualifying the sentence as “False” and as “Neither” are two possible ways of signaling a refusal and / or an inability to come up with a truth-value for the sentence.

The fact that participants gave different answers for vague descriptions and for presuppositional descriptions in the critical contexts suggests a different treatment of the two phenomena. From a truth-functional point of view, these results argue in favor of a system where, in crucial contexts: i) vague and presuppositional propositions are assigned different, specific truth-values; and ii) negation does not affect the truth-value of vague and presuppositional propositions. Consequently, the next chapter investigates the possibility of defining 4-valued systems which meet (i) and (ii).

---

<sup>14</sup>Note that in that case, one could say that a sentence is “False” if and only if one could say that a sentence is “Not true”.



## Chapter 6

# 4-Valued Systems for Vagueness and Presupposition

In Chapter 4, we motivated ST5 based on the following intuitions.

- a. **Non-Classicality:** both vague and presuppositional sentences trigger non-classical truth-value judgments in critical situations;
- b. **Distinction:** the set of non-classical truth-value judgments associated with vagueness is not the same as the set of non-classical truth-value judgments associated with presupposition;
- c. **Unpolarized Vagueness:** the affirmative and negative counterparts of a vague description of a borderline case trigger the same non-classical truth-value judgments;
- d. **Polarized Presupposition:** the affirmative and negative counterparts of a sentence with an unfulfilled presupposition trigger distinct sets of non-classical truth-value judgments

Importantly, the experiments discussed in the last chapter are compatible with **Non-Classicality**, **Distinction**, and **Unpolarized Vagueness** but they seem to contradict **Polarized Presupposition**. Indeed, in the critical contexts, the pattern of answers that the participants gave for the affirmative presuppositional descriptions was the same as the pattern of answers that the participants gave for the negative presuppositional descriptions. I will therefore favor **Unpolarized Presupposition**:

- e. **Unpolarized Presupposition:** the affirmative and negative counterparts of a sentence with an unfulfilled presupposition trigger the same non-classical truth-value judgments

I already showed in Chapt. 4, Sect. 4.2.2 why a *totally* ordered 4-valued system was unthinkable for the current problematic. However, the reasons why I excluded a partially ordered 4-valued alternative building on Dunn-Belnap are the following:

1. it assigns the same truth-value to positive and negative counterparts of propositions with unfulfilled presuppositions
2. it assigns 0 (ie. plain “false” judgments) to any hybrid conjunctions and 1 (ie. plain “true” judgments) to any hybrid disjunction

Because the previous experiments questioned **Polarized Presupposition**, the first of these two properties might in fact be seen as a welcome property. The second one still seems undesirable though, as one can see when evaluating (71-a) (an hybrid conjunction) and (71-b) (an hybrid disjunction) in a situation where the amplifiers are known to have never buzzed and to be borderline-loud now. Indeed, (71-a) does not seem plainly false, nor does (71-b) feel plainly true in this situation.

- (71)    a.    The amplifiers are loud and they have stopped buzzing.  
           b.    Either the amplifiers are loud or they have stopped buzzing.

Rather, it seems that in such situations, the truth-value judgments that we want to give for (71-a) and (71-b) are those that we give for simple presuppositional sentences in situations of presupposition failure. This chapter will propose alternative 4-valued systems where **Non-Classicality**, **Distinction** and **Unpolarized Vagueness** obtain but where **Polarized Presupposition** does not. In addition, this system will assign the unique truth-value associated with presupposition failures to both (71-a) and (71-b) in the described situations.

## 6.1 A Bi-Lattice

The system that I present in this section integrates two dimensions respectively dedicated to vagueness and to presupposition. In a way, this could be compared to the view of Dunn-Belnap bi-lattice as integrating a dimension of informativity and a dimension of truth in a single system. On the one hand, borderline cases (which yield value  $\mathcal{V}$ ) define an area *between* two extreme regions on a scale corresponding to the negative and the positive extensions of a vague predicate, and on the other hand presupposition failures (which yield value  $\mathcal{P}$ ) are perceived as an “*out of place*” phenomenon.  $\mathcal{V}$  and  $\mathcal{P}$  therefore enter in relation with the classical truth-values 1 and 0 along two *distinct* dimensions. This straightforwardly accounts for **Non-Classicality**

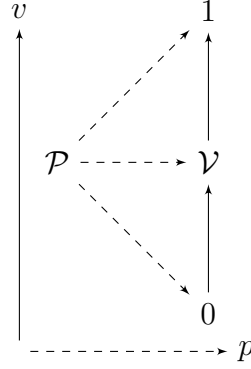


Figure 6.1: The 4-valued bilattice defined in Def. 6.1.1

and for **Distinction**. Formally, I will consider that each dimension defines a total order: on one dimension,  $\mathcal{V}$  lies between plain truth and plain falsity, and on another dimension,  $\mathcal{P}$  is left aside from 0,  $\mathcal{V}$  and 1 as “out of place”. Definition 6.1.1 formalizes this position and Figure 6.1 presents the associated bi-lattice.

**Definition 6.1.1** (4-Valued Bi-Lattice). *Let  $\{0, \mathcal{P}, \mathcal{V}, 1\}$  be our set of truth-values.  $\leq_p$  and  $\leq_v$  define two weak partial orders such that:*

- $\mathcal{P} \leq_p 0, \mathcal{V}, 1$  but  $0 \not\leq_p \mathcal{V}, 1$  and  $\mathcal{V} \not\leq_p 0, 1$  and  $1 \not\leq_p 0, \mathcal{V}$  (i.e. 0,  $\mathcal{V}$  and 1 are not ordered along  $\leq_p$ );
- $0 \leq_v \mathcal{V}$  and  $\mathcal{V} \leq_v 1$  but  $\mathcal{P} \not\leq_v 0, \mathcal{V}, 1$  and  $0, \mathcal{V}, 1 \not\leq_v \mathcal{P}$  (i.e. 0,  $\mathcal{P}$  is not ordered with 0,  $\mathcal{V}$  nor 1 along  $\leq_v$ )

The semantics I will give for the negation operator will have the effect of reversing the truth-value of a proposition along  $\leq_v$ . As a consequence,  $\mathcal{V}$  and  $\mathcal{P}$  will evenly be unaffected by this operator, allowing this system to account for **Unpolarized Vagueness** and **Unpolarized Presupposition**. To define conjunction, disjunction and implication, I will first introduce a total order based on  $\leq_p$  and  $\leq_v$ , along which each truth-value is distinct from the three others, and where  $\mathcal{P}$  is the lowest truth-value. As a consequence, I will give a semantics for these operators that associates hybrid conjunctions like (71-a) and hybrid disjunctions like (71-b) with the truth-value  $\mathcal{P}$ . Definition 6.1.2 defines the relation  $\leq_4$  such that  $\mathcal{P} \leq_4 0 \leq_4 \mathcal{V} \leq_4 1$ ; Definition 6.1.3 defines the semantics of the 4-valued system, and in particular of the operators.

**Definition 6.1.2** ( $\leq_4$ ). For any  $a, b \in \{0, \mathcal{V}, \mathcal{P}, 1\}$ ,  $a \leq_4 b$  if and only if  $a \leq_p b$  or  $a \leq_v b$ .

First, we can show that  $\leq_4$  is transitive:

- a. The relations  $\leq_p$  and  $\leq_v$  define partial orders and are thus transitive.
- b. For any  $a, b$  such that  $a \leq_4 b$ , either  $a \leq_p b$  and  $\leq_p$  is transitive or  $a \leq_v b$  and  $\leq_v$  is transitive too.
- c. Hence,  $\leq_4$  is transitive  $\square$

Then, we can show that  $\leq_4$  defines a weak *total* order where  $\mathcal{P} \leq_4 0 \leq_4 \mathcal{V} \leq_4 1$ :

- a.  $\mathcal{P} \leq_4 0, \mathcal{V}, 1$  because  $\mathcal{P} \leq_p 0, \mathcal{V}, 1$ .
- b.  $0 \leq_4 \mathcal{V}, 1$  because  $0 \leq_v \mathcal{V}, 1$ .
- c.  $\mathcal{V} \leq_4 1$  because  $\mathcal{V} \leq_v 1$ .
- d. By transitivity of  $\leq_4$ , we have  $\mathcal{P} \leq_4 0 \leq_4 \mathcal{V} \leq_4 1$ .
- e.  $0 \not\leq_4 \mathcal{P}$  because  $0 \not\leq_p \mathcal{P}$  and  $0 \not\leq_v \mathcal{P}$ .
- f.  $\mathcal{V} \not\leq_4 \mathcal{P}, 0$  because  $\mathcal{V} \not\leq_p \mathcal{P}$  and  $\mathcal{V} \not\leq_v \mathcal{P}$ , and  $\mathcal{V} \not\leq_p 0$  and  $\mathcal{V} \not\leq_v 0$ .
- g.  $1 \not\leq_4 \mathcal{P}, 0, \mathcal{V}$  because  $1 \not\leq_p \mathcal{P}$  and  $1 \not\leq_v \mathcal{P}$ , and  $1 \not\leq_p 0$  and  $1 \not\leq_v 0$ , and  $1 \not\leq_p \mathcal{V}$  and  $1 \not\leq_v \mathcal{V}$ .
- h. Hence,  $\leq_4$  defines a weak total order where  $\mathcal{P} \leq_4 0 \leq_4 \mathcal{V} \leq_4 1$   $\square$

**Definition 6.1.3** (Semantics). Let  $\{0, \mathcal{P}, \mathcal{V}, 1\}$  be our set of truth-values and let  $\mathcal{L}$  be our language containing vague predicates and presuppositional propositions. For any model  $\mathcal{M}$  whose interpretation function is  $\mathcal{I}$

- i. for any vague predicate  $P \in \mathcal{L}$  and any individual name  $a \in \mathcal{L}$ ,  $\mathcal{I}(Pa) = \mathcal{V}$  iff  $\mathcal{I}(a)$  is a borderline case for  $\mathcal{I}(P)$ ,  $\mathcal{I}(Pa) \in \{0, 1\}$  otherwise.
- ii. for any atomic presuppositional proposition  $\phi \in \mathcal{L}$ ,  $\mathcal{I}(\phi) = \mathcal{P}$  iff its presupposition is not fulfilled in  $\mathcal{M}$ ,  $\mathcal{I}(\phi) \in \{0, 1\}$  otherwise.
- iii. for any wff  $\phi$ ,
  - a.  $\mathcal{I}(\neg\phi) = 1$  iff  $\mathcal{I}(\phi) = 0$

- b.  $\mathcal{I}(\neg\phi) = 0$  iff  $\mathcal{I}(\phi) = 1$
- c.  $\mathcal{I}(\neg\phi) = \mathcal{V}$  iff  $\mathcal{I}(\phi) = \mathcal{V}$
- d.  $\mathcal{I}(\neg\phi) = \mathcal{P}$  iff  $\mathcal{I}(\phi) = \mathcal{P}$ .

iv. for two wff  $\phi$  and  $\psi$

- a.  $\mathcal{I}(\phi \wedge \psi) = \mathcal{I}(\phi)$  iff  $\mathcal{I}(\phi) \leq_4 \mathcal{I}(\psi)$ ,  $\mathcal{I}(\phi \wedge \psi) = \mathcal{I}(\psi)$  otherwise
- b.  $\mathcal{I}(\phi \vee \psi) = \mathcal{I}(\neg(\neg\phi \wedge \neg\psi))$
- c.  $\mathcal{I}(\phi \rightarrow \psi) = \mathcal{I}(\neg\phi \vee \psi)$ .

### 6.1.1 $\mathcal{P}$ as an Absorbing Truth-Value

Let us see what truth-values these definitions derive for negation, conjunction, disjunction and implication depending on the truth-values of the members:

$\phi$	$\psi$	$\neg\phi$	$\neg\psi$	$\phi \wedge \psi$	$\neg\phi \wedge \neg\psi$	$\phi \vee \psi$	$\neg\phi \vee \psi \equiv \phi \rightarrow \psi$
0	0	1	1	0	1	0	1
0	$\mathcal{V}$	1	$\mathcal{V}$	0	$\mathcal{V}$	$\mathcal{V}$	1
0	$\mathcal{P}$	1	$\mathcal{P}$	$\mathcal{P}$	$\mathcal{P}$	$\mathcal{P}$	$\mathcal{P}$
0	1	1	0	0	0	1	1
$\mathcal{V}$	0	$\mathcal{V}$	1	0	$\mathcal{V}$	$\mathcal{V}$	$\mathcal{V}$
$\mathcal{V}$	$\mathcal{V}$	$\mathcal{V}$	$\mathcal{V}$	$\mathcal{V}$	$\mathcal{V}$	$\mathcal{V}$	$\mathcal{V}$
$\mathcal{V}$	$\mathcal{P}$	$\mathcal{V}$	$\mathcal{P}$	$\mathcal{P}$	$\mathcal{P}$	$\mathcal{P}$	$\mathcal{P}$
$\mathcal{V}$	1	$\mathcal{V}$	0	$\mathcal{V}$	0	1	1
$\mathcal{P}$	0	$\mathcal{P}$	1	$\mathcal{P}$	$\mathcal{P}$	$\mathcal{P}$	$\mathcal{P}$
$\mathcal{P}$	$\mathcal{V}$	$\mathcal{P}$	$\mathcal{V}$	$\mathcal{P}$	$\mathcal{P}$	$\mathcal{P}$	$\mathcal{P}$
$\mathcal{P}$	$\mathcal{P}$	$\mathcal{P}$	$\mathcal{P}$	$\mathcal{P}$	$\mathcal{P}$	$\mathcal{P}$	$\mathcal{P}$
$\mathcal{P}$	1	$\mathcal{P}$	0	$\mathcal{P}$	$\mathcal{P}$	$\mathcal{P}$	$\mathcal{P}$
1	0	0	1	0	0	1	0
1	$\mathcal{V}$	0	$\mathcal{V}$	$\mathcal{V}$	0	1	$\mathcal{V}$
1	$\mathcal{P}$	0	$\mathcal{P}$	$\mathcal{P}$	$\mathcal{P}$	$\mathcal{P}$	$\mathcal{P}$
1	1	0	0	1	0	1	1

Table 6.1: Truth-tables obtained from Definition 6.1.3

The predominance of  $\mathcal{P}$  in these truth-tables shows why  $\mathcal{P}$  can be described as an *absorbing* truth-value in this system. The formal reasons for this lie in four points:

- Negation does not affect the truth-value  $\mathcal{P}$
- $\mathcal{P}$  is the lowest value under  $\leq_4$

- Conjunction returns the lowest truth-value under  $\leq_4$  among the truth-values of its arguments
- Disjunction and implication are defined exclusively in terms of negation and conjunction

This system meets our intuitions concerning hybrid conjunctions and hybrid disjunctions. Let us take  $\phi \wedge \psi$  and  $\phi \vee \psi$  as the respective translations of the hybrid conjunction (71-a) and the hybrid disjunction (71-b); and let us subsequently assume that in situations where the amplifiers have never buzzed and are currently borderline-loud,  $\phi$  gets the value  $\mathcal{V}$  and  $\psi$  gets the value  $\mathcal{P}$ . We see from the truth-tables above that in these situations, this system associates both the hybrid conjunction (71-a) and the hybrid disjunction (71-b) with the truth-value  $\mathcal{P}$ . As a result, this 4-valued system associates these hybrid sentences with judgments specific of a presupposition failure. In contrast, as I showed in Chapt. 4, a Dunn-Belnap 4-valued system would assign 0 to  $\phi \wedge \psi$  and 1 to  $\phi \vee \psi$  when  $\phi$  gets one of the non-bivalent truth-values and when  $\psi$  gets the other non-bivalent truth-value.

But because this system makes  $\mathcal{P}$  an absorbing truth-value, in particular it assigns  $\phi \rightarrow \psi$  the value  $\mathcal{P}$  when  $\phi$  gets 0 and when  $\psi$  gets  $\mathcal{P}$ . In regard of (72), already discussed in Chapt. 2 and which could be translated as  $\phi \rightarrow \psi$ , this might sound too strong, for we do not seem to have truth-value judgments typical of a presupposition failure when we think there is no wooden planet (which is a situation where we would give  $\phi$  the truth-value 0 and  $\psi$  the truth-value  $\mathcal{P}$ ).

(72) If there are wooden planets, then the wooden planets are burnable.

We saw in Chapt. 2 that (George 2008) proposed to derive and use the Peters truth-tables (also called the *Middle Kleene* truth-tables) instead of the Strong Kleene truth-tables to model the *incremental* aspect of presupposition.<sup>1</sup> The next section discusses how the repair strategy can derive the present 4-valued system and how we can further introduce minor changes to have an incremental approach of presupposition.

---

<sup>1</sup>Even though, for simplicity, I will consider incrementality as reflecting a *linear* parsing of sentences, recall that George remains neutral as to whether we should use linear or syntactical considerations in implementing the incremental aspect of presupposition.

## 6.2 The Repair Strategy

### 6.2.1 Weak Kleene and Strong Kleene

A very interesting property of the 4-valued system I just defined is how it is related to the Weak and Strong Kleene systems. If we remove from the previous 4-valued truth-tables all the lines where either  $\phi$  or  $\psi$  gets the value  $\mathcal{P}$ , we get trivalent truth-tables which correspond to the Strong Kleene truth-tables, where  $\mathcal{V}$  stands for the third truth-value.

$\phi$	$\psi$	$\neg\phi$	$\neg\psi$	$\phi \wedge \psi$	$\phi \vee \psi$	$\phi \rightarrow \psi$
0	0	1	1	0	0	1
0	$\mathcal{V}$	1	$\mathcal{V}$	0	$\mathcal{V}$	1
0	1	1	0	0	1	1
$\mathcal{V}$	0	$\mathcal{V}$	1	0	$\mathcal{V}$	$\mathcal{V}$
$\mathcal{V}$	$\mathcal{V}$	$\mathcal{V}$	$\mathcal{V}$	$\mathcal{V}$	$\mathcal{V}$	$\mathcal{V}$
$\mathcal{V}$	1	$\mathcal{V}$	0	$\mathcal{V}$	1	1
1	0	0	1	0	1	0
1	$\mathcal{V}$	0	$\mathcal{V}$	$\mathcal{V}$	1	$\mathcal{V}$
1	1	0	0	1	1	1

Table 6.2:  $\mathcal{P}$ -free lines from Table 6.1

But now, if we remove from the 4-valued truth-tables all the lines where either  $\phi$  or  $\psi$  gets *the value*  $\mathcal{V}$ , we get trivalent truth-tables which correspond to the Weak Kleene truth-tables, where  $\mathcal{P}$  corresponds to the third truth-value.<sup>2</sup>

$\phi$	$\psi$	$\neg\phi$	$\neg\psi$	$\phi \wedge \psi$	$\phi \vee \psi$	$\phi \rightarrow \psi$
0	0	1	1	0	0	1
0	$\mathcal{P}$	1	$\mathcal{P}$	$\mathcal{P}$	$\mathcal{P}$	$\mathcal{P}$
0	1	1	0	0	1	1
$\mathcal{P}$	0	$\mathcal{P}$	1	$\mathcal{P}$	$\mathcal{P}$	$\mathcal{P}$
$\mathcal{P}$	$\mathcal{P}$	$\mathcal{P}$	$\mathcal{P}$	$\mathcal{P}$	$\mathcal{P}$	$\mathcal{P}$
$\mathcal{P}$	1	$\mathcal{P}$	0	$\mathcal{P}$	$\mathcal{P}$	$\mathcal{P}$
1	0	0	1	0	1	0
1	$\mathcal{P}$	0	$\mathcal{P}$	$\mathcal{P}$	$\mathcal{P}$	$\mathcal{P}$
1	1	0	0	1	1	1

Table 6.3:  $\mathcal{V}$ -free lines from Table 6.1

The 4-valued system I defined can therefore be thought of as a particular way of merging the Weak and Strong Kleene truth-tables: the lines where  $\mathcal{V}$  and  $\mathcal{P}$  appear

---

<sup>2</sup>It is noticeable that this property distinguishes the present 4-valued bi-lattice from the Dunn-Belnap bi-lattice which results in the Strong Kleene truth-tables regardless of which non-bivalent truth-value we remove (see (Muskens 1999) on this point).

together (ie. the line where  $[I(\phi) = \mathcal{P} \text{ and } I(\psi) = \mathcal{V}]$  and the line where  $[I(\phi) = \mathcal{V} \text{ and } I(\psi) = \mathcal{P}]$ ) are absent from the two trivalent truth-tables and constitute the specific predictions of the 4-valued system. There are obviously many other formal possibilities for filling these lines, and this is what makes this system specific. But note that actually I have not simply merged the two truth-tables above and stipulated some completions for the missing lines. Rather, I considered a specific configuration of the truth-values themselves (that I claim to be conceptually motivated) and defined the set of operators on this basis.

## 6.2.2 Merging Weak and Strong Kleene

There is however a way to derive the exact same 4-valued truth-tables by defining a *repair function* on purely bivalent grounds. In (George 2008), repair functions take boolean functions as arguments and return 3-valued versions of the functions. The repair function that I will define also takes boolean functions as arguments but returns *4-valued* versions of the functions. The repair function actually consists in merging the repair function that derives the Weak Kleene truth-tables and the repair function that derives the Strong Kleene truth-tables. Hereafter, we introduce a repair function that derives the Weak Kleene truth-tables with  $\mathcal{P}$  as the third-value, and a repair function that derives the Strong Kleene truth-tables with  $\mathcal{V}$  as the third-value.

**Definition 6.2.1** (Weak Kleene Repair Function  $\mathcal{R}_w$ ). *For any function  $f : \overrightarrow{Bool} \rightarrow Bool$ ,  $\mathcal{R}_w(f)$  is the function such that, for any vector  $\vec{v}$  of truth-values in  $\{0, \mathcal{P}, 1\}$  and of the same length as the vectors in the domain of  $f$ :*

- $\mathcal{R}_w(f)(\vec{v}) = \mathcal{P}$  if and only if  $\vec{v}$  contains  $\mathcal{P}$ ,
- $\mathcal{R}_w(f)(\vec{v})$  is the unique value that  $f$  returns for  $\vec{v}$  otherwise.

**Definition 6.2.2** (Strong Kleene Repair Function  $\mathcal{R}_s$ ). *For any function  $f : \overrightarrow{Bool} \rightarrow Bool$ ,  $\mathcal{R}_s(f)$  is the function such that, for any vector  $\vec{v}$  of truth-values in  $\{0, \mathcal{V}, 1\}$  and of the same length as the vectors in the domain of  $f$ :*

- $\mathcal{R}_s(f)(\vec{v}) = \mathcal{V}$  if and only if there exist two bivalently repaired vectors  $\vec{v}_1'$  and  $\vec{v}_2'$ , i.e. two vectors whose elements are the same as those of  $\vec{v}$  except that each  $\mathcal{V}$  has been arbitrarily replaced either by 0 or by 1, and such that  $f(\vec{v}_1') \neq f(\vec{v}_2')$ ,
- $\mathcal{R}_s(f)(\vec{v})$  is the unique value that  $f$  returns for any such bivalently repaired vector otherwise.



Now we would like to have a general repair function returning functions defined on  $\{0, \mathcal{P}, \mathcal{V}, 1\}$ . But in order to return a repaired function,  $\mathcal{R}_w$  and  $\mathcal{R}_s$  require that respectively  $\mathcal{P}$ -free and  $\mathcal{V}$ -free vectors contain only 0 and 1. Indeed, if we extended the set of truth-values considered by  $\mathcal{R}_w$  to  $\{0, \mathcal{V}, \mathcal{P}, 1\}$  with no further modification,  $\mathcal{R}_w$  would crash with vectors containing  $\mathcal{V}$  but no  $\mathcal{P}$ , for it would try to return the unique value returned by the initial boolean function, but this latter function is actually undefined for  $\mathcal{V}$ . Similarly, if we extended the set of truth-values considered by  $\mathcal{R}_s$  to  $\{0, \mathcal{V}, \mathcal{P}, 1\}$  with no further modification,  $\mathcal{R}_s$  would crash with vectors containing  $\mathcal{P}$ , for it would try to compute the truth-values that the initial boolean function would return for  $\mathcal{V}$ -free vectors, but these  $\mathcal{V}$ -free vectors would still contain  $\mathcal{P}$  and the initial boolean function is undefined for  $\mathcal{P}$ . Fortunately, there is a straightforward way to solve this conflict, and it exploits two properties of  $\mathcal{R}_w$ : i) it does not call the initial boolean function to determine the conditions in which the repaired function should return  $\mathcal{P}$ ; ii) the repaired function does not return  $\mathcal{P}$  if and only if the passed vector does not contain  $\mathcal{P}$ . Therefore, we can define the general repair function so that it behaves as  $\mathcal{R}_w$  with vectors containing  $\mathcal{P}$ , and as  $\mathcal{R}_s$  with vectors where  $\mathcal{P}$  does not appear. This is what Definition 6.2.3 does: the first statement corresponds to the first statement in the definition of  $\mathcal{R}_w$ , the second and the third statements correspond to the two statements in the definition of  $\mathcal{R}_s$ .

**Definition 6.2.3** (General Repair Function  $\mathcal{R}_g$ ). *For any function  $f : \overrightarrow{Bool} \rightarrow Bool$ ,  $\mathcal{R}_g(f)$  is the function such that, for any vector  $\vec{v}$  of truth-values in  $\{0, \mathcal{V}, \mathcal{P}, 1\}$  and of the same length as the vectors in the domain of  $f$ :*

- $\mathcal{R}_g(f)(\vec{v}) = \mathcal{P}$  if and only if  $\vec{v}$  contains  $\mathcal{P}$ ,
- otherwise,  $\mathcal{R}_g(f)(\vec{v}) = \mathcal{V}$  if there exist two bivalently repaired vectors  $\vec{v}_1'$  and  $\vec{v}_2'$ , i.e. two vectors whose elements are the same as those of  $\vec{v}$  except that each  $\mathcal{V}$  has been arbitrarily replaced either by 0 or by 1, and such that  $f(\vec{v}_1') \neq f(\vec{v}_2')$ ,
- $\mathcal{R}_g(f)(\vec{v})$  is the unique value that  $f$  returns for any such bivalently repaired vector otherwise.

Proving that this general repair function derives the 4-valued truth-tables above is quite straightforward. First, note that the trivalent truth-tables above are derived by the general repair function. Indeed, when the vectors do not contain  $\mathcal{V}$ ,  $\mathcal{R}_g$  has the same effects as  $\mathcal{R}_w$ , and when the vectors do not contain  $\mathcal{P}$ ,  $\mathcal{R}_g$  has the same effects as  $\mathcal{R}_s$ . Given that these functions correctly derive their respective trivalent

truth-tables, we only need to check the vectors corresponding to the two lines where  $\phi$  or  $\psi$  receives the value  $\mathcal{V}$  while the other receives the value  $\mathcal{P}$ . As we saw above,  $\mathcal{R}_w$  makes  $\mathcal{P}$  an absorbing truth-value, and  $\mathcal{R}_g$  shares this property. As a consequence, when  $\mathcal{P}$  appears in a conjunction, in a disjunction or in an implication, the complex proposition gets the value  $\mathcal{P}$ . Therefore, we do get the two specific lines that we observe in the 4-valued truth-table.

### 6.2.3 Merging Strong and Middle Kleene

I will now revise the general repair function so as to implement the incremental aspect of presupposition. First, Definition 6.2.4 is a reformulation of Def. 2.2.2 (*George's Peters Repair Function*) from Chapt. 2, with  $\mathcal{P}$  being the non-bivalent truth-value.

**Definition 6.2.4** (Middle Kleene Repair Function  $\mathcal{R}_m$ ). *For any function  $f : \overrightarrow{Bool} \rightarrow Bool$ ,  $\mathcal{R}_m(f)$  is the function such that, for any vector  $\vec{v}$  of truth-values in  $\{0, \mathcal{P}, 1\}$  and of the same length as the vectors in the domain of  $f$ :*

- $\mathcal{R}_m(f)(\vec{v}) = \mathcal{P}$  if and only if there exist two incrementally bivalently repaired vectors  $\vec{v}'_1$  and  $\vec{v}'_2$ , i.e. two vectors whose elements are the same as those of  $\vec{v}$  until the first  $\mathcal{P}$ , starting from which each element has been arbitrarily replaced by either 0 or 1, and such that  $f(\vec{v}'_1) \neq f(\vec{v}'_2)$ ,
- $\mathcal{R}_m(f)(\vec{v})$  is the unique value that  $f$  returns for any such incrementally bivalently repaired vector otherwise.

I report Peters/Middle Kleene truth-tables below, with  $\mathcal{P}$  as the third truth-value.

$\phi$	$\psi$	$\neg\phi$	$\neg\psi$	$\phi \wedge \psi$	$\phi \vee \psi$	$\phi \rightarrow \psi$
0	0	1	1	0	0	1
0	$\mathcal{P}$	1	$\mathcal{P}$	0	$\mathcal{P}$	1
0	1	1	0	0	1	1
$\mathcal{P}$	0	$\mathcal{P}$	1	$\mathcal{P}$	$\mathcal{P}$	$\mathcal{P}$
$\mathcal{P}$	$\mathcal{P}$	$\mathcal{P}$	$\mathcal{P}$	$\mathcal{P}$	$\mathcal{P}$	$\mathcal{P}$
$\mathcal{P}$	1	$\mathcal{P}$	0	$\mathcal{P}$	$\mathcal{P}$	$\mathcal{P}$
1	0	0	1	0	1	0
1	$\mathcal{P}$	0	$\mathcal{P}$	$\mathcal{P}$	1	$\mathcal{P}$
1	1	0	0	1	1	1

Table 6.4: Peters/Middle Kleene truth-tables

As before, we want to get a general repair function which would behave as  $\mathcal{R}_s$  with  $\mathcal{V}$  and this time as  $\mathcal{R}_m$  with  $\mathcal{P}$ . Unfortunately,  $\mathcal{R}_m$  lacks the two properties of

$\mathcal{R}_w$  which we exploited to define a merged repair function. Namely,  $\mathcal{R}_m$  *does* call the initial boolean function to determine the conditions in which the repaired function should return  $\mathcal{P}$ ; and in addition the vectors for which the function repaired with  $\mathcal{R}_m$  returns another value than  $\mathcal{P}$  *can* contain  $\mathcal{P}$ . As a result, and contrary to the spirit of what we did when we merged  $\mathcal{R}_s$  and  $\mathcal{R}_w$ , there is no straightforward way to obtain a general repair function which would act as  $\mathcal{R}_s$  whenever  $\mathcal{R}_m$  would fail to return  $\mathcal{P}$ . To see this, note that the first statement in Def. 6.2.4 requires that each value before the first  $\mathcal{P}$  in the passed vector be either 0 or 1. Indeed, if we extended the truth-values considered by  $\mathcal{R}_m$  to  $\{0, \mathcal{V}, \mathcal{P}, 1\}$  with no further modification,  $\mathcal{R}_m$  would crash in particular with the vector  $\langle \mathcal{V}, \mathcal{P} \rangle$ , for it would try to determine whether the initial boolean function returns the same truth-value for the vectors  $\langle \mathcal{V}, 0 \rangle$  and  $\langle \mathcal{V}, 1 \rangle$  but this initial boolean function is actually undefined for these vectors.

One solution to this problem is to define a general repair function defined like  $\mathcal{R}_m$ , with the exception that instead of referring to the initial boolean truth-value it would refer to the function repaired with  $\mathcal{R}_s$ , so that it can handle vectors containing  $\mathcal{V}$ . But there is an alternative solution, which actually consists in doing the opposite: the general repair function would be defined like  $\mathcal{R}_s$ , with the exception that instead of referring to the initial boolean function it would refer to the function repaired with  $\mathcal{R}_m$ , so that it can handle vectors containing  $\mathcal{P}$ . And it turns out that these two implementations produce repaired function which only differ with respect to vectors where  $\mathcal{V}$  appears before  $\mathcal{P}$ . Definition 6.2.5 and Definition 6.2.6 present these two alternatives.

**Definition 6.2.5** (General Repair Function  $\mathcal{R}_{ms}$ ). *For any function  $f : \overrightarrow{Bool} \rightarrow Bool$ ,  $\mathcal{R}_{ms}(f)$  is the function such that, for any vector  $\vec{v}$  of truth-values in  $\{0, \mathcal{V}, \mathcal{P}, 1\}$  and of the same length as the vectors in the domain of  $f$ :*

- $\mathcal{R}_{ms}(f)(\vec{v}) = \mathcal{P}$  *if and only if there exist two incrementally  $\mathcal{P}$ -free repaired vectors  $\vec{v}'_1$  and  $\vec{v}'_2$ , i.e. two vectors whose elements are the same as those of  $\vec{v}$  until the first  $\mathcal{P}$ , starting from which each element has been arbitrarily replaced by either 0 or 1, and such that  $\mathcal{R}_s(f)(\vec{v}'_1) \neq \mathcal{R}_s(f)(\vec{v}'_2)$ ,*
- $\mathcal{R}_{ms}(f)(\vec{v})$  *is the unique value that  $\mathcal{R}_s(f)$  returns for any such incrementally  $\mathcal{P}$ -free repaired vector otherwise.*

**Definition 6.2.6** (General Repair Function  $\mathcal{R}_{sm}$ ). *For any function  $f : \overrightarrow{Bool} \rightarrow Bool$ ,  $\mathcal{R}_{sm}(f)$  is the function such that, for any vector  $\vec{v}$  of truth-values in  $\{0, \mathcal{V}, \mathcal{P}, 1\}$  and of the same length as the vectors in the domain of  $f$ :*

- $\mathcal{R}_{sm}(f)(\vec{v}) = \mathcal{V}$  if and only if there exist two  $\mathcal{V}$ -free repaired vectors  $\vec{v}'_1$  and  $\vec{v}'_2$ , i.e. two vectors whose elements are the same as those of  $\vec{v}$  except that each  $\mathcal{V}$  has been arbitrarily replaced either by 0 or by 1, and such that  $\mathcal{R}_m(f)(\vec{v}'_1) \neq \mathcal{R}_m(f)(\vec{v}'_2)$ ,
- $\mathcal{R}_{sm}(f)(\vec{v})$  is the unique value that  $\mathcal{R}_m(f)$  returns for any such  $\mathcal{V}$ -free repaired vector otherwise.

Now let us see that the functions repaired with  $\mathcal{R}_{ms}$  and the functions repaired with  $\mathcal{R}_{sm}$  differ only for vectors which contain a  $\mathcal{V}$  before a  $\mathcal{P}$ .

- For any vector of *bivalent truth-values*  $\vec{v}$  and any boolean function  $f$ , it is straightforward that  $\mathcal{R}_{ms}(f)(\vec{v}) = \mathcal{R}_{sm}(f)(\vec{v})$ 
  - Because there is no  $\mathcal{P}$  in  $\vec{v}$ ,  $\mathcal{R}_{ms}(f)(\vec{v})$  returns the value that  $\mathcal{R}_s(f)(\vec{v})$  returns, which is actually  $f(\vec{v})$  for there is no  $\mathcal{V}$  in  $\vec{v}$
  - Because there is no  $\mathcal{V}$  in  $\vec{v}$ ,  $\mathcal{R}_{sm}(f)(\vec{v})$  returns the value that  $\mathcal{R}_m(f)(\vec{v})$  returns, which is actually  $f(\vec{v})$  for there is no  $\mathcal{P}$  in  $\vec{v}$
  - Hence,  $\mathcal{R}_{ms}(f)(\vec{v}) = \mathcal{R}_{sm}(f)(\vec{v}) = f(\vec{v})$ .
- For any vector  $\vec{v}$  where  $\mathcal{P}$  is the first non-bivalent truth-value to appear and any boolean function  $f$ ,  $\mathcal{R}_{sm}(f)(\vec{v}) = \mathcal{R}_{ms}(f)(\vec{v})$ 
  - Because there is no  $\mathcal{V}$  before the first  $\mathcal{P}$  in  $\vec{v}$  and because any incrementally  $\mathcal{P}$ -free vector  $\vec{v}'$  repaired from  $\vec{v}$  replaces each element starting from the first  $\mathcal{P}$  by either 0 or 1,  $\vec{v}'$  necessarily contains only bivalent truth-values. Therefore, for any such  $\vec{v}'$ ,  $\mathcal{R}_s(f)(\vec{v}') = f(\vec{v}')$ . We thus have  $\mathcal{R}_{ms}(f)(\vec{v}) = \mathcal{R}_m(f)(\vec{v})$ .
  - Because any  $\mathcal{V}$ -free vector  $\vec{v}'$  repaired from  $\vec{v}$  has the same series of bivalent truth-values before the first  $\mathcal{P}$ ,  $\mathcal{R}_m(f)$  returns the same value for any such  $\vec{v}'$  (which is *not* a bivalent vector). Therefore we have  $\mathcal{R}_{sm}(f)(\vec{v}) = \mathcal{R}_m(f)$ .
  - Hence,  $\mathcal{R}_{ms}(f)(\vec{v}) = \mathcal{R}_{sm}(f)(\vec{v}) = \mathcal{R}_m(f)(\vec{v})$ .
- But for some vectors where  $\mathcal{V}$  appears before the first  $\mathcal{P}$  and some boolean functions  $f$ ,  $\mathcal{R}_{sm}(f)(\vec{v}) \neq \mathcal{R}_{ms}(f)(\vec{v})$ .  
Take in particular the boolean function  $f_\wedge$  such that for any boolean vector  $\vec{v}$ ,  $f_\wedge(\vec{v}) = 1$  if and only if every element in  $\vec{v}$  is 1, 0 otherwise, and take in particular the vector  $\langle \mathcal{V}, \mathcal{P} \rangle$ .

- $\langle \mathcal{V}, \mathcal{P} \rangle$  has two incrementally  $\mathcal{P}$ -free repaired vectors, namely  $\langle \mathcal{V}, 0 \rangle$  and  $\langle \mathcal{V}, 1 \rangle$ .  $\mathcal{R}_s(f_\wedge)(\langle \mathcal{V}, 0 \rangle) = 0$  because  $\langle \mathcal{V}, 0 \rangle$  has two  $\mathcal{V}$ -free repaired vectors, namely  $\langle 0, 0 \rangle$  and  $\langle 0, 1 \rangle$ , and  $f_\wedge(\langle 0, 0 \rangle) = f_\wedge(\langle 0, 1 \rangle) = 0$ .  $\mathcal{R}_s(f_\wedge)(\langle \mathcal{V}, 1 \rangle) = \mathcal{V}$  because  $\langle \mathcal{V}, 1 \rangle$  has two  $\mathcal{V}$ -free repaired vectors, namely  $\langle 0, 1 \rangle$  and  $\langle 1, 1 \rangle$ , and  $f_\wedge(\langle 0, 1 \rangle) = 0$  but  $f_\wedge(\langle 1, 1 \rangle) = 1$ . Therefore, there are two incrementally  $\mathcal{P}$ -free vectors  $\vec{v}_1'$  and  $\vec{v}_2'$  repaired from  $\langle \mathcal{V}, \mathcal{P} \rangle$  such that  $\mathcal{R}_s(f_\wedge)(\vec{v}_1') \neq \mathcal{R}_s(f_\wedge)(\vec{v}_2')$ . Hence,  $\mathcal{R}_{ms}(f_\wedge)(\langle \mathcal{V}, \mathcal{P} \rangle) = \mathcal{P}$ .
- $\langle \mathcal{V}, \mathcal{P} \rangle$  has two  $\mathcal{V}$ -free repaired vectors, namely  $\langle 0, \mathcal{P} \rangle$  and  $\langle 1, \mathcal{P} \rangle$ .  $\mathcal{R}_m(f_\wedge)(\langle 0, \mathcal{P} \rangle) = 0$  because  $\langle 0, \mathcal{P} \rangle$  has two incrementally  $\mathcal{P}$ -free repaired vectors, namely  $\langle 0, 0 \rangle$  and  $\langle 0, 1 \rangle$ , and  $f_\wedge(\langle 0, 0 \rangle) = f_\wedge(\langle 0, 1 \rangle) = 0$ .  $\mathcal{R}_m(f_\wedge)(\langle 1, \mathcal{P} \rangle) = \mathcal{P}$  because  $\langle 1, \mathcal{P} \rangle$  has two incrementally  $\mathcal{P}$ -free repaired vectors, namely  $\langle 1, 0 \rangle$  and  $\langle 1, 1 \rangle$ , and  $f_\wedge(\langle 1, 0 \rangle) = 0$  but  $f_\wedge(\langle 1, 1 \rangle) = 1$ . Therefore, there are two  $\mathcal{V}$ -free vectors  $\vec{v}_1'$  and  $\vec{v}_2'$  repaired from  $\langle \mathcal{V}, \mathcal{P} \rangle$  such that  $\mathcal{R}_m(f_\wedge)(\vec{v}_1') \neq \mathcal{R}_m(f_\wedge)(\vec{v}_2')$ . Hence,  $\mathcal{R}_{sm}(f_\wedge)(\langle \mathcal{V}, \mathcal{P} \rangle) = \mathcal{V}$ .

The 4-valued truth-tables in Table 6.5 have been derived using  $\mathcal{R}_{ms}$  and  $\mathcal{R}_{sm}$ . Note that the only line which exhibits different truth-values depending on which repair function was used is the line where  $\phi$  gets  $\mathcal{V}$  and  $\psi$  gets  $\mathcal{P}$ : the truth-values that appear on the left of an oblique bar result from a function repaired with  $\mathcal{R}_{ms}$  and the truth-values that appear on the right of an oblique bar result from a function repaired with  $\mathcal{R}_{sm}$ .

I motivated a *Middle* Kleene approach of presupposition on the observation that sentences like (72), repeated in (73), do not trigger truth-value judgments associated with presupposition failures.

(73) If there are wooden planets, then the wooden planets are burnable.

In this respect, note that  $\mathcal{R}_{ms}$  as well as  $\mathcal{R}_{sm}$  allow us to account for this observation. To see this, just note that when we focus on the  $\mathcal{V}$ -free lines in Table 6.5, we fall back on Peters truth-tables. I already showed in Chapt. 2 how such an approach deal with the projection of presuppositions in a 3-valued framework.

But what is new in Table 6.5 are the predictions that the systems based on  $\mathcal{R}_{ms}$  and  $\mathcal{R}_{sm}$  make for hybrid sentences like the hybrid conjunction (71-a) repeated in (74-a) and the hybrid disjunction (71-b) repeated in (74-b). Interestingly, these are precisely the types of sentences about which the two approaches disagree.

$\phi$	$\psi$	$\neg\phi$	$\neg\psi$	$\phi \wedge \psi$	$\neg\phi \wedge \neg\psi$	$\phi \vee \psi$	$\neg\phi \vee \psi \equiv \phi \rightarrow \psi$
0	0	1	1	0	1	0	1
0	$\mathcal{V}$	1	$\mathcal{V}$	0	$\mathcal{V}$	$\mathcal{V}$	1
0	$\mathcal{P}$	1	$\mathcal{P}$	0	$\mathcal{P}$	$\mathcal{P}$	1
0	1	1	0	0	0	1	1
$\mathcal{V}$	0	$\mathcal{V}$	1	0	$\mathcal{V}$	$\mathcal{V}$	$\mathcal{V}$
$\mathcal{V}$	$\mathcal{V}$	$\mathcal{V}$	$\mathcal{V}$	$\mathcal{V}$	$\mathcal{V}$	$\mathcal{V}$	$\mathcal{V}$
$\mathcal{V}$	$\mathcal{P}$	$\mathcal{V}$	$\mathcal{P}$	$\mathcal{P}/\mathcal{V}$	$\mathcal{P}/\mathcal{V}$	$\mathcal{P}/\mathcal{V}$	$\mathcal{P}/\mathcal{V}$
$\mathcal{V}$	1	$\mathcal{V}$	0	$\mathcal{V}$	0	1	1
$\mathcal{P}$	0	$\mathcal{P}$	1	$\mathcal{P}$	$\mathcal{P}$	$\mathcal{P}$	$\mathcal{P}$
$\mathcal{P}$	$\mathcal{V}$	$\mathcal{P}$	$\mathcal{V}$	$\mathcal{P}$	$\mathcal{P}$	$\mathcal{P}$	$\mathcal{P}$
$\mathcal{P}$	$\mathcal{P}$	$\mathcal{P}$	$\mathcal{P}$	$\mathcal{P}$	$\mathcal{P}$	$\mathcal{P}$	$\mathcal{P}$
$\mathcal{P}$	1	$\mathcal{P}$	0	$\mathcal{P}$	$\mathcal{P}$	$\mathcal{P}$	$\mathcal{P}$
1	0	0	1	0	0	1	0
1	$\mathcal{V}$	0	$\mathcal{V}$	$\mathcal{V}$	0	1	$\mathcal{V}$
1	$\mathcal{P}$	0	$\mathcal{P}$	$\mathcal{P}$	0	1	$\mathcal{P}$
1	1	0	0	1	0	1	1

Table 6.5: 4-valued truth-tables derived with  $\mathcal{R}_{ms}$  and  $\mathcal{R}_{sm}$

- (74) a. The amplifiers are loud and they have stopped buzzing.  
b. Either the amplifiers are loud or they have stopped buzzing.

Let us translate (74-a) as  $\phi \wedge \psi$  and (74-b) as  $\phi \vee \psi$ . In situations where the amplifiers are known to be borderline-loud and to have never buzzed, we will assign  $\mathcal{V}$  to  $\phi$  and  $\mathcal{P}$  to  $\psi$ . This configuration precisely corresponds to the line in Table 6.5 where  $\mathcal{R}_{ms}$  and  $\mathcal{R}_{sm}$  diverge.  $\mathcal{R}_{ms}$  derives  $\mathcal{P}$  for both  $\phi \wedge \psi$  and  $\phi \vee \psi$  whereas  $\mathcal{R}_{sm}$  derives  $\mathcal{V}$  for both  $\phi \wedge \psi$  and  $\phi \vee \psi$ . This means that adopting the repair function  $\mathcal{R}_{ms}$  would lead us to predict the observation of truth-value judgments typical of a presupposition failure for both (74-a) and (74-b) in the described situations, whereas adopting the repair function  $\mathcal{R}_{sm}$  would lead us to predict the observation of truth-value judgments typical of the description of a borderline case for both (74-a) and (74-b) in the described situations. In the end,  $\mathcal{R}_{ms}$  is at advantage, for our intuitions tell us that in the described situations, we would give (74-a) and (74-b) truth-value judgments that we would give for a simple sentence whose presupposition is unfulfilled.

Let us say that in the present framework, a sentence  $A$  translated as  $\phi$  presupposes  $B$  if and only if  $\phi$  gets  $\mathcal{P}$  whenever  $\text{not} - B$ . This way, we can look at Table 6.5 to determine what (74-a) and (74-b) presuppose: they respectively presuppose that  $\phi \wedge \psi$  and  $\phi \vee \psi$  do not get  $\mathcal{P}$ . If we were to use  $\mathcal{R}_{sm}$  as our repair function, we would thus say that (74-a) presupposes that either the amplifiers are not clearly loud or they buzzed

before; and we would say that (74-b) presupposes that either the amplifiers are at least borderline-loud or they buzzed before. But we argued that we should prefer to use  $\mathcal{R}_{ms}$ . Therefore, we say that (74-a) presupposes that either the amplifiers are clearly not loud or they buzzed before; and we say that (74-b) presupposes that either the amplifiers are clearly loud or they buzzed before. As we see,  $\mathcal{R}_{ms}$  and  $\mathcal{R}_{sm}$  not only yield different predictions about what truth-value judgments are to be observed for hybrid sentences, they can also make different predictions about the projection of presuppositions in hybrid sentences.

### 6.3 Conclusions and Comparison with ST5

The main weakness of ST5 is that it makes presuppositional propositions sensitive to negation in a way that it naturally predicts “True” judgments for negative presuppositional sentences. Observing the absence of such truth-value judgments for this kind of sentences in two experiments, I proposed to model vagueness and presupposition with a 4-valued bi-lattice that views these phenomena as two distinct non-bivalent aspects of one system. I claim that this 4-valued bi-lattice allows for a natural definition of negation that makes predications of borderline cases and propositions with unfulfilled presuppositions unaffected when appearing as arguments of the negation operator.

One appealing aspect of ST5 is its total order, which allows us to define two notions of satisfaction that can subsequently be bridged to truth-value judgments. With the 4-valued bi-lattice in Fig. 6.1, we find another motivation for predicting non-bivalent, specific truth-value judgments for vagueness and presupposition. On the one hand,  $\mathcal{V}$  is somewhere between plain truth and plain falsity of  $\leq_v$ , as it is in the original system ST. It is therefore natural to consider that vague predications of borderline cases are neither completely true nor completely false, but also partially true and partially false. On the other hand,  $\mathcal{P}$  is deemed apart from the spectrum of truth, and in particular from 0 and 1. It is therefore natural to consider that sentences with unfulfilled presuppositions are neither true nor false, and the configuration of the 4-valued bi-lattice would offer no justification for (unobserved) “Both true and false” judgments for these sentences.

It should be noted however that the absence of “True” judgments for the negative presuppositional descriptions in the previous experiments does not establish the *impossibility* of such judgments for negative sentences with unfulfilled presuppositions. In fact, as noted in Chapter 3, (Abrusán & Szendrői 2013) observed that participants

readily give a “True” judgment for negative presuppositional sentences of specific forms even when they believe in the falsity of the presupposition. In addition, I partly motivated ST5 on my intuitions that this judgment was available in these situations. In their discussions of negation, (Horn 1989) and (Beaver 2001) note that the acceptance of negated sentences with unfulfilled presupposition could be due to what they call a *metalinguistic* use of negation. To this extent, natural language could have a usual negation operator which would have the effect of reversing the truth of a proposition, and a metalinguistic negation operator which would have the effect of reversing the assertability of a proposition. The 4-valued bi-lattice in Fig. 6.1 calls for the definition of a metalinguistic negation: it would map propositions of value  $\mathcal{P}$  (i.e. “out of place” propositions) to the truth-value 1 and propositions of value 0,  $\mathcal{V}$  or 1 (i.e. “in place” propositions) to the truth-value 0. Importantly, that this operator treats  $\mathcal{V}$  on a par with 0 and 1 means that we could not use it to deny the felicitousness of a vague description of a borderline case. In contrast in ST5, and under the assumption that sentences are felicitous when they have a bivalent truth-value, that  $\mathcal{P}^0$ ,  $\mathcal{V}$  and  $\mathcal{P}^1$  are all between 0 and 1 makes it natural to consider that a proposition of one of these truth-values is not totally felicitous.

Rather than directly defining a metalinguistic negation, one could entertain the possibility of defining covert operators sensitive to the non-bivalent truth-values and which could take scope under negation, like the *meta-assertion operator* also discussed in (Beaver 2001). This operator would only leave propositions of value 1 unchanged and would map any proposition of a different value to 0. As a consequence, it would be possible to deny the felicitousness of propositions with unfulfilled propositions as well as the felicitousness of vague predications of borderline cases by embedding this operator under negation. In addition, the “covert operators” strategy allows us to define a covert *somewhat* operator that would account for the tolerant reading of vague predicates. This *somewhat* operator would exploit  $\leq_4$  and map propositions of a value on the top half to 1 but propositions of a value on the bottom half to 0. As a consequence, one would reject the vague description of a borderline case to the extent that the *meta-assertion* operator makes it false but would accept it to the extent that the *somewhat* operator makes it true. In contrast, both the *meta-assertion* and the *somewhat* operators would map propositions with unfulfilled presuppositions to 0, and one would have no choice but to reject these propositions as false when giving a bivalent truth-value judgment for them.

The derivation of  $\leq_4$  also provided us with the grounds to define the connectives for the 4-valued bi-lattice. Because  $\leq_4$  defines a total order, this system benefits



from the same feature as ST5 of directly dealing with hybrid propositions. But contrary to ST5,  $\mathcal{P}$  is here initially an absorbing truth-value, whereas ST5 treats  $\mathcal{P}^0$  and  $\mathcal{P}^1$  in a more Strong-Kleenean way. It is still an unanswered empirical question to determine which position is to be preferred, but some evidence from the study of presupposition projection argue for an intermediate position, where  $\mathcal{P}$  should be approached in a Middle-Kleenean way. One way to pursue this objective in either system is to implement an incremental algorithm *à la George* ((George 2008)), as what is suggested in (Fox 2008) for instance. This is precisely what I investigated for the 4-valued view after having observed the correspondence between the truth-tables derived from the bi-lattice system and the Strong and Weak Kleene truth-tables.

Finally, remember that the reason why ST5 associates  $\mathcal{P}^1$  with “True” judgments is because it also associates the lower value  $\mathcal{V}$  with “True” judgments. This latter association is motivated under the assumption that we observe “True” and “Both true and false” judgments for vague descriptions of borderline cases. I considered the existence of a *somewhat* operator in a 4-value system for the same reason. But the experiments reported in Chapter 5 provide no concrete evidence for these judgments. In addition, I noticed in Chapter 3 that the acceptance rate for glutty descriptions (“X is ADJ and not ADJ”) of borderline cases that (Alxatib & Pelletier 2011) found is lower than the one they found for the acceptance rate for gappy descriptions (“X is neither ADJ nor not ADJ”). They do not say whether this difference was significant, but on this basis one might question the accessibility of glutty judgments for vagueness and therefore the discarding of ST5. The next chapter addresses this issue by experimentally investigating the role of antonymy in speakers’ acceptance of gappy and glutty descriptions for borderline cases.

# Chapter 7

## Antonyms<sup>1</sup>

As the experiments of Chapter 5 revealed, empirical investigation can lead to question assumptions that are essential for the building of a theoretical system. My initial intuitions about the truth-value judgments that we give in situations of presupposition failure were not supported by the results that I obtained. As a consequence, the 5-valued system that I developed on the basis of these intuitions proved to be unable to account for the truth-value judgments that participants gave in situations of presupposition failure.

Even though the truth-value judgments that participants gave for borderline cases were not incompatible with the predictions of ST5, and more specifically with the prediction that in these situations speakers could give a “Both true and false” judgment (henceforth *glutty* judgment), the preceding experiments brought no evidence for their availability. Thus far, I have built a defense for my intuitions on this question on the basis of experimental results from the literature suggesting that speakers can describe borderline cases for a vague adjective *A* as *both A and not A*. More precisely, these results come from (Egré et al. 2013), (Ripley 2011), (Alxatib & Pelletier 2011) and (Serchuk et al. 2011).

### 7.1 Borderline Contradictions

In (Egré et al. 2013)’s experiment, there was a condition where participants were asked to agree or disagree with descriptions of the form *the square is COLOR and not COLOR* against a series of patches whose color varied along a scale from the color

---

<sup>1</sup>As noted in Footnote 1 from Chapter 5, in this chapter I present descriptive analyses, but most aspects of the discussion rely on robust observations, so whether more elaborate analyses bring new significant differences or dismiss an alleged effect should not have any dramatic impact on the content of my discussion. In any event, the data that I collected are available for statistical treatment – readers wishing to examine the data should contact me at [jeremy.e.zehr@gmail.com](mailto:jeremy.e.zehr@gmail.com).

*COLOR* to another color. Whether the patches were presented in random, ascending or descending order proved to have no crucial effect. They found that participants agreed with these glutty descriptions more than half the time for patches in the central region, and even significantly more often than simple descriptions of the form *the square is COLOR* or *the square is not COLOR* judged in isolation for the same patches. These are very compelling results. For the present purposes however, it should be noted that descriptions of the form *the square is neither COLOR nor not COLOR* (henceforth *gappy* descriptions) were not presented to participants.

(Ripley 2011) found results similar to the ones just discussed and he also investigated gappy descriptions. All the participants were presented with the same picture (projected onto a screen) representing seven pairs of a square and a circle aligned horizontally, with the distance between them varying across the pairs. The pair presenting the highest distance between the square and the circle was at the top of the picture, and the distance decreased with each pair so as to have side by side figures at the bottom of the picture. Depending on the group they belonged, they they were asked to rate a gappy or a glutty description constructed with *near* on a seven point scale for each pair (there were four groups, with two different gappy descriptions and two different glutty descriptions). In each group, there was a pair for which the mean rate that participants gave was important. Individually, over half of the participants rated the description 6/7 or 7/7 for one of the seven pairs. All of the four highest mean scores (5.2/7, 5.3/7, 5.7/7, 5.1/1=7) were significantly above 4, suggesting that participants did not answer by chance for borderline-near figures but actually felt the description to be admissible. In addition, the statistical analyses revealed no significant difference between the groups' answers, which means that both glutty and gappy descriptions were as readily accepted.

(Alxatib & Pelletier 2011)'s results are less clear. Every participants were presented with a picture representing five men of various heights and for each of them they had to indicate their judgments on a hard-copy sheet regarding 4 descriptions: *the man is tall*, *the man is not tall*, *the man is tall and not tall* and *the man is neither tall nor not tall*. They could check *true*, *false* or *can't tell*. It turned out that 44.7% of participants accepted the glutty description as true for the man of average height (contra 14.5% for the smallest man and 5.3% for the tallest man) and 53.9% of participants accepted the gappy description as true for the same man (contra 27.6% for the smallest man and 6.6% for the tallest man). The authors do not discuss this contrast and do not provide any statistical analysis for it, but this seems to suggest

an overall contrast between gappy and gluttony descriptions, with gappy descriptions being more readily accepted than gluttony descriptions.

(Serchuk et al. 2011) found an even bigger contrast between these two types of descriptions. They conducted four experiments on vague adjectives: two of them involved gluttony descriptions and one of these two also involved gappy descriptions. They explicitly asked participants to imagine a woman, Susan, who was described as somewhere between clear instances of the adjective and clear counter-instances of the adjective. The adjective was either *rich* or *heavy*, and the description to be evaluated was either of the form *Susan is ADJECTIVE* or *Susan is definitely ADJECTIVE*. Then participants had to check one truth-value judgment among *true*, *not true*, *but also not false*, *partially true and partially false*, *false*, *both true and false*, *true or false*, *but I don't know which*. The presence of *definitely* had the effect that a vast majority of participants checked *false* for either adjective. When *definitely* was absent, the answers for the description containing *heavy* distributed over all the judgments, with a tendency toward *not true*, *but also not false* and *partially true and partially false* and against *true* and *both true and false*. The answers for the description containing *rich* were similar, except that participants preferred *false* to *partially true and partially false*. Since they were only interested in the effect of the presence/absence of *definitely*, they did not discuss the specific truth-value judgments nor provide statistical analysis for them. They also conducted a variant of the experiment where they investigated the role of the negation, where they simply substituted descriptions involving negations to the affirmative descriptions. One of the descriptions involving negations was *Susan is rich and Susan is not rich*. A vast majority of participants judged this description *false*. This experiment did not investigate the description *Susan is neither rich nor not rich* however. Even though these results are hard to interpret, in part because the experiments were not designed and the analyses were not driven with the same questions as our current questions in mind, they seem compatible with a view where gluttony judgments are dispreferred to gappy judgments.

## 7.2 Two Conceptions of Antonyms

Given the data discussed in the previous paragraph, one question is whether there is a real contrast between gappy and gluttony judgments for vague predicates, and if so, how it might be explained. On this issue, the study of negation and antonyms can prove to be informative. Indeed, the specificity of gappy and gluttony descriptions is that they either affirm or deny an adjective and its negation at a same time. A

classical approach of negation makes these descriptions contradictions, for classically a predicate and its negation are mutually exclusive and cover their whole domain of definition. However, things appear to be more complex for gradable adjectives. There are at least two ways of negating an adjective: i) using a syntactic negation as in the descriptions considered so far (I will call *not tall* the *syntactic antonym* of *tall*), or ii) using a lexical antonym (I will call *short* the *lexical antonym* of *tall*).

(Ruytenbeek 2013) discusses two major theoretical approaches to the question of antonyms: a Gricean approach (as developed by (Horn 1984)) and (Krifka 2007)'s approach. Both approaches exploit the semantic/pragmatic distinction: at the semantic level, a gradable adjective and its syntactic antonym are mutually exclusive and their extensions entirely cover their domain of definition, but at the pragmatic level, their ranges of application get narrowed and we thus observe a gap between them. The main distinction between the two approaches lies in the status they give to *lexical* antonyms. Under the Gricean view, they are pure *contraries*, which means that both antonyms cannot be true of the same entity at the same time: accordingly, their semantic extensions already define a gap under this view. In contrast, they are pure *contradictories* under Krifka's view, which means they cannot be true of the same entity at the same time nor false of the same entity at the same time: syntactic and lexical antonyms are semantically equivalent under this view, they cover their whole domain of definition.

Krifka again exploits the pragmatic level to derive a gap between an adjective and its lexical antonym, whereas the gap is present starting from the semantic level in the Gricean approach. At this point it should be noted that in presenting his approach, Krifka considers the pair of antonyms *happy/unhappy*, where *unhappy* is a morphologically derived lexical antonym of *happy* (in contrast with *sad* which is morphologically atomic).<sup>2</sup> Krifka endorses this position as an epistemicist. Even though he does not make it explicit, this position makes it possible to define the negative prefix of antonyms such as *unhappy* as a classical negation operator. But in the end, both approaches derive the same ranges of applications for adjectives and their antonyms. Figure 7.2 exemplifies the two approaches by considering the triplet (*tall, short, not tall*): the semantic extensions are represented at the top and the pragmatic ranges of applications are represented at the bottom.<sup>3</sup>

---

<sup>2</sup>However, there is one pair of antonymic determiners (*many/few*) in the list of examples that he gives.

<sup>3</sup>Krifka attacks the Gricean approach on the claim that it would incorrectly predict a common use of *not unhappy* to refer to a neutral state of happiness, whereas it is actually used to refer to a "typically mild" (*sic*) state of happiness. But this sentence from (Ruytenbeek 2013)'s succinct

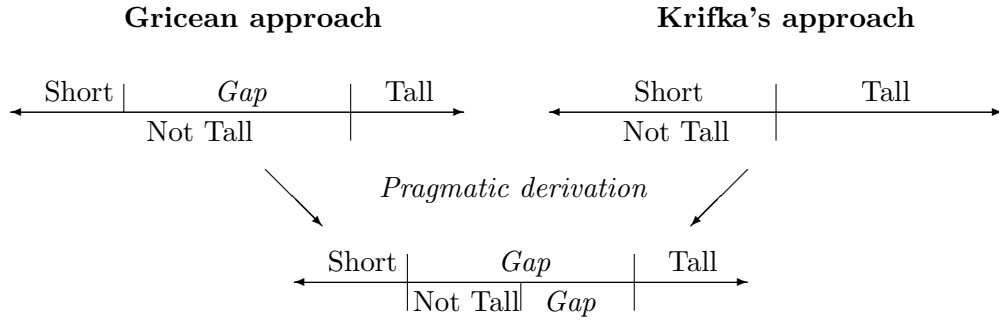


Figure 7.1: Gricean and Krifka's approaches of *tall*, *short* and *not tall*. Both approaches derive the same pragmatic usages but differ on their semantic representations.

Whereas, from a semantic perspective, syntactic antonyms are contradictories and should, as such, yield contradictions when they are used to build gappy and glutty descriptions, they receive a narrower range of applications at the pragmatic level. As things appear to be, an adjective and its antonyms (syntactic or lexical) are mutually exclusive but do *not* cover their entire domain of definition. This makes important predictions about the acceptance of gappy and glutty descriptions. Namely, we should expect gappy descriptions to fit a variety of situations but we should expect the corresponding glutty descriptions to be totally out of order. The experimental results discussed above seem to meet the prediction about gappy descriptions, but they enter in direct conflict with the prediction about glutty descriptions.

Paul Egré and I therefore decided to experimentally test whether gappy descriptions were indeed more appropriate than glutty descriptions when applied to borderline cases. In addition, to the extent that the ranges of applications of an adjective and its lexical antonym seem to lie further away from each other than the ranges of application of an adjective and its syntactic antonym, we expected gappy descriptions built on lexical antonyms to cover a wider range of cases than gappy descriptions built on syntactic antonyms.

### 7.3 Pilot Experiment

Before running a proper experiment, we decided to conduct a pilot experiment. In particular, we were concerned with the question of whether judging the descriptions with one type of antonyms (syntactic or lexical) would have an influence on the

---

presentation of the Horn-Gricean approach suggest that partisans of the latter may disagree with Krifka: “The pragmatic strengthening from ‘medium or short’ to ‘short’ is a case of the general phenomenon of conversational implicature.”.

judgments that participant would then give for the descriptions with the other type of antonyms. In addition, we wanted to investigate whether a yes/no task would be sensitive enough to provide sufficiently fine-grained answers.

### 7.3.1 Design

The pilot experiment was a yes/no task. Subjects were provided with a short text asking them to imagine a borderline case on a scale associated with a pair of antonymous adjectives. Just under the text they were asked to indicate their acceptance of a gappy description and their acceptance of the corresponding glutty description by checking either *yes* or *no*. The two kinds of descriptions were always presented simultaneously in random order and always involved the same type of antonyms. Subjects were divided in two groups: one group of subjects saw a block of items containing the lexical antonyms first and a block of items containing the corresponding syntactic antonyms then; the other group was presented with the same items but with the block in the reversed order. Figure 7.2 exemplifies an item.

Consider the scale of **age**. You have people whose age is very high, and people whose age is very low. Then there are people who lie *in the middle* between these two areas.

Imagine that Sam is one of those people. Can you say the following?

Sam is old and young	<input type="radio"/> Yes	<input type="radio"/> No
Sam is neither old nor young	<input type="radio"/> Yes	<input type="radio"/> No

→ Click here to continue

Figure 7.2: Example of an item from the pilot experiment

The dependent measure was the difference between the ratio of acceptance of gappy and glutty descriptions. This led to a 2\*2 interaction design, with the between-subjects factor **Group** whose modalities were *Lexical first* or *Syntactic first* and the within-subjects, within-items factor **Antonymy** whose modalities were *Lexical* and *Syntactic*.

Alternatively, one can measure the ratios of acceptance of each description type in isolation, adding **Description Type** as a within-subjects, within-items factor whose modalities were *Glutty* and *Gappy*. The former measure is a good choice to analyze

the possible global effect of **Group** and its possible interaction with **Antonymy**. If we observe a strong effect of **Group**, we can decide to focus on the results of the first block of each group to analyze how the descriptions were evaluated when participants had no prior and use the latter measure to investigate a possible global effect of **Antonymy** and its possible interaction with **Description Type**.

### 7.3.1.1 Predictions

As can be seen in Fig. 7.2, none of the two theories of antonyms that we discussed makes an adjective and its antonyms overlap at any level. To this extent, they predict a total rejection of gluttony descriptions. However, one can imagine that each speaker shows some flexibility in where she puts the line separating gradable contradictories. As a consequence, the possible acceptance of gluttony descriptions could result from fixing the line at different positions for the adjective and its contradictory. For instance, one could describe a 1.75cm tall man as “both tall and not tall” by associating the first occurrence of *tall* with a threshold of 1.70cm and the second occurrence of *tall* with a threshold of 1.80cm. Given that syntactic antonyms are contradictories at the semantic level in both theories (i.e. they cover their whole domain), they would accordingly predict a non-zero rate of acceptance of gluttony descriptions conjoining an adjective and its syntactic negation. Regarding lexical antonyms however, only Krifka assumes a level where they are contradictories (the semantic level), whereas they are always contraries for Horn-Gricean approaches. When implemented in Krifka’s theory, the flexibility-of-the-line reasoning would thus predict a relative acceptance of gluttony descriptions conjoining an adjective and its lexical antonym, whereas Horn-Gricean approaches would still consider them as totally out of order. Note that if we make it possible to independently fix the line separating contradictories at each occurrence of a contradictory, so as to get gluttony descriptions, we also expect to get gappy descriptions.

When combining the derivation of a pragmatic gap and the flexibility-of-the-line strategy, we get a picture which is reminiscent of the TCS system developed in (Cobrerros et al. 2012). TCS derives two modes of assertion in addition to the classical one: a strict mode of assertion and a tolerant mode of assertion. One can view the classical mode of assertion as representing the semantic level of gradable vague predicates. When one speaks strictly, one revises the extension of gradable vague predicates so as to exclude borderline cases. When one speaks tolerantly, one revises the extension of gradable vague predicates so as to include borderline cases.



Table 7.1 illustrates the different extensions for *Tall* and *Not tall* in TCS depending on the mode of assertion under consideration.




	Counter instances	Borderline cases	Instances
Classical			
Strict			
Tolerant			

Table 7.1: TCS extensions for *tall* and *not tall* in classical, strict and tolerant modes of assertion. The strict mode defines a gap and the tolerant mode defines a glut.

The classical extensions that TCS assigns to *Tall* and *Not tall* are the same as the ones that Horn-Gricean and Krifka’s approaches give for these expressions at the semantic level. The strict extensions define a gap, as do the pragmatic ranges of applications in the two latter approaches. Importantly, we had to introduce a flexibility-of-the-line mechanism to derive gluts in the latter approaches, but in TCS this comes straightforwardly with the tolerant extensions. The only difference that we’re concerned with between Krifka’s and Horn-Gricean approaches is the semantic extensions they postulate for the lexical antonym (here, *short*). Sticking to the parallel between this notion of semantic extensions and TCS notion of classical extensions, we can represent these two positions in TCS. Table 7.2 illustrates a possible implementation of Horn-Gricean semantic extensions for antonyms in TCS; Table 7.3 illustrates a possible implementation of Krifka’s semantic extensions for antonyms in TCS.

If one views the strict extensions as roughly corresponding to ranges of application that the pragmatic approaches derive, one can see in Table 7.2 and in Table 7.3 that both Horn-Gricean and Krifka’s accounts predict felicitous use of gappy descriptions with either type of antonyms to describe borderline cases (the three black areas overlap). However, when we look at the classical Krifkaian extensions, we see no overlapping, whereas the classical Horn-Gricean extensions make the black areas of the lexical antonyms overlap. When we look at tolerant extensions now, we see that both Horn-Gricean and Krifka’s accounts predict felicitous use of gappy descriptions with *syntactic* antonyms to describe borderline cases (their black areas overlap) but only Krifkaian approaches predict felicitous use of gappy descriptions with *lexical*

	Counter instances	Borderline cases	Instances
Classical	Short		
			Tall
	Not tall		
Strict	Short		
			Tall
	Not tall		
Tolerant	Short		
			Tall
	Not tall		

Table 7.2: Derivation of the strict and tolerant extensions of *short*, *tall* and *not tall* from Horn-Gricean classical extensions. *Short* and *tall* classically define a gap.

antonyms (their tolerant Horn-Gricean extensions do no overlap, contrary to their tolerant Krifkaian extensions).

In sum, both approaches take syntactic antonyms to be contradictories at the semantic level and therefore predict felicitous use of gappy and glutty syntactic descriptions for borderline cases by resorting to the flexibility-of-the-line strategy. But only Krifka’s approach takes lexical antonyms to be contradictories at the semantic level.

Table 7.4 summarizes the acceptability of glutty and gappy descriptions built with syntactic and lexical antonyms, as it is predicted by each approach at the semantic and at the pragmatic level. A “Yes” in a line beginning with “Glutty” or “Gappy” indicates that the description is true of a borderline case at the level and under the approach corresponding to the cell; a “No” indicates that it is false; “With flex” indicates that it is true if we enrich the approach with the flexibility-of-the-line strategy. The lines beginning with “Diff” indicate whether the approach assigns different status to gappy and glutty descriptions at the corresponding level. It appears that the two theories differ only in the status they assign to gappy descriptions with lexical antonyms at the semantic level. As a consequence, Krifka would predict similar observations for all syntactic and lexical descriptions, whereas a Horn-Gricean approach would predict *lexical gappy descriptions* to be more readily accepted than any other description. In conclusion, Horn-Gricean approaches predict an effect of **Antonymy** while Krifka’s approach predicts no effect at all.

	Counter instances	Borderline cases	Instances
Classical	Short		
			Tall
	Not tall		
Strict	Short		
			Tall
	Not tall		
Tolerant	Short		
			Tall
	Not tall		

Table 7.3: Derivation of the strict and tolerant extensions of *short*, *tall* and *not tall* from Krifkaian classical extensions. *Short* and *tall* classically exhaust their domain.

		Horn-Grice		Krifka	
		Semantic	Pragmatic	Semantic	Pragmatic
Syntactic	Glutty	With flex	No	With flex	No
	Gappy	With flex	Yes	With flex	Yes
	<i>Diff</i>	<i>No</i>	<i>Yes</i>	<i>No</i>	<i>Yes</i>
Lexical	Glutty	No	No	With flex	No
	Gappy	Yes	Yes	With flex	Yes
	<i>Diff</i>	<i>Yes</i>	<i>Yes</i>	<i>No</i>	<i>Yes</i>
Effect of antonymy		Yes		No	

Table 7.4: Predictions of Krifka’s and Horn-Grice approaches. The two approaches crucially differ in their semantic representations for lexical antonyms.

## 7.3.2 Methods

Even though we were not familiar with his work when we built the materials for the pilot, (Ruytenbeek 2013) discusses some aspects of gradable adjectives and their antonyms that can play a crucial role in our design. The next subsection is dedicated to these aspects, and the subsection that comes after it presents the materials that were used in the pilot experiment.

### 7.3.2.1 Possible Biases in the Adjectives’ and their Antonyms’ Features

One worry that we had concerning the descriptions with syntactic antonyms was that facing apparently contradictory descriptions, participants might interpret the negated and the non-negated occurrences of the same adjective in two different ways.

For instance, one person could be described as *both scary and not scary* given that some people are scared by this person but others are not, or because this person is scary in some situations but not in others, or because this person fakes his/her scariness. With these readings in mind, the possible acceptance of gappy and glutty descriptions would not have necessarily resulted from the consideration of a borderline case.

We conducted the pilot experiment before being acquainted with (Ruytenbeek 2013)’s thesis, which reviews several aspects of adjectives that can be responsible for such a shifting in their interpretation. Two of them are *subjectivity* and *evaluativity*. An adjective *ADJ* is considered *subjective*, Ruytenbeek says, when sentences of the form *X finds Y ADJ* are felt natural. For instance, *tall* is subjective because *John finds Bill tall* is felt natural, but *even* is not, because *John finds 2 even* sounds unnatural. Besides, an adjective *ADJ* is considered *evaluative*, Ruytenbeek says, when sentences of the form *X finds Y ADJ-er/more ADJ than Z* are felt natural. For instance, *beautiful* is evaluative because *John finds Bill more beautiful than Sam* is felt natural, but *tall* is not, because *John finds Bill taller than Sam* sounds degraded.

A possible core difference between subjective and evaluative adjectives is that the scale associated with the former would be the same for any speaker (even though each speaker would still have liberty in placing a threshold on this constant scale), whereas the very criteria used to build the scale associated with the latter may change depending on the evaluator. Indeed, there is just one criterion used to build the scale of heights, which precisely is height.<sup>4</sup> But there are several, maybe even an infinity of criteria used to determine whether someone is beautiful or not. All gradable adjectives that we considered succeed in the test of subjectivity.<sup>5</sup> This means that the adjectives we used in our experiment were subjective, and that in theory subjects had the possibility to shift their interpretation in the descriptions with syntactic antonyms. Even though some contextualists could claim that one speaker can actually interpret two occurrences of the same subjective adjective in one sentence with two different thresholds in mind, we suspect that subjectivity is constant through the evaluation of one sentence, and for this reason we assume that using subjective adjectives did not *per se* favor the acceptance of conflicting descriptions. As for evaluativity, the 8 adjectives that we used were all non-evaluative. This means that in theory participants did not have the possibility to imagine different criteria to build two different

---

<sup>4</sup>I here focus on the totally ordered aspect of the scale, regardless of ratio/metric considerations.

<sup>5</sup>This suggests that the two property might be tightly connected. (Burnett 2012)’s typology of gradable adjectives can be seen as an approach exploiting this connection.

scales for the evaluation of the two occurrences of the adjective.

We were also concerned with what could be called a *stage/individual* level distinction (not discussed in (Ruytenbeek 2013)). We say that *scary* is a *stage level* adjective because the measure it returns for a given entity changes over time and situations. In contrast, we say that *tall* is an *individual level* adjective because the measure it returns for a given entity (at least for adults) is constant over time and situations. Of course this distinction is itself gradable: some adjectives neither clearly situate at an individual level nor at a stage level. For instance the measure that *smart* returns for a given person can vary across situations but not really across time: one person can be smart with orienting by reading a map but not smart when it is about orienting by looking at the stars, however it is not very plausible to imagine that this person is smart one day but not smart the day after. We consequently tried to select adjectives that were closer to the individual (constant) level than to the stage (variable) level.

In building the descriptions with syntactic antonyms, we also had to choose an adjective from the pair of lexical antonyms. It turns out that we always chose the unmarked member of the pair even though we were not aware of the criterion we used to make our decision. (Ruytenbeek 2013)’s criterion to tag a non-evaluative adjective *ADJ* as marked is the inference from the equative construction *X is as ADJ as Y* to *X and Y are ADJ*. For instance, *tall* is not marked because the inference from *John is as tall as Bill* to *John and Bill are tall* does not hold, but *short* is marked because the inference from *John is as short as Bill* to *John and Bill are short* holds.

In addition, a lexical antonym can be positive or negative. Ruytenbeek investigates two diagnostics for the positivity of an adjective *ADJ* in his experiments: an acceptability judgment for sentences of the form *X n’est pas très ADJ* (French for *X is not very ADJ*) and an acceptability judgment for exclamations of the form *C’est fou à quel point X n’est pas ADJ !* (French for *It’s crazy how X is not ADJ!*). For non-evaluative adjectives, he found a correlation between positivity and acceptance of the exclamations but no correlation between positivity and acceptance of constructions with *pas très*. For instance, *tall* is positive whereas *short* is negative, because the exclamation *It’s crazy how John is not tall!* appears more acceptable than the exclamation *It’s crazy how John is not short!*<sup>6</sup> Positivity might influence what enti-

---

<sup>6</sup>Note that when we remove the negation as in (i), the contrast between the two antonyms disappears.

- (i) a. It’s crazy how John is tall!
- b. It’s crazy how John is short!

It is therefore crucial to keep the negation in the exclamations to establish a diagnostic of positivity

ties you count as borderline. For instance, someone whose incomes are in the mean of the population’s incomes may count as a borderline case for the positive adjective *rich* but not as a borderline case for the negative adjective *poor* for people would be reluctant to assign negative properties. It seems that all the adjectives we used in the descriptions with syntactic antonyms yield acceptable exclamations, suggesting that we always chose the positive member of the pair. However, it should be noted that whether one accepts these exclamations as natural or not highly depends on one’s expectations concerning the measure that the adjective is about. For instance, the exclamation *It’s crazy how this soviet building is not big!* seems more natural than the exclamation *It’s crazy how this soviet building is not small!* because soviet buildings tend to be big, suggesting that *big* would be positive whereas *small* would be negative. But the exclamation *It’s crazy how this electronic chip is not big!* sounds much less natural than the exclamation *It’s crazy how this electronic chip is not small!* because electronic chips tend to be small, suggesting that *small* would be positive whereas *big* would be negative. In addition, depending on the season, both the exclamation *It’s crazy how the weather is not hot today!* and the exclamation *It’s crazy how the weather is not cold today!* sound natural, suggesting that both *hot* and *cold* are positive. These observations may suggest that positivity is a highly context-dependent property, but for our concerns this has the consequence to diminish the control we have on the positivity of the adjective we choose for syntactic antonyms.

### 7.3.2.2 Materials

As noted earlier, when we built the materials for the pilot study the only above consideration that we had in mind was the distinction between stage-level and individual-level adjectives. Fortunately, the adjectives we chose satisfy the desiderata we have to conduct such an experiment. They are reported in Appendix B.2

We used 8 pairs of antonyms to get 8 measures per subject for each condition (i.e. 16 data points per subjects). The lexical antonyms were all morphologically simple. The corresponding syntactic antonyms were always built on the unmarked member of the lexical antonyms. As noted above, the positivity of an adjective is not very easy to determine, but the members we used to build the syntactic antonyms were apparently all positive too. The 16 resulting items were randomly ordered inside each block, and the gluttony and the gappiness descriptions were randomly ordered for each item. There was no control (because we were interested in the differences in acceptance) and no filler.

---

for an adjective.

We paid particular attention to the texts that we used to make participants imagine borderline cases. Because we asked participants to imagine an entity on the middle of a scale, we had to make sure that such entities would indeed be borderline cases for the adjectives that were used in the descriptions. But this would not have systematically obtained if we had referred to the scale with the nominalization of the adjective that we tested. For instance, an individual whose age lies in the middle of the scale of *oldness* could maybe count as a borderline case of *old* but it is less obvious that this individual would count as a borderline case for *young*. And symmetric considerations apply to the scale of *youth*. Rather, in this case, we chose to talk about the scale of *ages*, where an individual who lies in the middle would reasonably count as a borderline case both for *old* and *young*. In addition, we thought that using to a neutral noun would help people adapt the scale under consideration in order to pick an entity that would make the descriptions as felicitous as possible regardless of whether they involved syntactic or lexical antonyms.

### 7.3.2.3 Procedure

40 participants were recruited on Amazon Mechanical Turk. Two announcements were created, one for each group of subjects. Subjects were warned that they would not be paid if they participated to both versions.<sup>7</sup>

There was no training session and they had to answer two series of 8 items each with no transition screen between the two series. At the end of the experiment, participants had to enter their Amazon Mechanical Turk ID and to indicate whether they were native speakers of English to validate their participation. They could also indicate their sex (Male or Female), their age and leave a comment.

35 Amazon workers (no requirement specified) participated whereas each announcement was designed for 20 participants. It turned out that 5 participants went through both versions of the experiment: they were paid only once and were excluded from the analysis. 15 of the remaining participants saw the block of the 8 lexical antonyms first and the block of the 8 syntactic antonyms then, and the 15 others saw the two blocks in the reversed order. 14 participants identified themselves as females, 13 as males and 3 did not indicate their sex affiliation. 24 participants out of 30 defined themselves as native speakers of English. Participants reported to range between 20 and 52 year old. The average duration of the experiment was 3min53".

---

<sup>7</sup>The text they read before the experiments is reported in Appendix B.1.

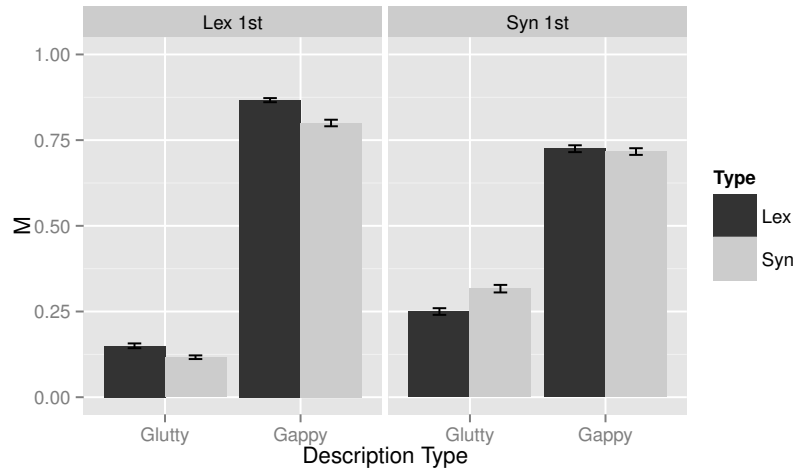


Figure 7.3: Mean acceptance rates of glutty and gappy descriptions across group and antonymy. The left column and the right column respectively correspond to the participants who were first presented with the descriptions built with the lexical and the syntactic antonyms.

### 7.3.3 Results

As shown in Figure 7.3, participants globally accepted gappy descriptions more often than they accepted glutty descriptions. Figure 7.4 plots the difference between the acceptance rates of gappy and glutty descriptions across **Group** and **Antonymy**: 1 corresponds to a total acceptance of gappy descriptions and a total rejection of glutty descriptions, 0 corresponds to an equal treatment of gappy and glutty descriptions (possibly including total acceptance of total rejection of both), and -1 would correspond to a total acceptance of glutty descriptions and a total rejection of gappy descriptions. It reveals a strong effect of **Group** but no clear effect of **Antonymy**. Figure 7.5 represents the distribution of these differences across the conditions: participants tended to have a maximal contrast but some of them had more nuanced judgments, and some participants even turned out to prefer glutty descriptions over gappy descriptions and thus have a negative score. Figure 7.6 plots the variance of these differences: it was low for the group who saw the block of lexical antonyms first and high for the group who saw the block of syntactic antonyms first. Figure 7.7 focuses on the rates of acceptance in the first block of each group of participants and reveals an interaction between **Antonymy** and **Description Type**.

### 7.3.4 Discussion

The clear global effect of **Group** suggests that the evaluation of the antonyms in the second block was influenced by the evaluation of the antonyms in the first block.



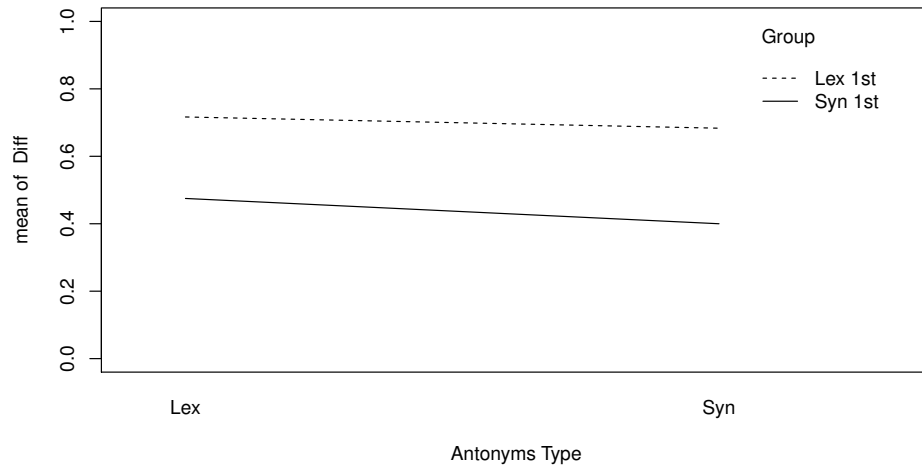


Figure 7.4: Mean contrasts between glutty and gappy descriptions across group and antonymy. The lines report the differences between the mean acceptances of the gappy descriptions and of the glutty descriptions in each condition. The dashed line and the plain line respectively correspond to the participants who were first presented with the descriptions built with the lexical and the syntactic antonyms.

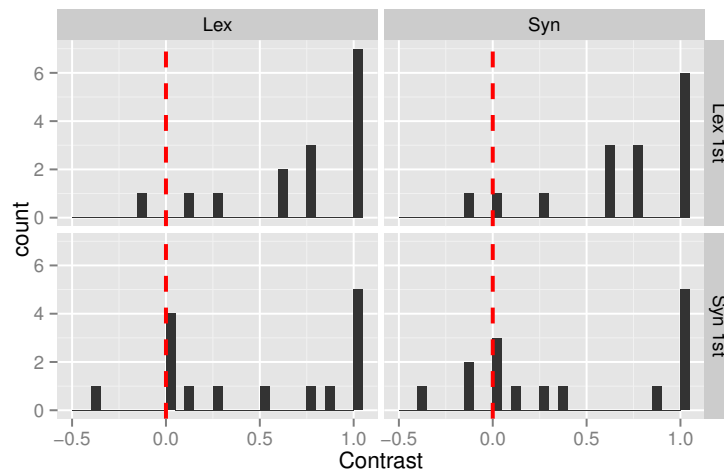


Figure 7.5: Distribution of the contrasts across the conditions. The negative contrasts, on the left of the red dashed lines, report a preference for glutty descriptions over gappy descriptions. The first and the second rows respectively correspond to the participants who were first presented with the descriptions built with the lexical and the syntactic antonyms.

None of the two theories we discussed directly predicts this pattern. However they consider that the range of applications of syntactic antonyms is wider than the range of applications of lexical antonyms: participants who evaluated the lexical antonyms first might have accordingly narrowed the range of applications of the syntactic antonyms,

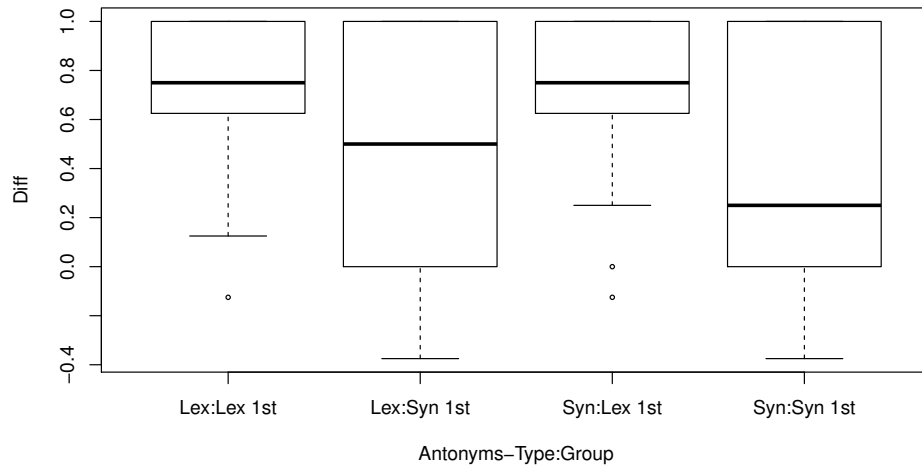


Figure 7.6: Variation of contrasts between glutty and gappy descriptions across group and antonymy. The contrasts of the participants who were first presented with the descriptions built with the lexical antonyms (1<sup>st</sup> and 3<sup>rd</sup> boxes) are less scattered than those of the participants who were first presented with the descriptions built with the syntactic antonyms (2<sup>nd</sup> and 4<sup>th</sup> boxes).

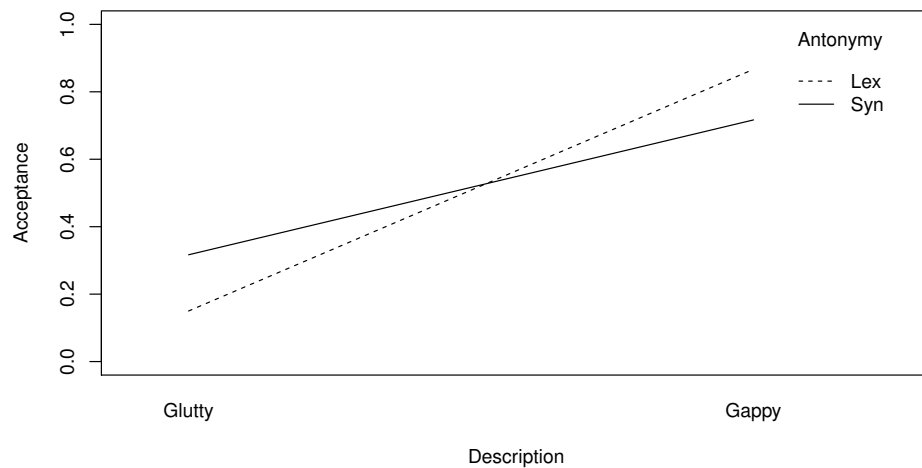


Figure 7.7: Mean acceptance rates of glutty descriptions and gappy descriptions in the first block of each group across antonymy. The dashed line and the plain line respectively correspond to the participants who were first presented with the descriptions built with the lexical and the syntactic antonyms.

thus showing a strong contrast between the gappy and the glutty descriptions for both types of antonyms. This does not explain though why participants who were presented with the syntactic antonyms first showed a lower contrast between the gappy and the

glutty descriptions with lexical antonyms. Indeed, since the range of applications of lexical antonyms is narrower than the range of applications of syntactic antonyms, it would have to be *widened* to make the two correspond. Horn-Gricean theories are not in a good position to account for this: under this approach, lexical antonyms are pure contraries and therefore cannot make glutty descriptions licit. In contrast, the two types of antonyms are initially synonymous under Krifka's approach, making it in a better position to explain how the first evaluation of either type can influence the evaluation of the other one.

To the extent that this contrast stands as evidence for the flexibility in the interpretation of either type of antonyms, it is worth noting that asking participants to evaluate them in two specific successive blocks did not favor a contrast between them, even though it could have been expected. It is reasonable to assume that the results for the first blocks in each group represent the typical (non-influenced) interpretation of the two types of antonyms. Consequently, this calls for a between-subject comparison of the two sets of results. Figure 7.7 compares the results for the first block in each group (i.e. between-subjects results). It shows an effect of **Antonymy**: the judgments that participants gave for glutty and gappy descriptions tended to be more contrastive for the lexical descriptions than for the syntactic descriptions (again, these are by-subjects results). Fig. 7.7 further reveals that this effect of **Antonymy** on mean contrasts actually corresponds to an interaction between **Antonymy** and **Description Type** on mean acceptance rates: the participants who judged the glutty descriptions with the *lexical* adjectives in the first block rejected them more often than the participants who judged the glutty descriptions with the *syntactic* adjectives in the first block; but the participants who judged the gappy descriptions with the *lexical* adjectives in the first block rejected them less often than the participants who judged the gappy descriptions with the *syntactic* adjectives in the first block.

These preliminary results seem to favor a Horn-Gricean approach which predicted an effect of **Antonymy**. Indeed, that the mean rate of acceptance of glutty descriptions increased when going from lexical to syntactic descriptions was expected: there is no level where lexical glutty descriptions are to be accepted for Horn-Griceans, whereas syntactic glutty descriptions are acceptable at the semantic level (with the flexibility-of-the-line strategy). In contrast, Krifka's account predicts both lexical and syntactic glutty descriptions to be acceptable at the semantic level, and we should therefore observe no significant difference between them. Besides, the Horn-Gricean derivation could explain that gappy descriptions are less accepted with syntactic

antonyms than with lexical antonyms: at the semantic level, syntactic gappy descriptions are only acceptable when resorting to the flexibility-of-the-line strategy whereas lexical gappy descriptions are already perfectly fine. This contrast does not obtain in Krifka’s configuration, where both types of descriptions need the strategy to be acceptable.

## 7.4 Main Experiment

### 7.4.1 Design

We used the same design as in the pilot experiment, but we added a third, control description to each slide. The control descriptions were always of the form “X is extremely ADJ” or of the form “X is not extremely ADJ”, where ADJ corresponded to the target adjective in the blocks testing the lexical antonyms and to the lexical antonym in the blocks testing the syntactic antonyms. The purpose of these control items was twofold. First, we wanted to analyze the acceptance rate of glutty descriptions with lexical antonyms, an aspect which does not appear to have been tested before our study. A significant acceptance of glutty descriptions with lexical antonyms would dismiss a Horn-Gricean approach (to the extent that it views lexical antonyms as pure contraries, thus forming a gap even at the semantic level) and favor Krifka’s approach (to the extent that it views lexical antonyms as semantic contradictories and with the flexibility-of-the-line strategy in mind). Second, we suspected that always presenting subjects with both types of antonyms would neutralize the global **Group** effect, leaving the **Antonymy** effect intact. Figure 7.8 illustrates an item.

#### 7.4.1.1 Predictions

The predictions are basically the same as in the pilot: Horn-Gricean approaches predict an effect of **Antonymy**, Krifka’s account does not. From the addition of control descriptions which contain the type of antonym absent from the target descriptions co-occurring on the same item, we expect no order effect, that is to say no order of **Group**. If there were an effect of **Group** though, we would focus on the results from the first blocks of each group and conduct a between-subjects analysis. Given that the extensions of syntactic antonyms can overlap at the semantic level in both approaches (with the addition of a flexibility-of-the-line mechanism), we expect syntactic glutty descriptions to be accepted more often than false control sentences. For the same reason, we expect syntactic gappy descriptions to be accepted less often than true

Consider the scale of **age**. You have people whose age is very high, and people whose age is very low. Then there are people who lie *in the middle* between these two areas.

Imagine that Sam is one of those people. Can you say the following?

Sam is neither old nor not old    ☐ Yes    ☐ No

Sam is not extremely old        ☐ Yes    ☐ No

Sam is old and not old           ☐ Yes    ☐ No

→ [Click here to continue](#)

Figure 7.8: Example of an item from the main experiment

control sentences. Contrary to Krifka’s position augmented with the flexibility-of-the-line, the Horn-Gricean position claims that the extensions of lexical antonyms can never overlap, and therefore predicts lexical glutty descriptions to be as rarely accepted as false control sentences and lexical gappy descriptions to be accepted as often as true control sentences.

## 7.4.2 Methods

### 7.4.2.1 Materials

We used the same materials as in the previous experiment, with the addition of the control descriptions. Each block contained 4 affirmative control descriptions (expected to be false) and 4 negative control descriptions (expected to be true). The polarity of the control descriptions was counterbalanced between items: in particular, the items that contained an affirmative control description in the first block contained a negative control description in the second block. The descriptions of the form *X is extremely ADJ* were expected to be false in the context provided, given that *X* was always described as an entity “in the middle” between two extremes on the scale associated with *ADJ*. Correspondingly, the negative counterparts of these descriptions, *X is not extremely ADJ*, were expected to be true. It may be that the negative controls came with an implicature that *X is somewhat ADJ* and that it could bias their control status.

### 7.4.2.2 Procedure

80 participants were recruited via Amazon Mechanical Turk platform. We used an announcement of the same form as in the pilot. But this time, we used it a first time to redirect subjects to a version of the experiment where the first block tested the lexical antonyms, and after 40 subjects completed the experiment we used the same announcement a second time to redirect subjects to a version of the experiment corresponding where the first block tested the syntactic antonyms. Using this procedure on Amazon Mechanical Turk allowed us to ensure no subject participate to both versions. Every participant was paid \$0.25 for an average time of 4min46. 8 participants were excluded from the analyses because they also took part to previous versions of the experiment. Among the 72 remaining participants (36 per group), we only considered those who scored with an accuracy of at least 75% on the control descriptions to conduct the analyses. Because of a probable flaw in the design, 46 participants were thus excluded and we ended up with two groups of 13 participants. 8 identified themselves as females, 16 as males and 2 did not indicate their sex affiliation. 22 participants out of 26 defined themselves as native speakers of English. These 26 participants reported to range between 21 and 61 year old.

### 7.4.3 Results

The results of the pilot experiment were replicated: the addition of the control descriptions seems to have had no significant impact on the rates of acceptance of the target descriptions (regardless of the inclusion or the exclusion of the non-accurate participants). Figure 7.9 shows that the 26 retained participants again globally accepted the gappy descriptions more often than they did the glutty descriptions. The mean acceptance of the gappy descriptions did not seem to differ from the mean acceptance of the true control descriptions in 3 of the 4 conditions. The participants who judged the former in the first block considered them *true* less often than they did the true control descriptions though. The participants who saw the lexical block first rejected the gappy descriptions as *false* about as often as they did the false control descriptions, but the participants who saw the syntactic block first rejected them as *false* less often than they did the false control descriptions. In every condition, the mean acceptance of the gappy descriptions differed from the mean acceptance of the glutty descriptions.<sup>8</sup> One can see in Figure 7.10 that there is once again a strong

---

<sup>8</sup>Adding the non-accurate subjects to the analyses has the effect that false control descriptions are globally accepted more often than glutty descriptions, and that true control descriptions are globally accepted less often than gappy descriptions.

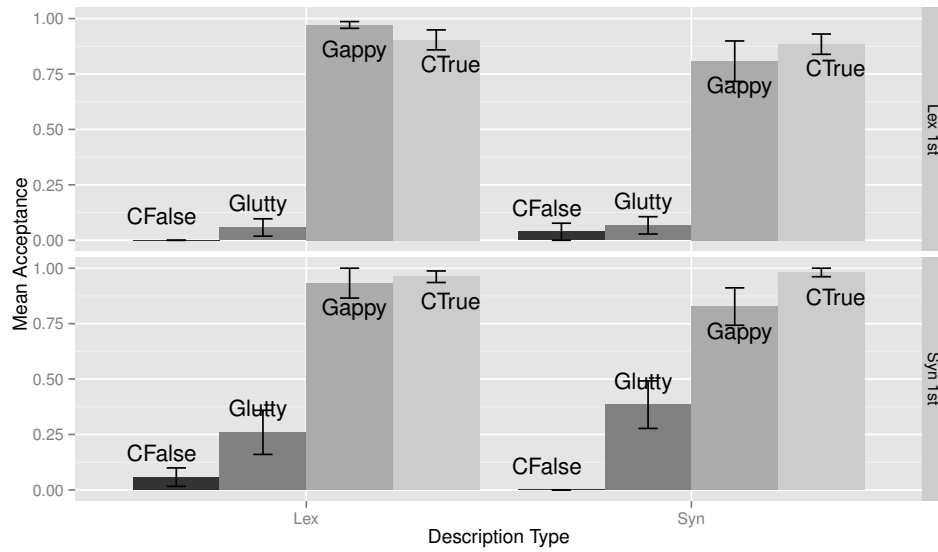


Figure 7.9: Mean acceptance rates of glutty, gappy and control descriptions across group and antonymy. The top row and the bottom row respectively correspond to the participants who were first presented with the descriptions built with the lexical and the syntactic antonyms.

effect of **Group**. The slope of the parallel lines is greater than in the pilot experiment and seems to confirm an effect of **Antonymy** which was unclear in the pilot. Figure 7.11 plots the variance of these differences: it was very low for the lexical descriptions evaluated by the 13 participants who saw the lexical block first, and rather important for the syntactic descriptions evaluated by the 13 participants who saw the syntactic block first. Figure 7.12 focuses on the rates of acceptance in the first block of each group of the 26 retained participants and reveals again an interaction between **Antonymy** and **Description Type**.

#### 7.4.4 Discussion

The first remark that we should make is that the material we added globally failed as control elements: about 2/3 of the participants showed a low accuracy on their judgments for these descriptions. Importantly, when we look at the distribution of the contrasts for the true control and the false control descriptions, two groups emerge: a group of participants with a contrast of 100% (these are the participants who score high on accuracy) and a group of participants with a contrast of 0%. A contrast of 0% means that the participants judged as many *true* judgments for the true control descriptions as they did for the control *false* descriptions. A possible explanation for this observation is that these participants did not pay attention to the presence or

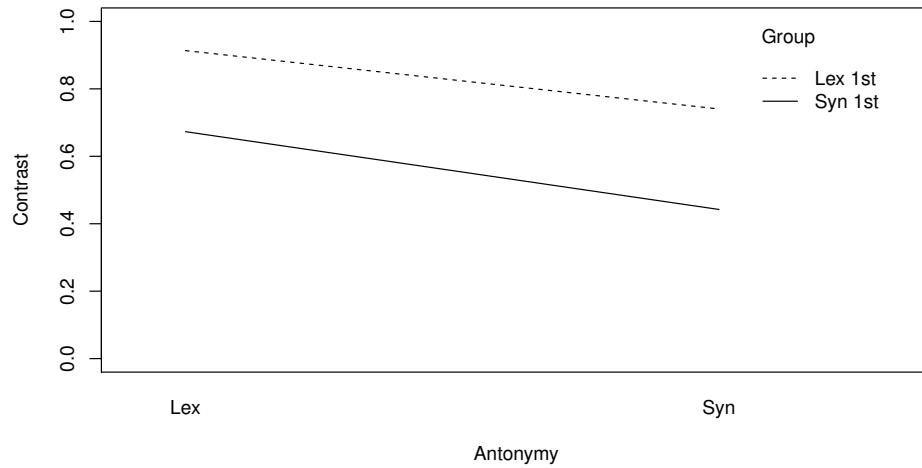


Figure 7.10: Mean contrasts for glutty and gappy descriptions across group and antonymy. The lines report the differences between the mean acceptances of the gappy descriptions and of the glutty descriptions in each condition. The dashed line and the plain line respectively correspond to the participants who were first presented with the descriptions built with the lexical and the syntactic antonyms.

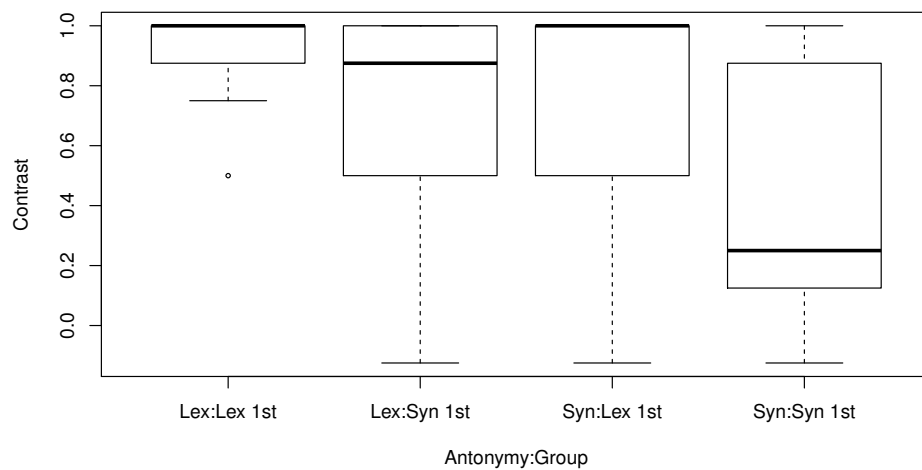


Figure 7.11: Variation of contrasts for glutty and gappy descriptions across group and antonymy. The contrasts that the participants reported for the descriptions built with the lexical antonyms in the first block (1<sup>st</sup> box) are less scattered than those that the participants reported in the other conditions. On the contrary, the contrasts that the participants reported for the descriptions built with the syntactic antonyms in the first block (4<sup>th</sup> box) are more scattered than those that the participants reported in the other conditions.

the absence of negation.



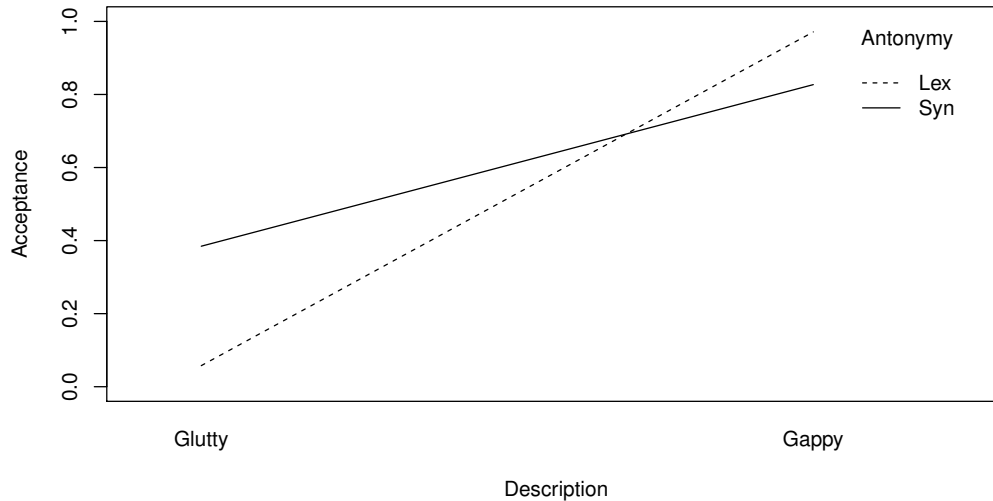


Figure 7.12: Mean acceptance rates of glutty descriptions and gappy descriptions in the first block of each group across antonymy. The dashed line and the plain line respectively correspond to the participants who were first presented with the descriptions built with the lexical and the syntactic antonyms.

Second, adding control descriptions containing an antonym of the other type than the one used in the target descriptions failed to neutralize the order effect: Fig. 7.10 shows a clear effect of **Group**). For this reason and as explained earlier, a between-subjects analysis of the results of the first blocks of each group will be more informative on the differences between lexical and syntactic antonyms. When we look at the leftmost and the rightmost boxes in Fig. 7.11, the effect of **Antonymy** appears clearly: the 13 control-accurate participants who evaluated the lexical descriptions in the first block had a maximal contrast between gappy and glutty descriptions, whereas the values for the 13 control-accurate participants who evaluated the syntactic descriptions in the first block broadly distribute around a low contrast.

These different contrasts result from an interaction between **Antonymy** and **Description type** as can be seen on Fig. 7.12: glutty descriptions globally tend to be rejected and gappy descriptions globally tend to be accepted, but both these tendencies are diminished for syntactic descriptions and exacerbated for lexical descriptions. The top-left graph from Fig. 7.9 suggests that the lexical glutty descriptions did not differ significantly from the false control descriptions and that the lexical gappy descriptions did not differ from the true control descriptions. This is in accordance with a maximal contrast between the target descriptions. The bottom-right graph from

Fig. 7.9 reveals that each kind of description differed from each other in the syntactic first block. Importantly, the syntactic gluttony descriptions seem to have been accepted a significant amount of times. This result argues in favor of a Horn-Gricean approach where the extensions of lexical antonyms never overlap and can therefore never yield true gluttony descriptions but where the extensions of syntactic antonyms can overlap at the semantic level with the help of a flexibility-of-the-line mechanism and can therefore yield true gappy descriptions. Given that Krifka posits the same semantic extensions for lexical and syntactic antonyms and that the flexibility-of-the-line mechanism would operate on the semantic extensions, his approach would predict lexical gluttony descriptions to be accepted as often as syntactic gluttony descriptions, contrary to what we observe.

# Chapter 8

## Conclusions

### 8.1 Theoretical Conclusions

I started this thesis with the aim of providing a theoretical explanation for the observation of non-bivalent truth-value judgments as a response to both vague and presuppositional sentences. However, I also wanted this explanation to take the specificities of vagueness and the specificities of presupposition into account. I started the discussion in the introduction by exposing two main strategies to guide the elaboration of a system that would jointly address both phenomena. The first strategy consists in defining a general algorithm that would apply equally to vague and presuppositional sentences, whereas the second strategy consists in defining specific algorithms to treat each type of sentence. Based on its generic aspect, it seemed that the first strategy would easily account for the observation of non-bivalent truth-value judgments in response to both vague and presuppositional sentences but would fail to distinguish between these two types of sentences. On the contrary, by defining specific algorithms dedicated to each phenomenon, the second strategy seemed in a better position to model vagueness and presupposition as distinct phenomena; it would, however, likely miss the commonality between vagueness and presupposition, namely that they both trigger non-bivalent truth-value judgments.

With the elaboration of ST5, a totally ordered 5-valued system, I showed that it was possible to give a truth-functional account of vagueness and presupposition along the lines of the first strategy. Contrary to what one might have expected, defining non-bivalent truth-values along a total order and providing a general principle to derive truth-value judgments did not result in an indistinguishability of the two phenomena. Unexpectedly, in light of the results of the experiments that I presented in Chapter 5, ST5 in fact even predict unobserved differences between vague and presuppositional sentences. More precisely, the system predicts that in situations of presupposition

failure, the truth-value judgments of presuppositional sentences should be sensitive to polarity; this prediction was not borne out. In experiment 2 in particular, participants did not choose to label the negative presuppositional sentences as “Completely true” in the critical contexts. In Chapter 7, I presented an experiment on gappy and glutty descriptions conducted in collaboration with Paul Egré. The results that we obtained can be taken as evidence for the possibility to give glutty truth-value judgments (i.e. judgments such as “Both true and false”) in reaction to vague descriptions of borderline cases. With these considerations in mind, remember that I explained in the discussion of Chapter 5 that ST5 was unable to derive gappy truth-value judgments for vague descriptions of borderline cases without also deriving “True” judgments for negative sentences with unfulfilled presuppositions. As a consequence, ST5 necessarily derives “True” judgments for negative sentences with unfulfilled presuppositions to the extent that it has to account for the “Both true and false” judgments for vague descriptions of borderline cases.

However, the results of the experiments discussed in Chapter 3 do not exclude that negation *can* have an impact on sentences with unfulfilled presuppositions. In particular, we saw that a *local accommodation* reading of negative presuppositional propositions is discussed in the literature, and the absence of evidence for this reading in my experiments does not mean that it is *never* to be observed. As mentioned in Chapter 3, (Chemla & Bott 2013) actually collected empirical data about local accommodation under negation and compared them with their data for global accommodation. On this point, it is worth noting that ST5 deems local accommodation readings as natural in situations of presupposition failure: a “True” judgment is naturally expected for negative sentences with unfulfilled presupposition, but not a “False” judgment. Because ST5 therefore needs an *additional* mechanism to derive “False” judgments, it would reasonably predict local accommodation readings to be faster to access than global accommodation readings. However, Chemla & Bott’s precisely obtained opposite results.

These considerations justified the exploration of the second strategy, that is to say defining a system which would dedicate an algorithm to vagueness and another algorithm to presupposition. In Chapter 6, I brought ontological grounds for a unified system incorporating specific algorithms. I showed that merging an algorithm for vagueness that derives the Strong Kleene truth-tables with an algorithm for presupposition that derives the Weak Kleene truth-tables results in an algorithm which itself derives the truth-tables that I had otherwise derived from a partially ordered 4-valued bi-lattice. This bi-lattice unifies vagueness and presupposition in that they both enter

in relation with plain truth and plain falsity, but it also accounts for their different status by defining this relation along a different dimension for each phenomenon. This bi-lattice also allowed me to straightforwardly define a total order which associates sentences whose presuppositions are unfulfilled with the lowest value ( $\mathcal{P}$ ), and vague descriptions of borderline cases with a middle-top value ( $\mathcal{V}$ ). Bearing in mind that vague sentences give rise to “True” judgments in a greater variety of situations than presuppositional sentences do, this configuration of truth-values constitutes further support for a system in line with this partially ordered 4-valued lattice. By deriving a total order on which the definitions of the logical vocabulary of the language rely, this approach shares with ST5 the virtue of making direct predictions about sentences containing both vague and presuppositional expressions (*hybrid* sentences) in a non-stipulative way. However, this non-stipulative aspect collapses once we revise the Weak Kleene algorithm dedicated to presupposition to take the linearity aspect of presupposition into account by substituting an algorithm that derives the Middle Kleene truth-tables. After doing so, we have to stipulate either that the final output should correspond to the algorithm dedicated to presupposition or that the final output should correspond to the algorithm dedicated to vagueness.

## 8.2 Theoretical Perspectives

From a more general perspective, I showed how the 4-valued systems investigated in Chapter 6 result from merging two algorithms dedicated to each phenomenon. But one could imagine that these specific algorithms in fact operate at a pragmatic level, building on a bivalent semantic representation of each type of sentences. Under this view, the 4-valued systems I discussed would in fact correspond to a semanticization of a set of pragmatic processes at play when a speaker evaluates a sentence. But other pragmatic approaches of each phenomenon exist. For instance, some authors view vagueness as ignorance (Williamson 1994), and some authors approach presupposition in the same way as they approach implicatures (Chemla 2009). It would be interesting to see how these pragmatic views interact and what predictions accordingly obtain in regard of hybrid sentences.

This thesis has focused on two possible truth-functional systems that would jointly address vagueness and presupposition, but it leaves several other theoretical options unexplored. By modeling vagueness with a non-bivalent truth-value, I did not address the particular question of what representation it should receive in the general

framework of dynamic semantics. Indeed, one core idea of the dynamic system proposed by (Heim 1983) is to replace the traditional truth-functional composition of meaning with a compositional approach in terms of updates on the contexts. Once we adopt a formalization of contexts as sets of possible worlds, we face the question of how to incorporate vague descriptions of borderline cases in this system, given that they were previously modeled as propositions of an intermediate value. One possible answer that would deserve attention consists in using (Pawlak 1997)’s notion of rough sets to model context sets capable of dealing with beliefs on borderline cases.

### 8.3 Empirical Conclusions

This thesis has presented two experimental studies that investigate speakers’ attitude toward non-bivalence. The experiments presented in Chapter 5 directly targeted non-bivalent truth-value judgments for vague and presuppositional sentences; and with the experiment presented in Chapter 7, Paul Egré and I collected bivalent truth-value judgments for contradictory descriptions built with vague adjectives. A general observation is that, as expected, speakers showed a *non-bivalent behavior* across all of these experiments: they gave non-bivalent truth-value judgments for both types of sentences in the first study, and they accepted to describe borderline cases with both gappy and glutty descriptions in the second study.

Overall, the results I obtained for presuppositional sentences in the first study replicate the results that (Abrusán & Szendrői 2013) obtained in their experiment on affirmative and negative presuppositional sentences. In particular in my second experiment, participants distributed their answers over “Completely false” and “Neither” for the presuppositional negative sentences in the critical contexts. In comparison, Abrusán & Szendrői observed a majority of “False” judgments and a certain amount of “Can’t say” judgments for sentences like (75). Importantly, for these sentences and in these conditions, subjects basically gave no “Completely true” judgment in my experiment and subjects gave few “True” judgments in Abrusán & Szendrői’s experiment.

(75) The king of France is not bald

On the other hand, (Chemla & Bott 2013) did obtain data points for a “True” judgment of sentences like (76), associated with the false presupposition that elephants are reptiles.

(76) Zoologists do not realize that elephants are reptiles.

This cross-experiment difference can be due to the different designs that were used, but also to the specificities of the presuppositional expressions that were used. If this last explanation is correct, that the results of my second experiment align with Abrusán & Szendrői's results suggests that the aspectual verb *stop* and the definite article *the* would belong to a first class of presuppositional expressions whereas the factive verb *know* would belong to a second class of presuppositional expressions. In any case, all these experiments confirm that “False” judgments are observable both for the affirmative and the negative counterparts of sentences with unfulfilled presupposition. In addition, they lead to the conclusion that if “True” judgments are available too, they are strongly dispreferred. As a consequence, and as I recalled above, these results argue against a theory of presupposition which give a privileged status to a local accommodation reading of negative presuppositional sentences. This is the case of ST5, but also in general of pragmatic approaches (see the studies discussed in Chapter 3 for a discussion of how pragmatic theories of presupposition can deal with these observations). The specificity of ST5 however is that this undesired treatment of negative presuppositional sentences comes along with a supposedly desired treatment of affirmative and negative vague sentences as triggering both glutty and gappy judgments. Even though the patterns of answers that I obtained for vagueness in my experiments are extremely clear and establish that affirmative and negative vague sentences both trigger non-bivalent truth-value judgments, the status of these non-bivalent truth-value judgments depend on how we interpret a click on the middle button for descriptions of borderline cases. Importantly, nothing in the designs that I used forces us to interpret these clicks as signaling a glutty truth-value judgments (“Both true and false”).

However, the experiment reported in Chapter 7 provides indirect evidence that speakers might access both gappy and glutty truth-value judgments when evaluating vague descriptions of borderline cases. The results that Paul Egré and I obtained with this study replicate those of previous experiments on gappy and glutty descriptions (Alxatib & Pelletier 2011), (Ripley 2011). However, contrary to these previous studies, the results of our experiment also suggest that gappy descriptions are more readily accepted than glutty descriptions. This observation is still to be explained, as the accounts capable to account for the acceptance of gappy and glutty descriptions treat them on a par. In addition, and after (Ruytenbeek 2013), we provided further evidence for the view according to which lexical and syntactic antonyms are used to refer to different regions on the scale they are associated with. Whereas participants accepted glutty descriptions formed with syntactic antonyms to describe borderline

cases, they refused them when they were formed with lexical antonyms. We claimed that this new observation argues against (Krifka 2007) who situates the difference between the two types of antonyms at the pragmatic level but treats both of them as semantic contradictories, and for a view in the lines of (Horn 1984) where lexical and syntactic antonyms are given different semantic representations.

## 8.4 Experimental Perspectives

Overall, the threefold choice task that (Chemla & Bott 2013) designed in order to test for truth-value gaps due to homogeneity proved efficient to elicit non-bivalent truth-value judgments due to vagueness as well as non-bivalent truth-value judgments due to presupposition. This makes it a good candidate to conduct further experimental work comparing non-bivalent truth-value judgments due to a variety of phenomena. In addition, as Chemla & Bott themselves note, homogeneity has been given accounts in terms of vagueness as well as accounts in terms of presupposition. The next logical step would therefore be to use this design to compare homogeneity with vagueness and presupposition.

Another crucial point for the joint study of vagueness and presupposition are hybrid sentences. Introspection does not provide us with clear intuitive truth-value judgments about sentences like (77-a) or (77-b) as evaluated in situations where the amplifiers are borderline loud and have never buzzed, but collecting massive data with a good design might help us to make the point clearer.

- (77) a. The amplifiers are loud<sub>vague</sub> and they have stopped<sub>presuppositional</sub> buzzing.  
 b. Either the amplifiers are loud<sub>vague</sub> or they have stopped<sub>presuppositional</sub> buzzing.

(78) exemplifies another type of hybrid sentences: sentences where the vague and the presuppositional expressions both appear in a single matrix clause. Using a visual world paradigm (as used in (Schwarz to appear) for instance) could help us to investigate the interaction between vagueness and presupposition.

- (78) Bill is tall<sub>vague</sub> too<sub>presuppositional</sub>.

One can imagine several critical contexts of evaluation for (78), but two appear to be of particular interest: a context where Bill is clearly tall but where the salient antecedent individual for *too* is borderline tall, and a context where the salient antecedent individual for *too* is borderline tall but where Bill is borderline tall. In the



former situation, the presupposition can be viewed as corresponding to a vague description of a borderline case (the salient antecedent individual is borderline tall) and the assertion can be viewed as plainly true (Bill is clearly tall); in the latter situation on the contrary, the presupposition can be viewed as clearly fulfilled (the salient antecedent individual is clearly tall) but the assertion can be viewed as a vague description of a borderline case (Bill is borderline tall). To determine what each system would predict regarding the interpretation of this sentence in these contexts, and even *whether* each system would predict anything would demand a deep discussion. ST5 however makes direct predictions: (78) would be associated with  $\mathcal{P}^1$  in the former context (it only tolerantly satisfies the presuppositional part but it strictly satisfies the assertive part) and (78) would be associated with  $\mathcal{V}$  in the latter context (it strictly satisfies the presuppositional part but it only tolerantly satisfies the assertive part). Given that  $\mathcal{P}^1$  is closer to 1 than  $\mathcal{V}$  is, we could expect that when subjects have to choose between a picture describing the former context and a picture describing the latter context, they would prefer to click the former to signal what they think (78) best describes.

Each experiment conducted in the context of this thesis eventually presented some problems with its design or with its stimuli. Even though the second experiment in the first study produced clearer results than the first experiment did, two sets of presuppositional items nonetheless appeared to group apart from the others. In the experiment on antonymy, a majority of participants failed to provide accurate answers for the control elements. These issues call for replications of the experiments with refined designs and stimuli. The replications should also investigate different *kinds* of stimuli: for instance, it would be interesting to conduct a replication of the second experiment of the first study, with presuppositional sentences containing a different expression from the aspectual verb *stop*; and it would be interesting to conduct a replication of the experiment on antonymy with morphologically complex lexical antonyms like *unhappy*.

# Appendix A

## Experiments on Vagueness and Presupposition

### A.1 Experiment 1

#### A.1.1 Instructions

In this experiment, which lasts less than 10 minutes, you will see pictures together with sentences that are used to describe them. Sometimes you might judge the description to be clearly true or clearly false. Other times, your judgment might be less clear.

If you think that the description of the picture is completely true, click on the rightmost button. If you think it is completely false, click on the leftmost one. You can give a more nuanced judgment by clicking on an intermediate button.

You don't need to remember the pictures you judge: even if some of them look extremely similar, the situations represented by the pictures are completely independent of each other.

The **Reset** button let you cancel whenever you want or restart from the very beginning by leading you back to this page. When the experiment is over, you will have to provide your AMT ID in order to validate your answers.

To start the experiment, please consider the following picture and indicate how you judge the description.

### A.1.2 Materials

The pictures on the left provided the *critical* contexts, the pictures in the middle provided the *counter-instance* contexts and the pictures on the right provided the *instance* contexts.

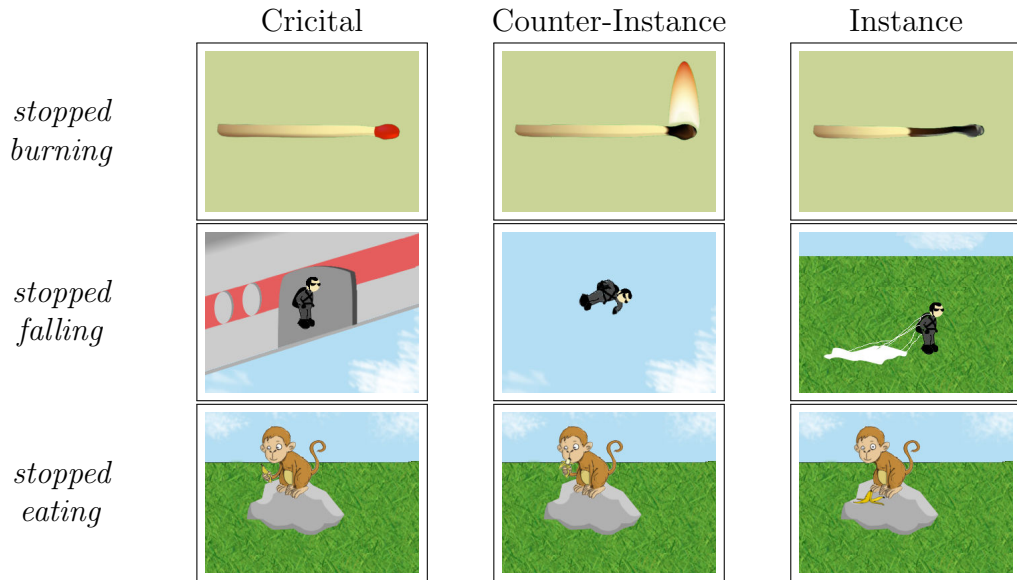


Figure A.1: The sets of pictures used with presuppositional descriptions

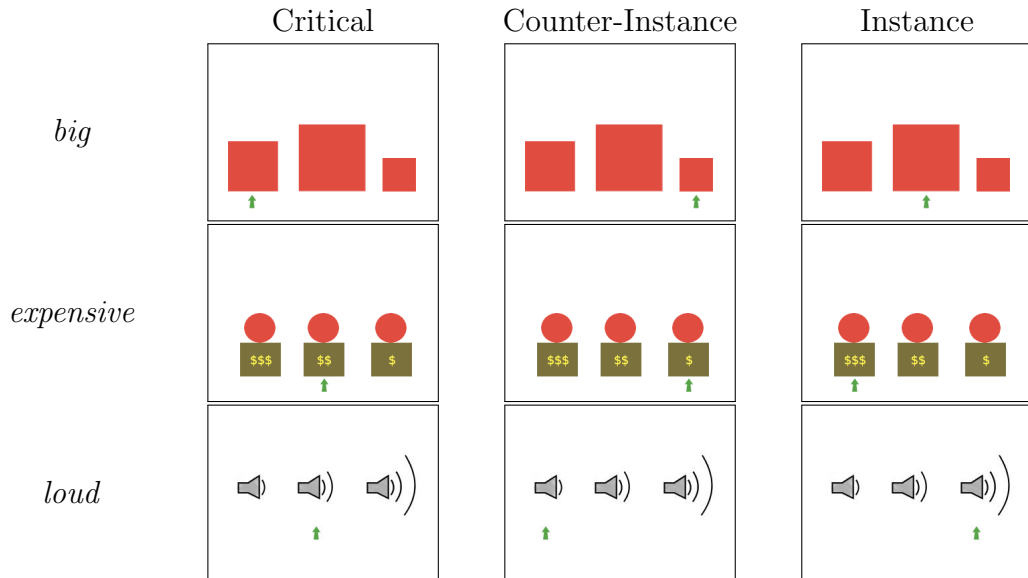
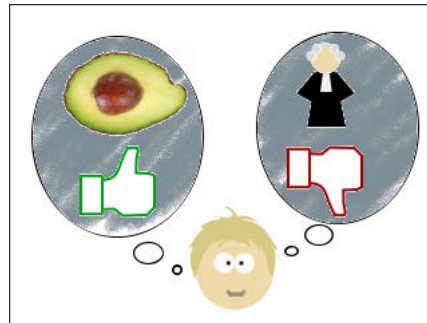


Figure A.2: The sets of pictures used with presuppositional descriptions

### A.1.3 Practice Item

Participants received no feedback after clicking.



*The boy likes avocados*

Completely false ○ ○ ○ ○ ○ Completely true

Figure A.3: Practice item

## A.2 Experiment 2

### A.2.1 Instructions

In this experiment (approximately 10 minutes), you will see several series of three pictures, together with a sentence that is used to describe the picture *with a border*.

Sometimes you might judge the description to be clearly true or clearly false. In those cases, click the corresponding button.

Other times, your judgment might be less clear: in those cases, click on *Neither*.

You will first have a training session of two trials to make sure you understand the task. Then, the experiment will start.

When the experiment is over, you will have to provide your AMT ID in order to validate your answers.

**CAUTION:** data are lost as soon as you leave or refresh this page, so please make sure to continue through the confirmation message. Please do not use the "Back" button of your browser to try to change your previous answers, as it will have the effect of leaving the page (and thus losing your data).

→ [Click here to continue](#)

### A.2.2 Materials

The materials from Experiment 1 were reused. The following sets of pictures (presented in a static sequence, see the trial items in A.2.3) were added.

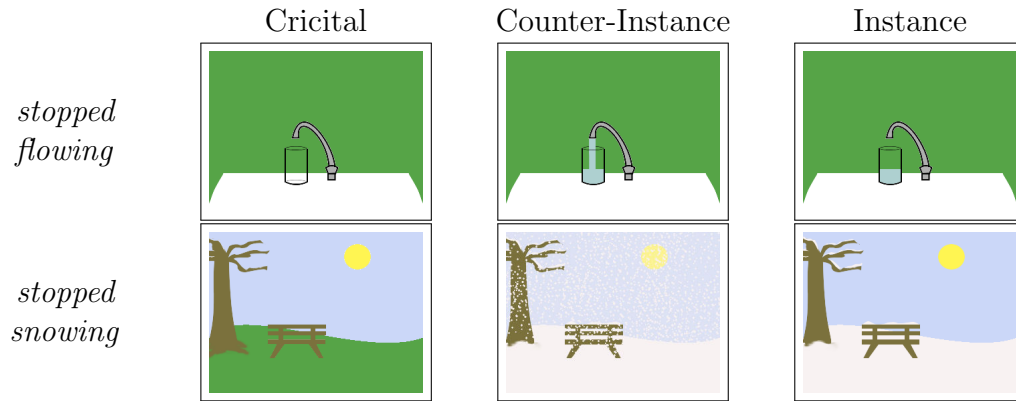


Figure A.4: The sets of additional pictures used with presuppositional descriptions

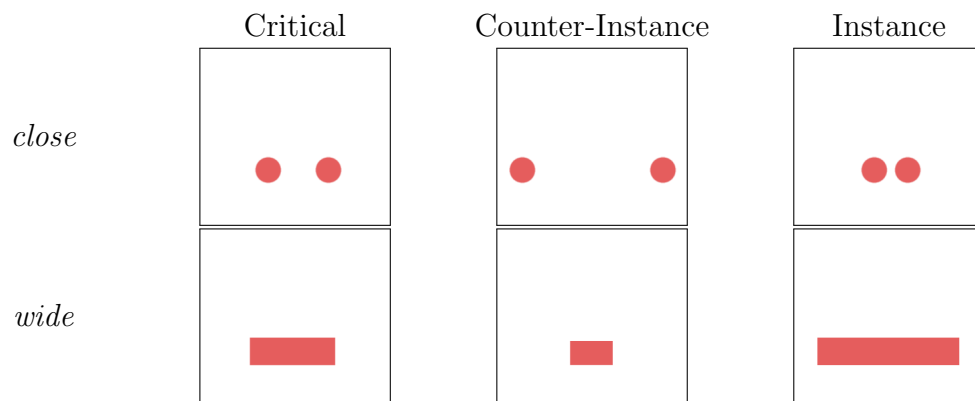
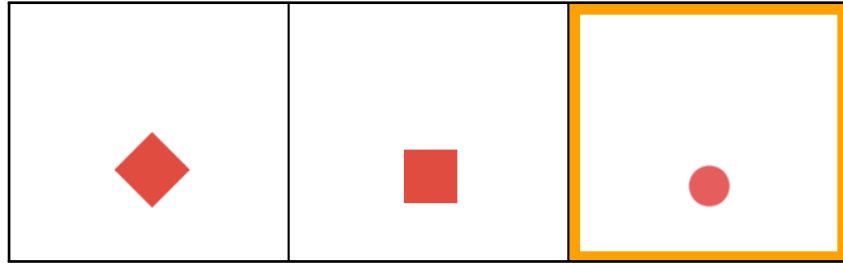


Figure A.5: The sets of additional pictures used with presuppositional descriptions

### A.2.3 Trial Items

Participants received feedback indicating whether they were wrong or correct.



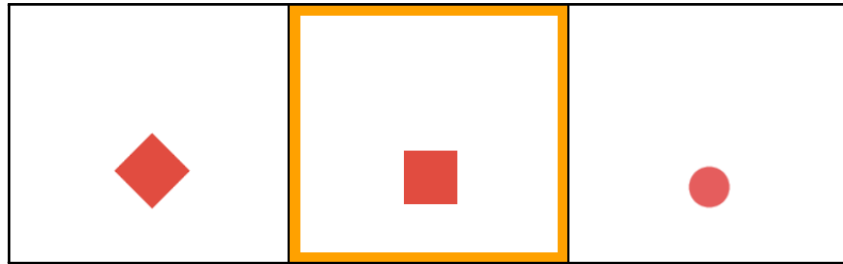
The figure is a square

Completely true

Neither

Completely false

Figure A.6: First trial item (correct answer: “Completely false”)



The figure is not a circle

Completely true

Neither

Completely false

Figure A.7: Second trial item (correct answer: “Completely true”)

# Appendix B

## Experiment on Antonyms

### B.1 Instructions

#### Answer this short survey

We invite you to participate in a research study on language production and comprehension. We will ask you to do a linguistic task such as reading sentences and giving your judgments about those sentences in specific contexts. There are two versions of this experiment: one is named “Version A” and the other is named “Version B”. If you participate to the present version, please do not participate to the other version: you will **not be paid** if you participate to **both versions** of the experiment. If you have read this form and have decided to participate in this experiment, please understand your participation is voluntary and you have the right to withdraw your consent or discontinue participation at any time.

Also, please note that in order to validate the Hit you need to complete the experiment, give your Worker ID and wait until the results are sent (this should only take a few seconds), otherwise we have no way to ensure that you participated at all.

There are no risks or benefits of any kind involved in this study. You will be paid for your participation at the posted rate.

Your participation in this study will remain confidential.

Your individual privacy will be maintained in all published and written data resulting from the study. You may print this form for you records.

If you have any comments, please feel free to contact us (procedure).

By clicking the link below, you agree to participate. Don't forget to validate your participation at the end by clicking the Submit button below.

[Survey link](#)

## B.2 Materials

The following triplets of antonyms were used both in the pilot and in the proper experiments along with the associated scale name.

Scale Name	Adjective	Lexical Antonym	Syntactic Antonym
wealth	rich	poor	not rich
height	tall	short	not tall
age	old	young	not old
weight	heavy	light	not heavy
volume	loud	soft	not loud
size	big	small	not big
temperature	hot	cold	not hot
price	expensive	cheap	not expensive

Table B.1: Antonyms used in the experiment



# Bibliography

- Abrusán, M. & Szendrői, K. (2013), ‘Experimenting with the King of France: Topics, verifiability and definite descriptions’, *Semantics and Pragmatics* **6**(10), 1–43.
- Alxatib, S. & Pelletier, F. J. (2011), ‘The psychology of vagueness: Borderline cases and contradictions’, *Mind & Language* **26**(3), 287–326.
- Beaver, D. I. (2001), *Presupposition and Assertion in Dynamic Semantics*, CSLI Publications, Stanford.
- Beaver, D. I. & Geurts, B. (2013), Presupposition, in E. N. Zalta, ed., ‘The Stanford Encyclopedia of Philosophy’, fall 2013 edn, Stanford.
- Bochvar, D. (1939), ‘Ob odnom trehznachom iscislennii i ego primeneii k analizu paradoksov klassičeskogo rassirennogo funkcional ‘nogo iscislennija’’, *Matematičeskij sbornik* **4**. English Translation: 1981, ‘On a Three-valued Calculus and Its Applications to the Analysis of the Paradoxes of the Classical Extended Functional Calculus’, *History and Philosophy of Logic* **2**, 87–112.
- Bonini, N., Osherson, D., Viale, R. & Williamson, T. (1999), ‘On the psychology of vague predicates’, *Mind and Language* **14**(4), 377–393.
- Burnett, H. (2012), *The Grammar of Tolerance: On Vagueness, Context-Sensitivity, and the Origin of Scale Structure*, PhD thesis, UCLA, Los Angeles, California, USA.
- Chemla, E. (2009), ‘Similarity: Towards a unified account of scalar implicatures, free choice permission and presupposition projection’, *Under revision for Semantics and Pragmatics* .
- Chemla, E. & Bott, L. (2013), ‘Processing presuppositions: dynamic semantics vs pragmatic enrichment’, *Language and Cognitive Processes* **28**(3), 241–260.

- Cobrerros, P., Egré, P., Ripley, D. & van Rooij, R. (forthcoming<sub>a</sub>), ‘Foreword: Three-valued logics and their applications’, *Journal of Applied Non-Classical Logics* .
- Cobrerros, P., Egré, P., Ripley, D. & van Rooij, R. (forthcoming<sub>b</sub>), ‘Vagueness, truth and permissive consequence’, *Unifying the Philosophy of Truth* .
- Cobrerros, P., Egré, P., Ripley, D. & van Rooij, R. (2012), ‘Tolerant, classical, strict’, *Journal of Philosophical Logic* **41**(2), 347–385.
- Egré, P. (to appear), ‘Vagueness: Why do we believe in tolerance?’, *Journal of Philosophical Logic* .
- Egré, P., de Gardelle, V. & Ripley, D. (2013), ‘Vagueness and order effects in color categorization’, *Journal of Logic, Language and Information* **22**(4), 391–420.
- Fine, K. (1975), ‘Vagueness, truth and logic’, *Synthese* **30**(3-4), 265–300.
- Fox, D. (2008), ‘Two short notes on Schlenker’s theory of presupposition projection’, *Theoretical Linguistics* **34**, 237–252.
- Fox, D. (2010), ““Schlenker 2008, basics” lecture notes of “Pragmatics in Linguistic Theory” course, MIT”, <http://ocw.mit.edu/courses/linguistics-and-philosophy/24-954-pragmatics-in-linguistic-theory-spring-2010/lecture-notes/>.
- Fox, D. (2012), ‘Presupposition projection from quantificational sentences: trivalence, local accommodation, and presupposition strengthening’, *MS the Hebrew University of Jerusalem. J. ROMOLI Gajewski, J* .
- George, B. R. (2008), Presupposition repairs: a static, trivalent approach to predicting projection., Master’s thesis, UCLA, Los Angeles, California, USA.
- Heim, I. (1983), On the projection problem for presuppositions, in P. Portner & B. H. Partee, eds, ‘Formal Semantics - the Essential Readings’, Vol. 2, Blackwell, pp. 249–260.
- Heim, I. & Kratzer, A. (1998), *Semantics in generative grammar*, Blackwell textbooks in linguistics, Blackwell publishers, Cambridge (Mass.), Oxford.
- Horn, L. (1984), ‘Toward a new taxonomy for pragmatic inference: Q-based and r-based implicature’, *Meaning, Form, and Use in Context* pp. 11–42.

- Horn, L. (1989), *A Natural History of Negation*, University of Chicago Press.
- Kamp, H. (1975), Two theories about adjectives, in ‘Formal semantics of Natural Language’, Cambridge University Press, pp. 123–155. Cambridge Books Online.
- Karttunen, L. (1973), ‘Presuppositions of Compound Sentences’, *Linguistic Inquiry* **4**, 167–193.
- Kleene, S. C. (1938), ‘On notation for ordinal numbers’, *Journal of Symbolic Logic* **3**(4), 150–155.
- Kleene, S. C. (1952), *Introduction to Metamathematics*, North Holland.
- Klein, E. (1980), ‘A semantics for positive and comparative adjectives’, *Linguistics and Philosophy* **4**(1), 1–45.
- Kooi, B. P. & Tamminga, A. M. (2013), ‘Three-valued logics in modal logic.’, *Studia Logica* **101**(5), 1061–1072.
- Krifka, M. (2007), ‘Negated antonyms: Creating and filling the gap’, *Presupposition and Implicature in Compositional Semantics* pp. 163–177.
- Križ, M. & Chemla, E. (2014), Two methods to find truth value gaps and their application to the projection problem of homogeneity. Ms. University of Vienna, LSCP.
- Lasersohn, P. (1993), ‘Existence presuppositions and background knowledge’, *Journal of Semantics* **10**(2), 113–122.  
**URL:** <http://jos.oxfordjournals.org/content/10/2/113.abstract>
- Lewis, D. (1982), ‘Logic for equivocators’, *Noûs* **16**(3), 431–441.
- Łukasiewicz, J. (1922), ‘On determinism’, in L. Borkowski, North-Holland, Amsterdam London eds, *Selected Works*, 1970, pp. 110–128.
- Mehlberg, H. (1958), *The Reach of Science*, University of Toronto Press.
- Muskens, R. (1999), ‘On Partial and Paraconsistent Logics’, *Notre Dame Journal of Formal Logic* **40**(3), 352–374.
- Nouwen, R. (2011), ‘Degree modifiers and monotonicity’, *Vagueness and Language Use* pp. 146–164.

- Pawlak, Z. (1997), Vagueness - a rough set view, *in* ‘Structures in Logic and Computer Science, A Selection of Essays in Honor of Andrzej Ehrenfeucht’, pp. 106–117.
- Percus, O. & Zehr, J. (2012), ‘TCS for presuppositions’, *Trivalent Logics and their applications, Proceedings of the ESSLLI 2012 Workshop* pp. 77–92.
- Peters, S. (1977), ‘A Truth-Conditional Formulation of Karttunen’s Account of Presupposition’, *Texas Linguistic Forum* 6 .
- Priest, G. (2006), *In Contradiction: A Study of the Transconsistent*, Oxford University Press.
- Ripley, D. (2011), Contradictions at the borders, *in* R. Nouwen, R. van Rooij, U. Sauerland & H.-C. Schmitz, eds, ‘Vagueness in Communication’, Vol. 6517 of *Lecture Notes in Computer Science*, Springer Berlin Heidelberg, pp. 169–188.
- Ripley, D. (2013), ‘Sorting out the sorites’, *Paraconsistency: Logic and Applications* **26**, 329–348.
- Russell, B. (1905), ‘On denoting’, *Mind* **14**, 479–493. Republished in 2005 in *Mind*, 114 (456): 873–887.
- Ruytenbeek, N. (2013), An experimental approach of negated gradable adjectives, Master’s thesis, ENS/EHESS/Paris V, Paris, France.
- Schlenker, P. (2008), ‘Be articulate: A pragmatic theory of presupposition projection.’, *Theoretical Linguistics* **34**(3), 157–212.
- Schwarz, F. (to appear), ‘False but slow: Evaluating statements with non-referring definites’, *Journal of Semantics* .
- Serchuk, P., Hargreaves, I. & Zach, R. (2011), ‘Vagueness, logic and use: Four experimental studies on vagueness’, *Mind & Language* **26**(5), 540–573.
- Sorensen, R. (2013), ‘Vagueness’, <http://plato.stanford.edu/archives/win2013/entries/vagueness/>.
- Spector, B. (2012), ‘Research seminar at the university of vienna on trivalent semantics for vagueness and presupposition’, <https://sites.google.com/site/bspectorpage/Home/teaching/>.
- Stalnaker, R. (1974), ‘Pragmatic presuppositions’, *Context and Content* pp. 47–62.

- Strawson, P. F. (1950), ‘On referring’, *Mind* **59**(235), 320–344.
- Thomason, R. H. (1972), ‘A semantic theory of sortal incorrectness’, *Journal of Philosophical Logic* **1**(2), 209–258.  
**URL:** <http://dx.doi.org/10.1007/BF00650499>
- Tye, M. (1994), ‘Sorites paradoxes and the semantics of vagueness’, *Philosophical Perspectives* **8**, 189–206.
- van Fraassen, B. C. (1966), ‘Singular terms, truth-value gaps, and free logic’, *Journal of Philosophy* **63**(17), 481–495.
- van Fraassen, B. C. (1968), ‘Presupposition, implication, and self-reference’, *Journal of Philosophy* **65**(5), 136–152.
- von Fintel, K. (2004), ‘Would you believe it? the king of france is back! presuppositions and truth-value intuitions’, *Descriptions and Beyond* **269**.
- Williamson, T. (1994), *Vagueness*, Routledge.